

INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS • 171

Quantum Field Theory and Critical Phenomena

FIFTH EDITION

JEAN ZINN-JUSTIN



OXFORD SCIENCE PUBLICATIONS

INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS

Series Editors

R. Friend	University of Cambridge
M. Rees	University of Cambridge
D. Sherrington	University of Oxford
G. Veneziano	CERN, Geneva

INTERNATIONAL SERIES OF MONOGRAPHS ON PHYSICS

172. J. Kübler: *Theory of itinerant electron magnetism, Second edition*
171. J. Zinn-Justin: *Quantum field theory and critical phenomena, Fifth edition*
170. V.Z. Kresin, S.G. Ovchinnikov, S.A. Wolf: *Superconducting state - mechanisms and materials*
169. P.T. Chruściel: *Geometry of black holes*
168. R. Wigmans: *Calorimetry - Energy measurement in particle physics, Second edition*
167. B. Mashhoon: *Nonlocal gravity*
166. N. Horing: *Quantum statistical field theory*
165. T.C. Choy: *Effective medium theory, Second edition*
164. L. Pitaevskii, S. Stringari: *Bose-Einstein condensation and superfluidity*
163. B.J. Dalton, J. Jeffers, S.M. Barnett: *Phase space methods for degenerate quantum gases*
162. W.D. McComb: *Homogeneous, isotropic turbulence - phenomenology, renormalization and statistical closures*
160. C. Barrabès, P.A. Hogan: *Advanced general relativity - gravity waves, spinning particles, and black holes*
159. W. Barford: *Electronic and optical properties of conjugated polymers, Second edition*
158. F. Strocchi: *An introduction to non-perturbative foundations of quantum field theory*
157. K.H. Bennemann, J.B. Ketterson: *Novel superfluids, Volume 2*
156. K.H. Bennemann, J.B. Ketterson: *Novel superfluids, Volume 1*
155. C. Kiefer: *Quantum gravity, Third edition*
154. L. Mestel: *Stellar magnetism, Second edition*
153. R.A. Klennm: *Layered superconductors, Volume 1*
152. E.L. Wolf: *Principles of electron tunneling spectroscopy, Second edition*
151. R. Blinc: *Advanced ferroelectricity*
150. L. Berthier, G. Biroli, J.-P. Bouchaud, W. van Saarloos, L. Cipelletti: *Dynamical heterogeneities in glasses, colloids, and granular media*
149. J. Wesson: *Tokamaks, Fourth edition*
148. H. Asada, T. Futamase, P. Hogan: *Equations of motion in general relativity*
147. A. Yaouanc, P. Dalmas de Réotier: *Muon spin rotation, relaxation, and resonance*
146. B. McCoy: *Advanced statistical mechanics*
145. M. Bordag, G.L. Klimchitskaya, U. Mohideen, V.M. Mostepanenko: *Advances in the Casimir effect*
144. T.R. Field: *Electromagnetic scattering from random media*
143. W. Götze: *Complex dynamics of glass-forming liquids - a mode-coupling theory*
142. V.M. Agranovich: *Excitations in organic solids*
141. W.T. Grandy: *Entropy and the time evolution of macroscopic systems*
140. M. Alcubierre: *Introduction to 3+1 numerical relativity*
139. A.L. Ivanov, S.G. Tikhodeev: *Problems of condensed matter physics - quantum coherence phenomena in electron-hole and coupled matter-light systems*
138. I.M. Vardavas, F.W. Taylor: *Radiation and climate*
137. A.F. Borghesani: *Ions and electrons in liquid helium*
135. V. Fortov, I. Iakubov, A. Khrapak: *Physics of strongly coupled plasma*
134. G. Fredrickson: *The equilibrium theory of inhomogeneous polymers*
133. H. Suhl: *Relaxation processes in micromagnetics*
132. J. Terning: *Modern supersymmetry*
131. M. Marinò: *Chern-Simons theory, matrix models, and topological strings*
130. V. Gantmakher: *Electrons and disorder in solids*
129. W. Barford: *Electronic and optical properties of conjugated polymers*
128. R.E. Raab, O.L. de Lange: *Multipole theory in electromagnetism*
127. A. Larkin, A. Varlamov: *Theory of fluctuations in superconductors*
126. P. Goldbart, N. Goldenfeld, D. Sherrington: *Stealing the gold*
125. S. Atzeni, J. Meyer-ter-Vehn: *The physics of inertial fusion*
123. T. Fujimoto: *Plasma spectroscopy*
122. K. Fujikawa, H. Suzuki: *Path integrals and quantum anomalies*
121. T. Giannì: *Quantum physics in one dimension*
120. M. Warner, E. Terentjev: *Liquid crystal elastomers*
119. L. Jacak, P. Sitko, K. Wieczorek, A. Wojs: *Quantum Hall systems*
117. G. Volovik: *The Universe in a helium droplet*
116. L. Pitaevskii, S. Stringari: *Bose-Einstein condensation*
115. G. Dissertori, I.G. Knowles, M. Schmelling: *Quantum chromodynamics*
114. B. DeWitt: *The global approach to quantum field theory*
112. R.M. Mazo: *Brownian motion - fluctuations, dynamics, and applications*
111. H. Nishimori: *Statistical physics of spin glasses and information processing - an introduction*
110. N.B. Kopnin: *Theory of nonequilibrium superconductivity*
109. A. Aharoni: *Introduction to the theory of ferromagnetism, Second edition*
108. R. Dobbs: *Helium three*

105. Y. Kuramoto, Y. Kitaoka: *Dynamics of heavy electrons*
104. D. Bardin, G. Passarino: *The Standard Model in the making*
103. G.C. Branco, L. Lavoura, J.P. Silva: *CP Violation*
101. H. Araki: *Mathematical theory of quantum fields*
100. L.M. Pismen: *Vortices in nonlinear fields*
99. L. Mestel: *Stellar magnetism*
98. K.H. Bennemann: *Nonlinear optics in metals*
94. S. Chikazumi: *Physics of ferromagnetism*
91. R.A. Bertlmann: *Anomalies in quantum field theory*
90. P.K. Gosh: *Ion traps*
87. P.S. Joshi: *Global aspects in gravitation and cosmology*
86. E.R. Pike, S. Sarkar: *The quantum theory of radiation*
83. P.G. de Gennes, J. Prost: *The physics of liquid crystals*
73. M. Doi, S.F. Edwards: *The theory of polymer dynamics*
69. S. Chandrasekhar: *The mathematical theory of black holes*
51. C. Møller: *The theory of relativity*
46. H.E. Stanley: *Introduction to phase transitions and critical phenomena*
32. A. Abragam: *Principles of nuclear magnetism*
27. P.A.M. Dirac: *Principles of quantum mechanics*
23. R.E. Peierls: *Quantum theory of solids*

QUANTUM FIELD THEORY AND CRITICAL PHENOMENA
FIFTH EDITION

JEAN ZINN-JUSTIN

IRFU/CEA, Paris-Saclay University
and
French Academy of Sciences

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Jean Zinn-Justin 2021

The moral rights of the author have been asserted

First Edition published in 1989
Second Edition published in 1993
Third Edition published in 1996
Fourth Edition published in 2002
Fifth Edition published in 2021

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021931817

ISBN 978-0-19-883462-5

DOI: 10.1093/oso/9780198834625.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

To Nicole

Preface

Introduced as a quantum extension of Maxwell's classical theory, quantum electrodynamics (QED) has been the first example of a local quantum field theory (QFT). Eventually, QFT has become the framework for the discussion of all fundamental interactions at the microscopic scale except, possibly, gravity. More surprisingly, it has also provided a framework for the understanding of second order phase transitions in statistical mechanics. In fact, as hopefully this work will illustrate, QFT is the natural framework for the discussion of most systems involving an infinite number of degrees of freedom with local couplings. These systems range from cold Bose gases at the condensation temperature (about ten nanokelvin) to conventional phase transitions (from a few degrees to several hundred) and high energy particle physics up to TeVs, altogether more than *twenty orders of magnitude* in the energy scale.

Therefore, although excellent textbooks about QFT had already been published, I thought, many years ago, that it might not be completely worthless to present a work in which the strong formal relations between particle physics and the theory of critical phenomena are systematically emphasized. This option explains some of the choices made in the presentation. A formulation in terms of field integrals is adopted to study the properties of QFT. Less important, perhaps, in general the space-time metric is chosen Euclidean, as is natural for statistical mechanics, and in particle physics often convenient for perturbative calculations, and necessary for numerical simulations. The language of partition and correlation functions is used throughout, even in applications of QFT to particle physics. Renormalization and renormalization group (RG) properties are systematically discussed, whereas limited space is devoted to scattering theory. Only formal aspects of QED are considered, since excellent textbooks cover this subject extensively.

For what follows, note that, in a deep quantum relativistic context, one can set $\hbar = c = 1$, and energies are then proportional to momenta and masses, and inverse of distances.

In QFT, the basic analytic tool to calculate physical quantities is an expansion in powers of the interactions. The initial (or bare) Lagrangian of QED generates an expansion in terms of the *bare fine-structure constant* $\alpha_0 = e_0^2/4\pi\hbar c$. In a straightforward perturbative calculation, one discovers that all physical quantities are infinite, the *locality* of QED generating *short-distance singularities* (one speaks about ultraviolet (UV) divergences).

This situation has to be contrasted with what happens in classical or non-relativistic quantum mechanics (QM); there, the replacement of macroscopic by point-like objects leads, in general, to no mathematical inconsistencies, and is often a very good approximation: the absence of this property would indeed have made progress in physics quite difficult. To summarize: in the latter theories, *phenomena at very different scales, to a good approximation, decouple*. Most surprisingly, this is no longer the case in QFT.

In QED, a remedy to the infinity problem was found empirically: one first *regularizes* QED (*i.e.* one renders the perturbative expansion finite) by artificially modifying the theory at short distance, or equivalently at large momentum, at a scale characterized by

a large momentum cut-off Λ (in general, introducing non-physical short-distance properties). Inspired by methods of condensed matter physics, one then re-expresses all physical quantities in terms of the measured fine structure constant and the physical masses of particles, in place of the original (bare) parameters of the Lagrangian. After this change of parametrization, the cut-off is removed, and somewhat miraculously, order by order in perturbation theory, all other physical quantities have a finite limit. Moreover, the limit is independent of the precise form of the regularization. This strange method, called *renormalization*, did soon find an experimental confirmation: it led to predictions agreeing with increasingly impressive precision with experiments.

Therefore, it became then natural to search for other *renormalizable* QFTs, to describe all interactions. This led to another major achievement: a renormalizable QFT for all three, strong, weak, and electromagnetic interactions. The so-called Standard Model (SM), whose formal structure was proposed more than forty years ago, completely describes physics at the microscopic scale, and has been confirmed in 2012, in a spectacular way, by the discovery, at the Large Hadron Collider of the European Center for Nuclear Research (CERN), of the last missing particle of the model, the Higgs scalar boson.

The impressive success of a strategy based on looking for renormalizable QFTs, that led to the SM, then slowly promoted *renormalizability* as a kind of additional law of nature. In particular, once the SM of weak, electromagnetic and strong interactions was established, much effort was devoted to cast gravity in the same framework. Despite ingenious attempts, no renormalizable form of quantum gravity has been found yet.

In a massless renormalizable QFT, it is necessary to introduce a reference physical energy scale at which the physical coupling constants are defined. It was realized early on, first as a mathematical curiosity, that an RG could be associated with a change of the reference scale at constant physical properties. The RG describes how the physical (or effective) coupling constants vary with the reference scale.

Eventually, it was realized that this property could also be used to discuss the short-distance properties of some physical processes. In *asymptotically free* QFTs (where the free QFT is an ultraviolet (UV) RG fixed point), these effective couplings become small at large-Euclidean momenta and, therefore, perturbation theory, improved by RG, can be used. In particle physics, the theory of strong interactions, based on $SU(3)$ gauge symmetry, shares this property. Later, Weinberg argued that the existence of UV RG fixed points (like in non-Abelian gauge theories, or non-linear σ models), that is, the existence of limits for the effective short-distance couplings, was a necessary condition for the consistency of a QFT on all scales. However, most of the field theories proposed to describe strong, electromagnetic and weak interactions are not asymptotically free.

Of course, the existence of other non-trivial fixed points cannot be established in the framework of perturbation theory. However, many numerical simulations of field theories on the lattice, which make non-perturbative explorations possible, have failed to discover non-trivial fixed points. Therefore, presumably, the present SM, which describes so precisely particle physics at present scale, is not consistent on all scales, and has to be modified at shorter distance.

This also suggests that the property of renormalizability has a different origin.

Somewhat surprisingly, in statistical physics, QFT has also become an essential tool for the understanding of the critical behaviour of a large class of second-order phase transitions with short-range interactions. Near the critical temperature, cooperative phenomena generate a large scale, associated with the so-called *correlation length*. Moreover, the large-scale properties of the system become independent of most of the details of the microscopic dynamics. First attempts to explain these properties were based on

usual ideas: a description only involving macroscopic degrees of freedom adapted to the scale of the correlation length. Such a description naturally emerges in simple approximations like *mean-field theory*. It is consistent with the general probabilistic idea that averages over a large number of independent stochastic variables obey a Gaussian distribution. The corresponding general ideas were summarized in Landau's theory of critical phenomena. However, it slowly became clear that the predictions of such a theory were too universal, conflicting with numerical calculations of critical exponents, experimental data and exact results in two dimensions. These results supported the concept of a more restricted *universality*: broad classes of systems have indeed the same large distance properties, but, unlike in mean-field theory, these properties seemed to depend on a small number of qualitative features, like dimension of space, number of components of the order parameter, symmetries, and so on. Actually, an analysis of the leading corrections to mean-field or Gaussian approximations indeed reveals that, at least in low-space dimensions, the short distance never completely decouple, a most unusual situation.

To explain this remarkable phenomenon, that is, that large distance properties of second order phase transitions are, to a large extent, short-distance insensitive, although degrees of freedom on all scales remain coupled, Wilson, partially inspired by some prior attempts of Kadanoff, introduced the RG idea: starting from a microscopic Hamiltonian, one integrates out, recursively, the degrees of freedom corresponding to short-distance fluctuations, and generates a scale-dependent effective Hamiltonian. Universality then relies upon the existence of IR fixed points in Hamiltonian space. One of the spectacular implications is that the universal properties of a large class of critical phenomena can be accurately predicted by the same QFT methods that had been invented for particle physics. At leading order in the critical domain, the physics of the fixed-point Hamiltonian can be reproduced by renormalizable (or super-renormalizable) QFTs.

Predictions obtained from a RG analysis of simple field theories like the $(\phi^2)^2$ QFT have been successfully compared to experiments as well as numerical data from lattice models. The same QFT methods have been shown to describe vastly different physical systems at criticality, like ferromagnets, liquid–vapour, binary mixtures, superfluid helium and, even more surprisingly, statistical properties of polymers.

If QFT has led to an understanding of the concept of universality and made the calculation of many universal physical quantities possible, conversely, critical phenomena have shed a new light on the origin of renormalizable QFTs and of the renormalization process in particle physics.

Let us describe again, in general terms, the renormalization method. Starting from an initial Lagrangian, assumed to be *renormalizable*, expressed in terms of *bare parameters*, one generates an expansion in powers of the interactions. The *perturbative expansion* is then plagued by UV divergences. To render the perturbative expansion finite, one introduces a large-momentum cut-off Λ , which generates an artificial, and somewhat arbitrary short-distance structure, with non-physical properties. One then calculates quantities at the physical scale and *tunes the bare parameters* as functions of the cut-off, in such way that observables at the physical scale have a finite infinite Λ limit. Since the bare parameters have generally no finite infinite Λ limit, this led some physicists to the paradoxical conclusion that the Lagrangian in relativistic QFT is non-physical. The insistence for taking the infinite cut-off limit was motivated by the wish to construct a renormalized QFT physically consistent on all scales. Formalisms were developed (like the Bogoliubov–Parasiuk–Hepp–Zimmermann (BPHZ) formalism) to generate directly a renormalized perturbation theory, in which the Lagrangian was reduced to a device for generating Feynman rules. However, another interpretation of the bare QFT and

the renormalization procedure, directly inspired from the functional RG as applied to critical phenomena by Wilson, has gained strength over the years. At very short distance (the Planck scale ?), the familiar notion of local QFT loses its meaning. However, the necessary non-local effects, which render the more fundamental theory finite, are limited to this microscopic scale (the equivalent of the condition of short-range forces in statistical systems). Dynamical effects, of a nature which at present can only be guessed, generate at distances large compared to the microscopic scale (equivalently, at lower energies) physics associated with the appearance of very low mass particles (compared to the Planck mass, for example, all known particles are almost massless). At large distance, physics can be described by an *effective local QFT*, whose action is a linear combination of all local monomials consistent with symmetries, and other possible requirements. With increasing distance, at least in a mean-field like approximation, all interactions scale as dimensional analysis (or power counting) indicates. Eventually, non-renormalizable interactions become very small, dimensionless renormalizable interactions survive, and unprotected terms, with positive dimension, like mass terms, diverge.

Actually, this picture has to be corrected, because the UV divergences and the necessity of introducing a large-momentum cut-off (a substitute for the unknown microscopic physics) show that microscopic scale and the physical scale do not completely decouple. The flow of interactions has to be described by a RG that interpolates between the microscopic scale and the physical scale. Still, as we shall argue, the flow of renormalizable interactions is only logarithmic, while the other terms have a power-law behaviour. Thus, the mean-field analysis of the hierarchy of interactions remains qualitatively correct.

This scheme has the following consequences: it explains the *emergence of renormalizable QFTs*. The role of renormalization theory is to prove, order by order in perturbation theory, that physics at the physical scale is, to a large extent, but not completely, short-distance insensitive (and thus insensitive to the precise cut-off implementation), but in a more subtle way as in classical physics. However, *the initial (bare) parameters have to be considered as being given*, and one has, therefore, to prove, beyond perturbation theory, that the change of parametrization, assumed in the renormalization theory, from bare to physical effective parameters, is possible. In particular, this leads to the *triviality issue* for IR-free theories (like QED): the physical charge decreases logarithmically with the cut-off and, therefore, the *cut-off cannot be sent to infinity*. The corresponding coupling constants are expected to be small (this is consistent with the small value of α). Moreover, most QFTs cannot be made consistent on all scales. The possibility of very small non-renormalizable interactions, since proportional to the inverse of the power of cut-off, has to be considered. General Relativity may be of this nature.

The origin of the Higgs boson, which is involved in most of the parameters of the SM, and the necessary *fine tuning* of its mass term in the Lagrangian, become essential physics issues. In contrast to critical phenomena in which a control parameter, like the temperature, can be adjusted to make the correlation length large, in particle physics the existence of small mass particles has to be explained by general properties of the unknown fundamental theory. This is the famous *hierarchy* problem. Spontaneous breaking of a continuous symmetry, gauge principle, and chiral invariance (but whose natural implementation seems to require additional space dimensions) are the known mechanisms which generate massless particles. Supersymmetry can be helpful to deal with scalar bosons. At present, the set of general conditions to be imposed on any fundamental theory, that is, in the language of critical phenomena the complete description of the universality class of particle physics has not been formulated. This is also a fundamental problem of the SM.

On the other hand, since the large distance physics is, to a large extent, short-distance insensitive, the real nature of the fundamental theory may remain, in the foreseeable future, elusive, in the same way as a precise knowledge of the critical exponents of the liquid–vapour phase transition gives limited information about real interactions in water.

This work, which does not claim to shed any light on these difficult problems, simply tries to describe particle physics and critical phenomena in statistical mechanics in a unified framework.

Chapters 1–7 deal with functional integrals, perturbation theory, functional methods and discuss general properties of scalar boson QFTs. Chapters 8–13 provide an introduction to renormalization theory with the simple example of the ϕ^4 QFT in four dimensions, and RG equations are derived. Composite operators and the short-distance expansion are discussed. Relativistic fermions are described. Renormalization properties of theories with symmetries are studied, and specific applications to particle physics are emphasized.

Chapters 14–19 are devoted to critical phenomena in macroscopic phase transitions: general properties, mean-field approximation, and mainly applications of QFT methods and RG, with the calculation of universal quantities, in particular, with Wilson–Fisher ε expansion, large N techniques and non-linear σ -model for $O(N)$ -symmetric models.

With Chapters 20–26, we return to particle physics. Chapter 20 deals with spontaneous fermion mass generation. We then discuss gauge theories, Abelian and non-Abelian. In particular, Chapter 22 briefly describes the SM of particle physics. Chapter 26 introduces Becchi–Rouet–Stora–Tyutin (BRST) symmetry, the Zinn–Justin (ZJ) equation and the proof of the renormalizability of gauge theories. Chapter 27 contains a short introduction to supersymmetry. Chapter 28 gathers the few elements of classical and quantum gravity needed in the work.

In Chapters 29–31, we focus on two-dimensional field theories, of relevance both for particle and statistical physics. Chapter 29 is devoted to QFTs defined on homogeneous spaces, and Chapters 30 and 31 describe exactly-solvable two-dimensional QFTs.

Chapter 32 provides an introduction to finite-size effects, and Chapter 33 to finite temperature relativistic QFT. Chapters 34–36 deal with stochastic evolution equations, and their application to critical dynamics in phase transitions. Chapters 37–42 describe the role of instantons in QM and QFT, the application of instanton calculus to the analysis of large-order behaviour of perturbation theory, and the problem of summation of the perturbative expansion. In particular, Chapter 41 applies this information to the evaluation of critical exponents and several other universal quantities.

I am fully aware that this work is largely incomplete. My ignorance or lack of understanding of many important topics is of course mostly responsible for this weakness. A lack of space has also forced me to remove an introduction to large random matrices, and prevented me from adding some other topics. Anyway, I believe that a complete survey of QFT and its applications is beyond the scope of a single physicist.

This work incorporates notes for lectures delivered in numerous summer schools, most notably, Cargèse 1973, Bonn 1974, Karpacz 1975, Basko Polje 1976, and Les Houches 1982, as well as for graduate lectures in universities like Princeton, Louvain-la-Neuve, Berlin, Lausanne, Cambridge (Harvard), Ecole Normale Supérieure, Paris 7, and so on.

Conversely, since some relevant material that I have gathered over the years can no longer find a place in this work, some elements have been published in four reviews (in the form of Physics Reports) and in three companion volumes, Refs. [6, 64] and J. Zinn-Justin, *From random walks to random matrices* (Oxford Univ. Press 2019).

Finally, comments or corrections are most welcome, and can be sent to the email address: jean.zinn-justin@cea.fr.

Acknowledgements

It is impossible to list all the physicists from whom I have benefited in my long career and whose influence can, therefore, be felt in one form or another in this work. My masters, M. Froissart and D. Bessis, guided my first steps in physics. E. Brézin and J.C. Le Guillou have collaborated with me for more than fifteen years, and without them, obviously, this work would never have been produced. I also think with a deep emotion of B.W. Lee: the year I spent working with him at Stony-Brook was one of the most exciting of my life as a physicist.

Wilson's RG ideas are a major source of inspiration for this work. S. Coleman, A.A. Slavnov, R. Stora, K. Symanzik, A.N. Vasilev, among others, played an important role in my understanding of several aspects of physics through their articles and lecture notes, as well as through private discussions. T.D. Lee and C.N. Yang have consistently honoured me with their friendship and hospitality in their institutions. Their deep remarks have been precious to me.

I have learned much from lectures in several summer schools, most notably, C.G. Callan, L.D. Faddeev, and D. Gross in Les Houches school 1975.

Several colleagues agreed to read part of the various editions before publication, and I have benefited from their criticisms, remarks and wisdom, E. Brézin, R. Stora, C. Bervillier, R. Guida, M. Moshe, O. Napol, A.N. Vasilev, P. Zinn-Justin, and J.-B. Zuber.

More generally, my colleagues of the Saclay theory group, with whom I have had so many discussions, in particular C. de Dominicis, E. Iancu, and C. Itzykson, have directly influenced this work. Finally, I also wish to thank all other colleagues with whom I have collaborated over the years and, more specifically, P. Ginsparg, R. Guida, S. Hikami, U.D. Jentschura, M. Moshe, and G. Parisi.

The many lectures I have attended in Les Houches during nine summers have provided me with additional inspiration, and a stay at the Massachusetts Institute of Technology (MIT), where lecture notes concerning finite temperature field theory were prepared, is gratefully acknowledged.

S. Zaffanella and M. Porneuf were most helpful in the preparation of the first edition.
All deserve my deepest gratitude.

Fully revised for the 5th edition, Paris-Saclay, 6 February 2021

Some general references for the whole work

In addition to the works explicitly quoted in the text, a number of textbooks or reviews have been a direct source of inspiration:

S. Weinberg, *The Quantum Theory of Fields*, 2 volumes, (Cambridge Univ. Press 1995, 1996);

A.N. Vasiliev, *Functional Methods in Quantum Field Theory and Statistical Physics*, (St Petersburg 1976), English translation (Gordon and Breach, Amsterdam 1998);

L.D. Faddeev and A.A. Slavnov, *Gauge Fields: Introduction to Quantum Field Theory* (Benjamin, Reading, MA 1980);

T.D. Lee, *Particle Physics and Introduction to Field Theory* (Harwood Academic, New York 1981);

A.M. Polyakov, *Gauge Fields and Strings* (Harwood Academic, New York 1988);

J. Drouffe and C. Itzykson, *Théorie Statistique des Champs* (InterEditions 1989), English version: *Statistical Field Theory*, (Cambridge Univ. Press 1989);

Construction of field theories from a more rigorous point of view is discussed in

R.F. Streater and A.S. Wightman, *PCT, Spin & Statistics and All That* (Benjamin, New York 1964);

G. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View* (Springer-Verlag, Berlin 1981).

Critical Phenomena, Random Systems, Gauge Theories, Proceedings of Les Houches Summer School 1984, K. Osterwalder and R. Stora eds. (Elsevier, Amsterdam 1986).

Contents

1 Gaussian integrals. Algebraic preliminaries	1
1.1 Gaussian integrals: Wick's theorem	1
1.2 Perturbative expansion. Connected contributions	3
1.3 The steepest descent method	4
1.4 Complex structures and Gaussian integrals	5
1.5 Grassmann algebras. Differential forms	6
1.6 Differentiation and integration in Grassmann algebras	8
1.7 Gaussian integrals with Grassmann variables	13
1.8 Legendre transformation	16
2 Euclidean path integrals and quantum mechanics (QM)	18
2.1 Markovian evolution and locality	19
2.2 Statistical operator: Path integral representation	20
2.3 Explicit evaluation of a path integral: The harmonic oscillator	24
2.4 Partition function: Classical and quantum statistical physics	25
2.5 Correlation functions. Generating functional	27
2.6 Harmonic oscillator. Correlation functions and Wick's theorem	30
2.7 Perturbed harmonic oscillator	33
2.8 Semi-classical expansion	35
A2 Additional remarks	38
A2.1 A useful relation between determinant and trace	38
A2.2 The two-point function: An integral representation	39
A2.3 Time-ordered products of operators	40
3 Quantum mechanics (QM): Path integrals in phase space	42
3.1 General Hamiltonians: Phase-space path integral	42
3.2 The harmonic oscillator. Perturbative expansion	45
3.3 Hamiltonians quadratic in momentum variables	47
3.4 The spectrum of the $O(2)$ -symmetric rigid rotator	51
3.5 The spectrum of the $O(N)$ -symmetric rigid rotator	52
A3 Quantization. Topological actions: Quantum spins, magnetic monopoles	56
A3.1 Symplectic form and quantization: General remarks	56
A3.2 Classical equations of motion and quantization	58
A3.3 Topological actions	60
4 Quantum statistical physics: Functional integration formalism	64
4.1 One-dimensional QM: Holomorphic representation	64
4.2 Holomorphic path integral	67

4.3 Several degrees of freedom. Boson interpretation	71
4.4 The Bose gas. Field integral representation	72
4.5 Fermion representation and complex Grassmann algebras	80
4.6 Path integrals with fermions	83
4.7 The Fermi gas. Field integral representation	87
5 Quantum evolution: From particles to non-relativistic fields	90
5.1 Time evolution and scattering matrix in quantum mechanics (QM)	90
5.2 Path integral and S -matrix: Perturbation theory	92
5.3 Path integral and S -matrix: Semi-classical expansions	95
5.4 S -matrix and holomorphic formalism	99
5.5 The Bose gas: Evolution operator	102
5.6 Fermi gas: Evolution operator	103
A5 Perturbation theory in the operator formalism	104
6 The neutral relativistic scalar field	105
6.1 The relativistic scalar field	105
6.2 Quantum evolution and the S -matrix	110
6.3 S -matrix and field asymptotic conditions	112
6.4 The non-relativistic limit: The ϕ^4 QFT	116
6.5 Quantum statistical physics	118
6.6 Källen–Lehmann representation and field renormalization	122
7 Perturbative quantum field theory (QFT): Algebraic methods	125
7.1 Generating functionals of correlation functions	126
7.2 Perturbative expansion. Wick’s theorem and Feynman diagrams	127
7.3 Connected correlation functions: Generating functional	129
7.4 The example of the ϕ^4 QFT	131
7.5 Algebraic properties of field integrals. Quantum field equations	133
7.6 Connected correlation functions. Cluster properties	139
7.7 Legendre transformation. Vertex functions	141
7.8 Momentum representation	144
7.9 Loop or semi-classical expansion	146
7.10 Vertex functions: One-line irreducibility	151
7.11 Statistical and quantum interpretation of the vertex functional	152
A7 Additional results and methods	155
A7.1 Generating functional at two loops	155
A7.2 The background field method	156
A7.3 Connected Feynman diagrams: Cluster properties	157
8 Ultraviolet divergences: Effective field theory (EFT)	160
8.1 Gaussian expectation values and divergences: The scalar field	161
8.2 Divergences of Feynman diagrams: Power counting	162
8.3 Classification of interactions in scalar quantum field theories	164
8.4 Momentum regularization	166
8.5 Example: The $\phi_{d=6}^3$ field theory at one-loop order	169
8.6 Operator insertions: Generating functionals, power counting	173
8.7 Lattice regularization. Classical statistical physics	175
8.8 Effective QFT. The fine-tuning problem	176
8.9 The emergence of renormalizable field theories	179
A8 Technical details	181

A8.1	Schwinger's proper-time representation	181
A8.2	Regularization and one-loop divergences	181
A8.3	More general momentum regularizations	184
9	Introduction to renormalization theory and renormalization group (RG)	185
9.1	Power counting. Dimensional analysis	186
9.2	Regularization. Bare and renormalized QFT	187
9.3	One-loop divergences	191
9.4	Divergences beyond one-loop: Skeleton diagrams	194
9.5	Callan–Symanzik equations	196
9.6	Inductive proof of renormalizability	198
9.7	The $\langle\phi^2\phi^2\rangle$ vertex function	203
9.8	The renormalized action: General construction	204
9.9	The massless theory	204
9.10	Homogeneous RG equations: Massive QFT	208
9.11	EFT and RG	210
9.12	Solution of bare RG equations: The triviality issue	212
A9	Functional RG equations. Super-renormalizable QFTs. Normal order	214
A9.1	Large-momentum mode integration and functional RG equations	214
A9.2	The ϕ^4 QFT in three dimensions: Divergences	216
A9.3	Super-renormalizable scalar QFTs in two dimensions: Normal order	218
10	Dimensional continuation, regularization, minimal subtraction (MS). Renormalization group (RG) functions	220
10.1	Dimensional continuation and dimensional regularization	220
10.2	RG functions	224
10.3	The structure of renormalization constants	226
10.4	MS scheme	227
10.5	RG functions at two-loop order: The ϕ^4 QFT	230
10.6	Generalization to N -component fields	235
A10	Feynman parametrization	239
11	Renormalization of local polynomials. Short-distance expansion (SDE)	240
11.1	Renormalization of operator insertions	240
11.2	Quantum field equations	245
11.3	Short-distance expansion of operator products	248
11.4	Large-momentum expansion of the SDE coefficients: CS equations	253
11.5	SDE beyond leading order. General operator product	255
11.6	Light-cone expansion of operator products	256
12	Relativistic fermions: Introduction	258
12.1	Massive Dirac fermions	258
12.2	Self-interacting massive fermions: Non-relativistic limit	263
12.3	Free Euclidean relativistic fermions	265
12.4	Partition function. Correlations	269
12.5	Generating functionals	270
12.6	Connection between spin and statistics	272
12.7	Divergences and momentum cut-off	274
12.8	Dimensional regularization	276
12.9	Lattice fermions and the doubling problem	276
A12	Euclidean fermions, spin group and γ matrices	280

A12.1 Spin group. Dirac γ matrices	280
A12.2 The example of dimension 4	288
A12.3 The Fierz transformation	289
A12.4 Traces of products of γ matrices	290
13 Symmetries, chiral symmetry breaking, and renormalization	292
13.1 Lie groups and algebras: Preliminaries	293
13.2 Linear global symmetries and WT identities	295
13.3 Linear symmetry breaking	298
13.4 Spontaneous symmetry breaking	301
13.5 Chiral symmetry breaking in strong interactions: Effective theory	304
13.6 The linear σ -model	306
13.7 WT identities	310
13.8 Quadratic symmetry breaking	313
A13 Currents and Noether's theorem	317
A13.1 Currents in classical-field theory	317
A13.2 The energy-momentum tensor	318
A13.3 Euclidean theory: Dilatation and conformal invariance	320
A13.4 QFT: Currents and correlation functions	322
A13.5 Energy-momentum tensor and QFT	323
14 Critical phenomena: General considerations. Mean-field theory (MFT)	324
14.1 The transfer matrix	325
14.2 The infinite transverse size limit: Ising-like systems	328
14.3 Continuous symmetries	331
14.4 Mean-field approximation	332
14.5 Universality within mean-field approximation	337
14.6 Beyond the mean-field approximation	342
14.7 Power counting and the role of dimension 4	345
14.8 Tricritical points	346
A14 Additional considerations	347
A14.1 High-temperature expansion	347
A14.2 Mean-field approximation: General formalism	348
A14.3 Mean-field expansion	351
A14.4 High-, low-temperature, and mean-field expansions	352
A14.5 Quenched averages	354
15 The renormalization group (RG) approach: The critical theory near four dimensions	357
15.1 RG: The general idea	358
15.2 The Gaussian fixed point	363
15.3 Critical behaviour: The effective ϕ^4 field theory	366
15.4 RG equations near four dimensions	368
15.5 Solution of the RG equations: The ϵ -expansion	370
15.6 Critical correlation functions with $\phi^2(x)$ insertions	372
15.7 The $O(N)$ -symmetric $(\phi^2)^2$ field theory	376
15.8 Statistical properties of long self-repelling chains	377
15.9 Liquid-vapour phase transition and ϕ^4 field theory	382
15.10 Superfluid transition	387
16 Critical domain: Universality, ϵ-expansion	391

16.1 Strong scaling above T_c : The renormalized theory	392
16.2 Critical domain: Homogeneous RG equations	396
16.3 Scaling properties above T_c	396
16.4 Correlation functions with ϕ^2 insertions	399
16.5 Scaling properties in a magnetic field and below T_c	400
16.6 The N -vector model	403
16.7 The general N -vector model	405
16.8 Asymptotic expansion of the two-point function	410
16.9 Some universal quantities as ϵ expansions	412
16.10 Conformal bootstrap	420
17 Critical phenomena: Corrections to scaling behaviour	421
17.1 Corrections to scaling: Generic dimensions	421
17.2 Logarithmic corrections at the upper-critical dimension	423
17.3 Irrelevant operators and the question of universality	426
17.4 Corrections coming from irrelevant operators. Improved action	428
17.5 Application: Uniaxial systems with strong dipolar forces	431
18 O(N)-symmetric vector models for N large	436
18.1 The large N action	436
18.2 Large N limit: Saddle point equations, critical domain	438
18.3 Renormalization group (RG) functions and leading corrections to scaling	445
18.4 Small-coupling constant, large-momentum expansions for $d < 4$	447
18.5 Dimension 4: Triviality issue for N large	448
18.6 The $(\phi^2)^2$ field theory and the non-linear σ -model for N large	449
18.7 The $1/N$ -expansion: An alternative field theory	453
18.8 Explicit calculations	455
19 The non-linear σ-model near two dimensions: Phase structure	458
19.1 The non-linear σ -model: Definition	459
19.2 Perturbation theory. Power counting	461
19.3 IR divergences	463
19.4 UV regularization	464
19.5 WT identities and master equation	466
19.6 Renormalization	469
19.7 The renormalized action: Solution to the master equation	471
19.8 Renormalization of local functionals	474
19.9 A linear representation	475
19.10 $(\phi^2)^2$ field theory in the ordered phase and non-linear σ -model	476
19.11 Renormalization, RG equations	479
19.12 RG equations: Solutions (magnetic terminology)	480
19.13 Results beyond one-loop order	486
19.14 The dimension 2: Asymptotic freedom	488
20 Gross–Neveu–Yukawa and Gross–Neveu models	489
20.1 The GNY model: Spontaneous mass generation	489
20.2 RG equations near four dimensions	494
20.3 The GNY model in the large N limit	498
20.4 The large N expansion	501
20.5 The GN model	504
21 Abelian gauge theories: The framework of quantum electrodynamics (QED)	507

21.1 The free massive vector field: Quantization	507
21.2 The Euclidean free action. The two-point function	509
21.3 Coupling to matter	512
21.4 The massless limit: Gauge invariance	514
21.5 Massless vector field, gauge invariance, and quantization	516
21.6 Equivalence with covariant quantization	519
21.7 Gauge symmetry and parallel transport	521
21.8 Perturbation theory: Regularization	522
21.9 WT identities and renormalization	526
21.10 Gauge dependence: The fermion two-point function	528
21.11 Renormalization and RG equations	531
21.12 One-loop β function and the triviality issue	532
21.13 The Abelian Landau–Ginzburg–Higgs model	535
21.14 The Landau–Ginzburg–Higgs model: WT identities	537
21.15 Spontaneous symmetry breaking: Decoupling gauge	538
21.16 Physical observables. Unitarity of the S -matrix	539
21.17 Stochastic quantization: The example of gauge theories	540
A21 Additional remarks	542
A21.1 Vacuum energy and Casimir effect	542
A21.2 Gauge dependence	545
A21.3 Divergences at one loop from Schwinger’s representation	546
22 Non-Abelian gauge theories: Introduction	548
22.1 Geometric construction: Parallel transport	548
22.2 Gauge-invariant actions	551
22.3 Hamiltonian formalism. Quantization in the temporal gauge	551
22.4 Covariant gauges	554
22.5 Perturbation theory, regularization	557
22.6 The non-Abelian Higgs mechanism	559
A22 Massive Yang–Mills fields	565
23 The Standard Model (SM) of fundamental interactions	567
23.1 Weak and electromagnetic interactions: Gauge and scalar fields	568
23.2 Leptons: Minimal SM extension with Dirac neutrinos	570
23.3 Quarks and weak-electromagnetic interactions	573
23.4 QCD. RG equations and β function	576
23.5 General RG β -function at one-loop order: Asymptotic freedom	578
23.6 Axial current, chiral gauge theories, and anomalies	582
23.7 Anomalies: Applications in particle physics	591
24 Large-momentum behaviour in quantum field theory (QFT)	593
24.1 The $(\phi^2)^2$ Euclidean field theory: Large-momentum behaviour	593
24.2 General ϕ^4 -like field theories: d=4	598
24.3 Theories with scalar bosons and Dirac fermions	600
24.4 Gauge theories	602
24.5 Applications: The theory of strong interactions	604
25 Lattice gauge theories: Introduction	607
25.1 Gauge invariance on the lattice: Parallel transport	607
25.2 The matterless gauge theory	609
25.3 Wilson’s loop and confinement	611

25.4 Mean-field approximation	617
A25 Gauge theory and confinement in two dimensions	621
26 Becchi–Rouet–Stora–Tyutin (BRST) symmetry. Gauge theories:	
Zinn–Justin equation (ZJ) and renormalization	623
26.1 ST identities: The origin	624
26.2 From ST symmetry to BRST symmetry	626
26.3 BRST symmetry: More general coordinates. Group structure	628
26.4 Stochastic equations	630
26.5 BRST symmetry, Grassmann coordinates, and gradient equations	632
26.6 Gauge theories: Notation and algebraic structure	635
26.7 Gauge theories: Quantization	636
26.8 WT identities and ZJ equation	639
26.9 Renormalization: General considerations	641
26.10 The renormalized gauge action	642
26.11 Gauge independence: Physical observables	647
A26 BRST symmetry and ZJ equation: Additional remarks	649
A26.1 BRST symmetry and ZJ equation	649
A26.2 Canonical invariance of the ZJ equation	650
A26.3 Elements of BRST cohomology	651
A26.4 From BRST symmetry to supersymmetry	654
27 Supersymmetric quantum field theory (QFT): Introduction	656
27.1 Scalar superfields in three dimensions	656
27.2 The $O(N)$ supersymmetric non-linear σ model	661
27.3 Supersymmetry in four dimensions	662
27.4 Vector superfields and gauge invariance	666
28 Elements of classical and quantum gravity	670
28.1 Manifolds. Change of coordinates. Tensors	671
28.2 Parallel transport: Connection, covariant derivative	673
28.3 Riemannian manifold. The metric tensor	677
28.4 The curvature (Riemann) tensor	678
28.5 Fermions, vielbein, spin connection	682
28.6 Classical GR. Equations of motion	684
28.7 Quantization in the temporal gauge: Pure gravity	687
28.8 Observational cosmology: A few comments	690
29 Generalized non-linear σ-models in two dimensions	692
29.1 Homogeneous spaces and Goldstone modes	692
29.2 WT identities and renormalization in linear coordinates	695
29.3 Renormalization in general coordinates: BRST symmetry	699
29.4 Symmetric spaces: Definition	703
29.5 Classical field equations. Conservation laws	704
29.6 QFT: Perturbative expansion and RG	706
29.7 Generalizations	711
A29 Homogeneous spaces: A few algebraic properties	713
A29.1 Pure gauge. Maurer–Cartan equations	713
A29.2 Metric and curvature in homogeneous spaces	714
A29.3 Explicit expressions for the metric	715
A29.4 Symmetric spaces: Classification	717

30 A few solvable two-dimensional quantum field theories (QFT)	721
30.1 The free massless scalar field	721
30.2 The free massless Dirac fermion	725
30.3 The gauge-invariant fermion determinant and the anomaly	728
30.4 The sG model	731
30.5 The Schwinger model	732
30.6 The massive Thirring model	736
30.7 A generalized Thirring model with two fermions	739
30.8 The $SU(N)$ Thirring model	742
A30 Two-dimensional models: A few additional results	745
A30.1 Four-fermion current interactions: RG β -function	745
A30.2 The Schwinger model: The anomaly	745
A30.3 Solitons in the sG model	746
31 O(2) spin model and the Kosterlitz–Thouless’s (KT) phase transition	747
31.1 The spin correlation functions at low temperature	748
31.2 Correlation functions in a field	749
31.3 The Coulomb gas in two dimensions	750
31.4 $O(2)$ spin model and Coulomb gas	755
31.5 The critical two-point function in the $O(2)$ model	756
31.6 The generalized Thirring model	758
32 Finite-size effects in field theory. Scaling behaviour	760
32.1 RG in finite geometries	760
32.2 Momentum quantization	764
32.3 The ϕ^4 field theory in a periodic hypercube	766
32.4 The ϕ^4 field theory: Cylindrical geometry	772
32.5 Finite size effects in the non-linear σ -model	776
A32 Additional remarks	782
A32.1 Perturbation theory in a finite volume	782
A32.2 Discrete symmetries and finite-size effects	783
33 Quantum field theory (QFT) at finite temperature: Equilibrium properties	786
33.1 Finite- (and high-) temperature field theory	786
33.2 The example of the $\phi_{1,d-1}^4$ field theory	790
33.3 High temperature and critical limits	796
33.4 The non-linear σ -model in the large N limit	799
33.5 The perturbative non-linear σ -model at finite temperature	804
33.6 The GN model in the large N expansion	810
33.7 Abelian gauge theories: The QED framework	817
33.8 Non-Abelian gauge theories	824
A33 Feynman diagrams at finite temperature	828
A33.1 One-loop calculations	828
A33.2 Group measure	830
34 Stochastic differential equations: Langevin, Fokker–Planck (FP) equations	831
34.1 The Langevin equation	831
34.2 Time-dependent probability distribution and FP equation	833
34.3 Equilibrium distribution. Correlation functions	835
34.4 A special class: Dissipative Langevin equations	838
34.5 The linear Langevin equation	839

34.6 Path integral representation	842
34.7 BRST and supersymmetry	843
34.8 Gradient time-dependent force and Jarzynski's relation	846
34.9 More general Langevin equations. Motion in Riemannian manifolds	848
A34 Markov's stochastic processes: A few remarks	852
A34.1 Discrete spaces: Markov's processes, phase transitions	852
A34.2 Stochastic process with prescribed equilibrium distribution	855
A34.3 Stochastic processes and phase transitions	856
35 Langevin field equations: Properties and renormalization	857
35.1 Langevin and Fokker–Planck (FP) equations	857
35.2 Time-dependent correlation functions and equilibrium	858
35.3 Renormalization and BRST symmetry	861
35.4 Dissipative Langevin equation and supersymmetry	864
35.5 Supersymmetry and equilibrium correlation functions	867
35.6 Stochastic quantization of two-dimensional chiral models	868
35.7 Langevin equation and Riemannian manifolds	871
A35 The random field Ising model: Supersymmetry	874
36 Critical dynamics and renormalization group (RG)	875
36.1 Dissipative equation: RG equations in dimension $d = 4 - \varepsilon$	876
36.2 Dissipative dynamics: RG equations in dimension $d = 2 + \varepsilon$	880
36.3 Conserved order parameter	882
36.4 Relaxational model with energy conservation	883
36.5 A non-relaxational model	886
36.6 Finite size effects and dynamics	888
A36 RG functions at two loops	894
A36.1 Supersymmetric perturbative calculations at two loops	894
37 Instantons in quantum mechanics (QM)	899
37.1 The quartic anharmonic oscillator for negative coupling	899
37.2 A toy model: A simple integral	901
37.3 QM: Instantons	902
37.4 Instanton contributions at leading order	904
37.5 General analytic potentials: Instanton contributions	908
37.6 Evaluation of the determinant: The shifting method	909
37.7 Zero temperature limit: The ground state	915
A37 Exact Jacobian. WKB method.	916
A37.1 The exact Jacobian	916
A37.2 The WKB method	917
38 Metastable vacua in quantum field theory (QFT)	919
38.1 The ϕ^4 QFT for negative coupling	920
38.2 General potentials: Instanton contributions	924
38.3 The ϕ^4 QFT in dimension 4	926
38.4 Instanton contributions at leading order	927
38.5 Coupling constant renormalization	931
38.6 The imaginary part of the n -point function	932
38.7 The massive theory	933
38.8 Cosmology: The decay of the false vacuum	934
A38 Instantons: Additional remarks	936

A38.1 Virial theorem	936
A38.2 Sobolev inequalities	937
A38.3 Instantons and RG equations	939
A38.4 Conformal invariance	940
39 Degenerate classical minima and instantons	942
39.1 The quartic double-well potential	942
39.2 The periodic cosine potential	945
39.3 Instantons and stochastic dynamics	948
39.4 Instantons in stable boson field theories: General remarks	951
39.5 Instantons in $CP(N-1)$ models	953
39.6 Instantons in the $SU(2)$ gauge theory	956
A39 Trace formula for periodic potentials	959
40 Large order behaviour of perturbation theory	960
40.1 QM	960
40.2 Scalar field theories: The example of the ϕ^4 field theory	963
40.3 The $(\phi^2)^2$ field theory in dimension 4 and $4 - \varepsilon$	964
40.4 Field theories with fermions	968
A40 Large-order behaviour: Additional remarks	974
41 Critical exponents and equation of state from series summation	975
41.1 Divergent series: Borel summability, Borel summation	975
41.2 Borel transformation: Series summation	978
41.3 Summing the perturbative expansion of the $(\phi^2)^2$ field theory	980
41.4 Summation method: Practical implementation	983
41.5 Field theory estimates of critical exponents for the $O(N)$ model	985
41.6 Other three-dimensional theoretical estimates	986
41.7 Critical exponents from experiments	987
41.8 Amplitude ratios	989
A41 Some other summation methods	990
A41.1 Order-dependent mapping method (ODM)	990
A41.2 Linear differential approximants	991
42 Multi-instantons in quantum mechanics (QM)	992
42.1 The quartic double-well potential	993
42.2 The periodic cosine potential	1000
42.3 General potentials with degenerate minima	1004
42.4 The $O(\nu)$ -symmetric anharmonic oscillator	1007
42.5 Generalized Bohr–Sommerfeld quantization formula	1009
A42 Additional remarks	1011
A42.1 Multi-instantons: The determinant	1011
A42.2 The instanton interaction	1012
A42.3 A simple example of non-Borel summability	1014
A42.4 Multi-instantons and WKB approximation	1016
Bibliography	1019
Index	1041

1 Gaussian integrals. Algebraic preliminaries

Since our study of perturbative aspects of quantum mechanics and quantum field theory (QFT) is largely based on functional (path or field) integrals and functional techniques, topics that may not be necessarily familiar to all readers, we begin this physics textbook with a discussion of the algebraic properties of Gaussian measures and Gaussian expectation values for a finite number of variables. The important role of Gaussian measures is not unrelated to the central limit theorem of probabilities, although the interesting physics is generally hidden in essential deviations from Gaussian distributions.

We first recall a few algebraic identities about Gaussian expectation values, in particular, Wick's theorem. We emphasize the role of cumulants. We discuss the steepest descent method, which reduces certain type of integrals to series whose terms are given by Gaussian expectation values.

The discussion of boson systems (see Chapter 4) also requires defining integrals over some type of formally complex conjugate variables.

By contrast, to discuss fermion systems, one first needs Grassmann or exterior algebras, and the corresponding generalization of the notions of differentiation and integration.

Both for complex and Grassmann integrals, we calculate Gaussian integrals and Gaussian expectation values, and prove generalized Wick's theorems.

Finally, as a preparation for the coming chapters, we recall the notions of generating functions and Legendre transformations.

All algebraic identities are derived for a finite number of variables, but the coming chapters will show that the extension to infinite systems is simple.

Notation. In most of this work, *for vectors and matrices, boldface will denote a matrix or a vector in its entirety, and the corresponding italics with indices will denote elements.*

1.1 Gaussian integrals: Wick's theorem

In this section, we briefly review a few basic algebraic properties of Gaussian integrals.

We first consider an n -dimensional Gaussian integral over real variables x_i , $i = 1, \dots, n$, of the form,

$$\mathcal{Z}(\mathbf{S}) = \int_{\mathbb{R}^n} d^n x e^{-S_2(\mathbf{x})}, \quad (1.1)$$

where S_2 is the real positive quadratic form,

$$S_2(\mathbf{x}) \equiv \frac{1}{2} \sum_{i,j=1}^n x_i S_{ij} x_j, \quad \text{with} \quad S_{ij} = S_{ji}, \quad \text{and } \mathbf{S} > 0. \quad (1.2)$$

Since the matrix \mathbf{S} is strictly positive, it can be diagonalized by an orthogonal transformation matrix \mathbf{O} , and has positive eigenvalues s_i . Changing variables, $x_i \mapsto x'_i$ with

$$\sum_j O_{ij} x_j = x'_i, \quad |\det \mathbf{O}| = 1,$$

a transformation of Jacobian unity, we obtain a product of independent x'_i integrals.

Each integral yields a factor $\sqrt{2\pi/s_i}$. The result thus involves the product of all eigenvalues, that is, the determinant. The result is

$$\mathcal{Z}(\mathbf{S}) = (2\pi)^{n/2} (\det \mathbf{S})^{-1/2}. \quad (1.3)$$

Moreover, since both the initial integral and the determinant are analytic functions of the coefficients of the matrix \mathbf{S} , the identity can be extended by analytic continuation to complex matrices (the global sign then requires some special care).

We now consider the more general integral

$$\mathcal{Z}(\mathbf{S}, \mathbf{b}) = \int d^n x e^{-S_2(\mathbf{x}) + \mathbf{b} \cdot \mathbf{x}}, \quad \text{where } \mathbf{b} \cdot \mathbf{x} \equiv \sum_{i=1}^n b_i x_i. \quad (1.4)$$

To calculate $\mathcal{Z}(\mathbf{S}, \mathbf{b})$, one first looks for the minimum of the quadratic form, given by

$$\frac{\partial}{\partial x_i} (S_2(\mathbf{x}) - \mathbf{b} \cdot \mathbf{x}) = 0 \quad \Rightarrow \quad \sum_j S_{ij} x_j = b_i.$$

The solution is

$$x_i = \sum_j \Delta_{ij} b_j, \quad \text{with} \quad \Delta \mathbf{S} = \mathbf{1}. \quad (1.5)$$

One then changes variables $\mathbf{x} \mapsto \mathbf{y}$, setting

$$x_i = \sum_j \Delta_{ij} b_j + y_i \quad \Rightarrow \quad -S_2(\mathbf{x}) + \mathbf{b} \cdot \mathbf{x} = w_2(\mathbf{b}) - S_2(\mathbf{y}), \quad (1.6)$$

with

$$w_2(\mathbf{b}) = \frac{1}{2} \sum_{i,j=1}^n b_i \Delta_{ij} b_j. \quad (1.7)$$

The integral becomes

$$\mathcal{Z}(\mathbf{S}, \mathbf{b}) = e^{w_2(\mathbf{b})} \int d^n y e^{-S_2(\mathbf{y})} = (2\pi)^{n/2} (\det \mathbf{S})^{-1/2} e^{w_2(\mathbf{b})}. \quad (1.8)$$

Gaussian expectation values. We consider the Gaussian expectation values,

$$\langle x_{k_1} x_{k_2} \cdots x_{k_\ell} \rangle \equiv \mathcal{Z}^{-1}(\mathbf{S}, 0) \int d^n x x_{k_1} x_{k_2} \cdots x_{k_\ell} e^{-S_2(\mathbf{x})}, \quad (1.9)$$

in which the normalization is determined by the condition $\langle 1 \rangle = 1$.

From expression (1.4), one derives

$$\frac{\partial}{\partial b_k} \mathcal{Z}(\mathbf{S}, \mathbf{b}) = \int d^n x x_k e^{-S_2(\mathbf{x}) + \mathbf{b} \cdot \mathbf{x}}. \quad (1.10)$$

Repeated differentiation with respect to \mathbf{b} then leads to the identity

$$\langle x_{k_1} x_{k_2} \cdots x_{k_\ell} \rangle = (2\pi)^{-n/2} (\det \mathbf{S})^{1/2} \left[\frac{\partial}{\partial b_{k_1}} \frac{\partial}{\partial b_{k_2}} \cdots \frac{\partial}{\partial b_{k_\ell}} \mathcal{Z}(\mathbf{S}, \mathbf{b}) \right] \Big|_{\mathbf{b}=0}.$$

Replacing the integral $\mathcal{Z}(\mathbf{S}, \mathbf{b})$ by its explicit form (1.8), one obtains

$$\langle x_{k_1} \cdots x_{k_\ell} \rangle = \left\{ \frac{\partial}{\partial b_{k_1}} \cdots \frac{\partial}{\partial b_{k_\ell}} e^{w_2(\mathbf{b})} \right\}_{\mathbf{b}=0}. \quad (1.11)$$

More generally, if $F(x)$ is a power series in the variables x_i , then

$$\langle F(x) \rangle = \left[F(\partial/\partial b) e^{w_2(\mathbf{b})} \right]_{\mathbf{b}=0}. \quad (1.12)$$

Wick's theorem. Using the identity (1.11), one derives a first form of Wick's theorem. Each time $\partial/\partial b$ acts on the exponential in the right-hand side, it generates a factor b . A second differential operator has to act on the same factor, otherwise the corresponding contribution vanishes when one sets $\mathbf{b} = 0$. This leads to the following expression for the expectation value of the product $x_{k_1} \cdots x_{k_\ell}$ with the normalized Gaussian weight $e^{-S_2(\mathbf{x})}$: one considers all possible pairings of the indices k_1, \dots, k_ℓ (ℓ must thus be even); to each pair k_p, k_q , one associates the element $\Delta_{k_p k_q}$ of the matrix Δ . Then,

$$\langle x_{k_1} \cdots x_{k_\ell} \rangle = \sum_{\substack{\text{all possible pairings} \\ P \text{ of } \{k_1, \dots, k_\ell\}}} \Delta_{k_{P_1} k_{P_2}} \cdots \Delta_{k_{P_{\ell-1}} k_{P_\ell}}, \quad (1.13)$$

$$= \sum_{\substack{\text{all possible pairings} \\ P \text{ of } \{k_1, \dots, k_\ell\}}} \langle x_{k_{P_1}} x_{k_{P_2}} \rangle \cdots \langle x_{k_{P_{\ell-1}}} x_{k_{P_\ell}} \rangle. \quad (1.14)$$

Equations (1.13) and (1.14), which express Wick's theorem, generalize immediately to an infinite number of variables [1] and, therefore, are useful in statistical and quantum theories, in their functional integral formulation.

1.2 Perturbative expansion. Connected contributions

1.2.1 Perturbation theory

We consider the more general integral

$$\mathcal{Z}(\lambda) = \int d^n x \exp(-S_2(\mathbf{x}) - \lambda V(x)), \quad (1.15)$$

in which $S_2(\mathbf{x})$ is the quadratic form (1.2), $V(x)$ a polynomial in the variables x_i , and λ a parameter. To calculate the integral, we expand the integrand in powers of λ . Formally,

$$\mathcal{Z}(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int d^n x e^{-S_2(\mathbf{x})} V^k(x).$$

The successive terms in the expansion are proportional to Gaussian expectation values of polynomials, which can be evaluated with the use of Wick's theorem (1.13):

$$\mathcal{Z}(\lambda) = \mathcal{Z}(0) \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \langle V^k(x) \rangle, \quad (1.16)$$

where $\langle \bullet \rangle$ means Gaussian expectation value. Since the function $e^{-\lambda V}$ has a power series expansion in x , we also obtain a formal expression of the integral by applying the identity (1.12) with $F = e^{-\lambda V}$ ($w_2(\mathbf{b})$ being defined in equation (1.7)):

$$\mathcal{Z}(\lambda) = \mathcal{Z}(0) \left\{ \exp \left[-\lambda V \left(\frac{\partial}{\partial b} \right) \right] e^{w_2(\mathbf{b})} \right\}_{\mathbf{b}=0}. \quad (1.17)$$

1.2.2 Connected contributions or cumulants

In the expansion (1.16), the coefficient of λ is $\langle V(x) \rangle$. At order λ^2 appears the expectation value $\langle V(x)V(x) \rangle$. Using Wick's theorem, we note that some contributions have a factorized form when pairings remain internal to each $V(x)$ factor. Their sum is simply $\langle V(x) \rangle \langle V(x) \rangle$. We call *connected* the genuine remaining contributions and use the notation $\langle \bullet \rangle_c$:

$$\langle V^2(x) \rangle = (\langle V(x) \rangle_c)^2 + \langle V^2(x) \rangle_c,$$

($\langle V(x) \rangle_c = \langle V(x) \rangle$). The argument generalizes to higher orders. At order k , we find a set of disconnected contributions, corresponding to all possible decompositions of k into a sum of positive integers $k = k_1 + k_2 + \dots + k_p$. The corresponding contribution

$$\langle V^{k_1}(x) \rangle_c \langle V^{k_2}(x) \rangle_c \cdots \langle V^{k_p}(x) \rangle_c,$$

has, when all k_i are different, a coefficient $1/k!$ from perturbation theory, multiplied by a combinatorial factor associated with all possible ways of gathering k objects into clusters of $k_1 + k_2 \dots$,

$$\frac{(-\lambda)^k}{k!} \times \frac{k!}{k_1!k_2!\dots k_p!} \langle V^{k_1}(x) \rangle_c \langle V^{k_2}(x) \rangle_c \cdots \langle V^{k_p}(x) \rangle_c.$$

Instead, if k_i appears m times in the decomposition, one finds the same contribution $m!$ times and, therefore, one has to divide by $m!$. Summing all the contributions, one obtains

$$\mathcal{W}(\lambda) \equiv \ln \mathcal{Z}(\lambda) = \ln \mathcal{Z}(0) + \sum_{k=1} \frac{(-\lambda)^k}{k!} \langle V^k(x) \rangle_c. \quad (1.18)$$

The new function $\mathcal{W}(\lambda)$, which is the sum of all connected expectation values, plays an important role in statistical physics and QFT.

1.3 The steepest descent method

To evaluate contour integrals in the complex domain, one can sometimes use the steepest descent method, which reduces their evaluation to Gaussian integrals. We consider the integral

$$\mathcal{I}(\lambda) = \int_{\mathbb{R}^n} d^n x e^{-S(\mathbf{x})/\lambda}, \quad (1.19)$$

where $S(\mathbf{x})$ is an analytic function of the variables x_i . In the limit $\lambda \rightarrow 0_+$, the integral is dominated by the saddle points \mathbf{x}^s solution of ($\nabla \equiv (\partial/\partial x_1, \dots, \partial/\partial x_n)$),

$$\nabla S(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^s} = 0. \quad (1.20)$$

To calculate the contribution of a leading *saddle point* \mathbf{x}^s , one changes variables $\mathbf{x} \mapsto \mathbf{y}$, and sets

$$\mathbf{x} = \mathbf{x}^s + \mathbf{y}\sqrt{\lambda}, \quad \text{and} \quad S_{i_1, \dots, i_k}^{(k)} = \left. \frac{\partial^k S}{\partial x_{i_1} \cdots \partial x_{i_k}} \right|_{\mathbf{x}=\mathbf{x}^s}.$$

One then expands $S(\mathbf{x})$ in powers of λ (and thus of \mathbf{y}):

$$S(\mathbf{x}) = S(\mathbf{x}^s) + \frac{\lambda}{2!} \sum_{i,j} S_{ij}^{(2)} y_i y_j + \sum_{k=3}^{\infty} \frac{\lambda^{k/2}}{k!} \sum_{i_1, \dots, i_k} S_{i_1, \dots, i_k}^{(k)} y_{i_1} \cdots y_{i_k}. \quad (1.21)$$

The change of variables is such that the term quadratic in \mathbf{y} in S/λ is independent of λ . The integral becomes

$$\begin{aligned} \mathcal{I}(\lambda) &= \lambda^{n/2} e^{-S(\mathbf{x}^s)/\lambda} \int d^n y \exp \left[-\frac{1}{2} \sum_{i,j} S_{ij}^{(2)} y_i y_j - R(\mathbf{y}) \right], \quad \text{with} \\ R(\mathbf{y}) &= \sum_{k=3}^{\infty} \frac{\lambda^{k/2-1}}{k!} \sum_{i_1, \dots, i_k} S_{i_1, \dots, i_k}^{(k)} y_{i_1} \cdots y_{i_k}. \end{aligned} \quad (1.22)$$

One then expands the integrand in powers of $\sqrt{\lambda}$. At the leading order, one obtains

$$\mathcal{I}(\lambda) \underset{\lambda \rightarrow 0}{\sim} (2\pi\lambda)^{n/2} \left(\det \mathbf{S}^{(2)} \right)^{-1/2} e^{-S(\mathbf{x}^s)/\lambda}. \quad (1.23)$$

At higher orders, the calculation involves Gaussian expectation values of polynomials, which can be evaluated using Wick's theorem.

1.4 Complex structures and Gaussian integrals

We will now introduce a formalism useful to describe boson systems (see Section 4.1). We consider the set of $2n$ complex variables $\{z_i, \bar{z}_i\}$. The variables z_i and \bar{z}_i are two *independent variables*, only formally conjugate, which provide a complex parametrization of phase space, related to the real position q_i and conjugate momentum p_i by

$$z_i = -i(p_i + iq_i)/\sqrt{2}, \quad \bar{z}_i = i(p_i - iq_i)/\sqrt{2}. \quad (1.24)$$

We consider the Gaussian integral,

$$\mathcal{Z}(\mathbf{S}) = \int \left(\prod_{i=1}^n \frac{dz_i d\bar{z}_i}{2i\pi} \right) \exp \left(- \sum_{i,j=1}^n \bar{z}_i S_{ij} z_j \right), \quad (1.25)$$

in which \mathbf{S} is a complex matrix with non-vanishing determinant. It can be calculated by changing to the real variables (p_i, q_i) , defined by equation (1.24) (then, $dz_i d\bar{z}_i = idp_i dq_i$), or by a change of variables like $\sum_j S_{ij} z_j = z'_i$. One finds

$$\mathcal{Z}(\mathbf{S}) = (\det \mathbf{S})^{-1}. \quad (1.26)$$

The generic Gaussian integral is

$$\mathcal{Z}(\mathbf{S}; \mathbf{b}, \bar{\mathbf{b}}) = \int \left(\prod_{i=1}^n \frac{dz_i d\bar{z}_i}{2i\pi} \right) \exp \left(- \sum_{i,j=1}^n \bar{z}_i S_{ij} z_j + \sum_{i=1}^n (\bar{b}_i z_i + b_i \bar{z}_i) \right), \quad (1.27)$$

The calculation of the integral proceeds as in the real case. One defines the complex matrix $\Delta = \mathbf{S}^{-1}$. The terms, linear in z_i and \bar{z}_i , can be eliminated by the change of variables $z_i \mapsto v_i$, $\bar{z}_i \mapsto \bar{v}_i$, with

$$z_i = v_i + \sum_j \Delta_{ij} b_j, \quad \bar{z}_i = \bar{v}_i + \sum_j \bar{b}_j \Delta_{ji}. \quad (1.28)$$

The result is

$$\mathcal{Z}(\mathbf{S}; \mathbf{b}, \bar{\mathbf{b}}) = (\det \mathbf{S})^{-1} \exp \left(\sum_{i,j=1}^n \bar{b}_i \Delta_{ij} b_j \right). \quad (1.29)$$

Wick's theorem. We consider expectations of polynomials in z, \bar{z} with the normalized weight $\exp(-\sum_{i,j} \bar{z}_i S_{ij} z_j) / \mathcal{Z}(\mathbf{S})$. They can be generated by differentiating the expression (1.29) with respect to b_i and \bar{b}_j , and then setting $\mathbf{b} = \bar{\mathbf{b}} = 0$.

Each derivative with respect to b has to be paired with a derivative with respect to \bar{b} , otherwise the contribution vanishes for $\mathbf{b} = \bar{\mathbf{b}} = 0$. Only monomials with an equal number of factors z and \bar{z} have a non-vanishing expectation value. Eventually, one derives Wick's theorem for complex integrals in the form

$$\begin{aligned} \langle z_{i_1} \bar{z}_{j_1} \cdots z_{i_\ell} \bar{z}_{j_\ell} \rangle &= \sum_{\substack{\text{all permutations} \\ P \text{ of } \{j_1, \dots, j_\ell\}}} \Delta_{i_1 j_{P_1}} \Delta_{i_2 j_{P_2}} \cdots \Delta_{i_\ell j_{P_\ell}} \\ &= \sum_{\substack{\text{all permutations} \\ P \text{ of } \{j_1, \dots, j_\ell\}}} \langle z_{i_1} \bar{z}_{j_{P_1}} \rangle \langle z_{i_2} \bar{z}_{j_{P_2}} \rangle \cdots \langle z_{i_\ell} \bar{z}_{j_{P_\ell}} \rangle. \end{aligned} \quad (1.30)$$

Conjugation. We define a conjugation, analogous to complex conjugation, by the transformation that acts as complex conjugation on complex numbers and exchanges z_i and \bar{z}_i . For example,

$$\sum_{i,j} \bar{z}_i S_{ij} z_j \mapsto \sum_{i,j} z_i S_{ij}^* \bar{z}_j.$$

We call a function of the $2n$ z, \bar{z} variables *formally real* if it is invariant under such a conjugation. For example,

$$\sum_{i,j} \bar{z}_i S_{ij} z_j = \sum_{i,j} z_i S_{ij}^* \bar{z}_j = \sum_{i,j} \bar{z}_i S_{ji}^* z_j,$$

and, thus, the matrix \mathbf{S} is Hermitian. One then verifies that every formally real function integrated over a conjugated pair z, \bar{z} , again yields a formally real function.

1.5 Grassmann algebras. Differential forms

Theories involving fermions require a parallel formalism (Section 4.5). Since fermion wave functions or field correlation functions are antisymmetric with respect to the exchange of two arguments, the construction of generating functionals requires the introduction of anticommuting classical functions, and thus Grassmann variables [2].

Grassmann algebra. A Grassmann (or exterior) algebra \mathfrak{A} over \mathbb{R} or \mathbb{C} (real or complex) is an associative algebra constructed from a unit 1 and a set of generators θ_i with anticommuting products:

$$\theta_i \theta_j + \theta_j \theta_i = 0, \quad \forall i, j. \quad (1.31)$$

As a consequence:

(i) if the number n of generators is finite, the elements of the algebra form a vector space of finite dimension 2^n over \mathbb{R} or \mathbb{C} . All elements can be written as linear combinations of the elements A_ν , $\nu = 1, \dots, 2^n$ of the form

$$A_\nu \in \{1 \text{ and } \{\theta_{i_1} \theta_{i_2} \cdots \theta_{i_p}\}, \text{ with } i_1 < i_2 < \cdots < i_p, 1 \leq p \leq n\}. \quad (1.32)$$

(ii) \mathfrak{A} is a graded algebra: to each monomial $\theta_{i_1} \theta_{i_2} \cdots \theta_{i_p}$, one can associate an integer p that counts the number of generators in a product. In particular, if A_p and A_q are two monomials of degree p and q , respectively, then

$$A_p A_q = (-1)^{pq} A_q A_p.$$

(iii) Elements of \mathfrak{A} are invertible if, and only if the term of degree 0 in the expansion on the basis (1.32) does not vanish.

For example, the element $1 + \theta$ is invertible and its inverse is $1 - \theta$, but θ is not invertible. The inverse can be calculated by expanding in a formal power series starting from the inverse of the term of degree 0.

(vi) All elements in a Grassmann algebra, considered as functions of a generator θ_i , are first degree polynomials, that is, affine functions.

Grassmannian parity. In the algebra \mathfrak{A} , a simple automorphism P can be defined:

$$P(\theta_i) = -\theta_i \Rightarrow P^2 = 1. \quad (1.33)$$

Then, on a monomial of degree p , P acts like

$$P(\theta_{i_1} \cdots \theta_{i_p}) = (-1)^p \theta_{i_1} \cdots \theta_{i_p}. \quad (1.34)$$

The reflection P divides the algebra \mathfrak{A} in two eigenspaces \mathfrak{A}^\pm containing the even or odd elements

$$P(\mathfrak{A}^\pm) = \pm \mathfrak{A}^\pm.$$

In particular, \mathfrak{A}^+ is a subalgebra, the subalgebra of commuting elements.

1.5.1 Differential forms

Notation. In this section, we use the convention of summation over *repeated upper and lower indices*.

Grassmann algebras can be used to represent explicitly differential forms (see also Section 28.1.3). The language of differential forms will not often be used in this work. However, it is interesting to recall one concept, the exterior derivative of forms whose generalization appears in the context of BRST symmetry (see Section 26.2). We consider totally antisymmetric tensors $\Omega_{\mu_1, \dots, \mu_l}(x)$, which are functions of n commuting variables x^μ . Associating n Grassmann generators θ^μ with the variables x^μ , we can write the corresponding l -form

$$\Omega = \Omega_{\mu_1, \dots, \mu_l}(x) \theta^{\mu_1} \cdots \theta^{\mu_l}, \quad (1.35)$$

where $l \leq n$, otherwise, the form vanishes.

One defines the differential operator d acting on forms by

$$d \equiv \theta^\mu \frac{\partial}{\partial x^\mu}. \quad (1.36)$$

We note that if Ω is an l -form, $d\Omega$ is an $(l+1)$ -form (see Chapter 28 for details). One immediately notes that the square of d vanishes:

$$d^2 = \theta^\mu \frac{\partial}{\partial x^\mu} \theta^\nu \frac{\partial}{\partial x^\nu} = 0, \quad (1.37)$$

because the product $\theta^\mu \theta^\nu$ is antisymmetric in $\mu \leftrightarrow \nu$.

It is a *cohomology operator*. A form Ω that satisfies $d\Omega = 0$ is called *closed* and a form Ω that can be written as $\Omega = d\Omega'$ is called *exact*. The property (1.37) implies that *any exact form is closed*.

Note that, in the case of differential forms, one often writes the generators of the exterior algebra dx^μ instead of θ^μ , and then uses the wedge notation \wedge for the product to indicate that it is antisymmetric.

1.6 Differentiation and integration in Grassmann algebras

To be able to construct parallel formalisms for bosons and fermions based on functional integrals, it is necessary to define differentiation and integration in Grassmann algebras.

1.6.1 Differentiation in Grassmann algebras

A straightforward generalization of the usual rules of differentiation is inconsistent, due to the non-commutative character of the algebra. A suitable definition is obtained in the following way: considered as functions of a specific generator θ_i , all elements A of \mathfrak{A} can be written as

$$A = A_1 + \theta_i A_2,$$

after some commutations, where A_1 and A_2 do not depend on θ_i . Then, one defines

$$\frac{\partial A}{\partial \theta_i} = A_2. \quad (1.38)$$

Note that the differential operator $\partial/\partial\theta_i$ shares one property of the form differentiation (equation (1.37)): its square vanishes, $(\partial/\partial\theta_i)^2 = 0$.

Left and right differentiation. Equation (1.38) defines a left differentiation in the sense that the action of $\partial/\partial\theta_i$ consists in bringing θ_i on the left in a monomial and suppressing it. Similarly, a right differentiation could have been defined by commuting θ_i to the right.

Chain rule. One verifies that the chain rule applies to Grassmann differentiation. If $\sigma(\theta)$ belongs to \mathfrak{A}^- and $x(\theta)$ belongs to \mathfrak{A}^+ , one finds

$$\frac{\partial}{\partial\theta} f(\sigma, x) = \frac{\partial\sigma}{\partial\theta} \frac{\partial f}{\partial\sigma} + \frac{\partial x}{\partial\theta} \frac{\partial f}{\partial x}. \quad (1.39)$$

For the second term in the right-hand side, the order between factors matters.

Formal construction. One can verify that any Grassmann differential operator D (also called anti-derivation) acting on \mathfrak{A} and defined as in equation (1.38), satisfies the two algebraic formal rules:

(i) It is a linear mapping of \mathfrak{A} , considered as a vector space, into itself:

$$D(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 D(A_1) + \lambda_2 D(A_2), \quad \text{for } \lambda_1, \lambda_2 \in \mathbb{R} \text{ or } \mathbb{C}. \quad (1.40)$$

(ii) By contrast with Leibniz's rule, it satisfies the condition

$$D(A_1 A_2) = P(A_1) D(A_2) + D(A_1) A_2. \quad (1.41)$$

The unusual form of equation (1.41) compared to the differentiation rule for commuting variables is required if one wants D to anticommute with P :

$$DP + PD = 0, \quad (1.42)$$

which means that the image of \mathfrak{A}^\pm by D belongs to \mathfrak{A}^\mp .

Note that if A belongs to \mathfrak{A}^+ and $F(x)$ is an ordinary function of real or complex variables, then,

$$D[F(A)] = D(A)F'(A), \quad \text{for } A \in \mathfrak{A}^+. \quad (1.43)$$

Finally, note that the form differentiation (1.36) shares all these properties, but acts on different variables.

Anticommutation relations. A short calculation shows that if D and D' are two operators satisfying conditions (1.40) and (1.41), then the anticommutator

$$\Delta = DD' + D'D \quad (1.44)$$

is a usual differential operator:

$$\begin{aligned} \Delta(\lambda_1 A_1 + \lambda_2 A_2) &= \lambda_1 \Delta(A_1) + \lambda_2 \Delta(A_2), \\ \Delta(A_1 A_2) &= \Delta(A_1) A_2 + A_1 \Delta(A_2). \end{aligned} \quad (1.45)$$

Furthermore,

$$\Delta P = P\Delta. \quad (1.46)$$

These properties, which are the consequence of the addition of relation (1.42) to the definitions (1.40) and (1.41), makes it possible to extend the notion of Lie algebra, and are directly relevant to the discussion of *supersymmetries* (see Section 27.1).

1.6.2 A basis of differential operators

Since a differential operator satisfies conditions (1.40) and (1.41), it is completely defined by its action on the generators θ_i . In addition, any differential operator left-multiplied by an element of \mathfrak{A}^+ still satisfies conditions (1.40) and (1.41). We conclude that any differential operator can be expanded on a basis of operators $\partial/\partial\theta_i$ defined by

$$\frac{\partial}{\partial\theta_i}\theta_j = \delta_{ij} \quad (1.47)$$

with left coefficients in \mathfrak{A}^+ . It is simple to verify that the differential operators $\partial/\partial\theta_i$ coincide with the operators defined by equation (1.38). The nilpotent differential operators $\partial/\partial\theta_i$, together with the generators θ_i considered as operators acting on \mathfrak{A} by left-multiplication, satisfy the *anticommutation relations* (identical to the relations between fermion creation and annihilation operators)

$$\theta_i\theta_j + \theta_j\theta_i = 0, \quad \frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j} + \frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_i} = 0, \quad \theta_i\frac{\partial}{\partial\theta_j} + \frac{\partial}{\partial\theta_j}\theta_i = \delta_{ij}. \quad (1.48)$$

Note the symmetry between the operators θ_i and $\partial/\partial\theta_i$.

The algebra of operators can be identified by introducing the linear combinations

$$D_i^\pm = \frac{\partial}{\partial\theta_i} \pm \theta_i,$$

which satisfy

$$\{D_i^\pm, D_j^\pm\} = \pm 2\delta_{ij}, \quad \{D_i^+, D_j^-\} = 0. \quad (1.49)$$

This shows that the operator algebra is the direct sum of two *Clifford algebras*.

1.6.3 Integration in Grassmann algebras

It is also convenient to define integration over Grassmann variables, for which the integral symbol notation is used, although integration and differentiation are identical operations,

$$\int d\theta_i A \equiv \frac{\partial}{\partial\theta_i} A, \quad \forall A \in \mathfrak{A}. \quad (1.50)$$

The integral or derivative symbols are used depending on the context.

General properties. We will now show that this operation satisfies the formal properties we expect from a *definite* integral. Quite generally, we associate an operator I to a given differential operator D , which has the following defining properties: it is a linear operator acting on \mathfrak{A} , that is,

$$I(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 I(A_1) + \lambda_2 I(A_2), \quad (1.51)$$

and satisfies the three properties,

$$ID = 0, \quad (1.52)$$

$$DI = 0, \quad (1.53)$$

and

$$D(A) = 0 \implies I(BA) = I(B)A. \quad (1.54)$$

In addition, it changes the grading in the same way as a differential operator:

$$PI + IP = 0.$$

Let us explain the conditions (1.52–1.54): condition (1.52) expresses that, in the absence of boundary terms, the integral of a total derivative vanishes; condition (1.53) expresses that if we integrate over a variable, the result no longer depends on this variable. Finally, condition (1.54) implies that a factor whose derivative vanishes can be taken out of the integral.

In the case of Grassmann algebras, if $D^2 = 0$, D itself satisfies all conditions. The differential operators $\partial/\partial\theta_i$ indeed have a vanishing square.

1.6.4 Change of variables in a Grassmann integral

We consider the integral

$$\int d\theta f(\theta), \quad (1.55)$$

and perform the (necessarily) affine change of variables:

$$\theta = a\theta' + b, \quad (1.56)$$

in which parity conservation implies that $a \in \mathfrak{A}^+$ and $b \in \mathfrak{A}^-$. The element a must be invertible, that is, its term of degree zero in the Grassmann variables must be different from zero. Then, using definition (1.50), one finds

$$\int d\theta f(\theta) = a^{-1} \int d\theta' f(\theta'a + b) = \int d\theta' \left(\frac{\partial \theta}{\partial \theta'} \right)^{-1} f(\theta(\theta')), \quad (1.57)$$

where the latter form is independent of the specific parametrization (1.56). We have derived a very important property of Grassmann integrals: the *Jacobian* is a^{-1} , while in the case of commuting variables it is a .

Generalization. More generally, a change of variables

$$\theta_i = \theta_i(\theta'), \quad \theta_i, \theta'_i \in \mathfrak{A}^-,$$

for which the matrix $\partial\theta_i/\partial\theta'_j$ has an invertible part of degree zero, leads to a Jacobian that is the *inverse* of the determinant of $\partial\theta_i/\partial\theta'_j$:

$$d\theta_1 \cdots d\theta_n = d\theta'_1 \cdots d\theta'_n J(\theta'), \quad (1.58)$$

with

$$J = \det \frac{\partial\theta'_i}{\partial\theta_j} = \left(\det \frac{\partial\theta_i}{\partial\theta'_j} \right)^{-1}. \quad (1.59)$$

The determinant is defined because all elements of the matrix $\partial\theta_i/\partial\theta'_j$ belong to \mathfrak{A}^+ .

The result can be derived by changing variables one at a time and using equation (1.57) repeatedly: $\theta_1 \mapsto \theta'_1$, then $\theta_2 \mapsto \theta'_2$, until, finally, $\theta_n \mapsto \theta'_n$. One obtains

$$J = \frac{\partial \theta'_1}{\partial \theta_1} \Big|_{\theta_2, \dots, \theta_n} \frac{\partial \theta'_2}{\partial \theta_2} \Big|_{\theta'_1, \theta_3, \dots, \theta_n} \cdots \frac{\partial \theta'_{n-1}}{\partial \theta_{n-1}} \Big|_{\theta'_1, \dots, \theta'_{n-2}, \theta_n} \frac{\partial \theta'_n}{\partial \theta_n} \Big|_{\theta'_1, \dots, \theta'_{n-1}} .$$

One recognizes one form of the Jacobian for complex variables, but for the change of variables $\theta'_i \mapsto \theta_i$. Indeed, if we introduce the matrices

$$M_{ij}^{(p)} = \frac{\partial \theta'_i}{\partial \theta_j}, \quad i, j \leq p \leq n,$$

using the chain rule (1.39), it is possible to verify the recursion relation

$$\frac{\partial \theta'_n}{\partial \theta_n} \Big|_{\theta'_1, \dots, \theta'_{n-1}} = \frac{\partial \theta'_n}{\partial \theta_n} \Big|_{\theta_1, \dots, \theta_{n-1}} - \sum_{i,j < n} \frac{\partial \theta'_n}{\partial \theta_i} [M^{n-1}]_{ij}^{-1} \frac{\partial \theta'_j}{\partial \theta_n} = \det M^{(n)} [M^{(n-1)}]^{-1}.$$

The expression (1.59) follows.

Example. A straightforward verification of equation (1.58) is provided by the following example:

$$1 = \int d\theta_1 \cdots d\theta_n \theta_n \cdots \theta_1 .$$

After the linear change of variables $\theta \mapsto \theta'$,

$$\theta_i = \sum_j M_{ij} \theta'_j,$$

the result relies upon the identity

$$\theta_n \cdots \theta_1 = \theta'_n \cdots \theta'_1 \det \mathbf{M} .$$

1.6.5 Mixed change of variables

One can meet integrals involving both commuting and anticommuting variables (bosons and fermions). Calculations may then involve mixed changes of variables.

Denoting by θ, θ' and x, x' , the anticommuting and commuting variables, respectively, we set (respecting parity):

$$x_a = x_a(x', \theta') \in \mathfrak{A}_+(\theta'), \quad \theta_i = \theta_i(x', \theta') \in \mathfrak{A}_-(\theta'). \quad (1.60)$$

We introduce the matrix \mathbf{M} of partial derivatives,

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

with

$$\mathbf{A}_{ab} = \frac{\partial x_a}{\partial x'_b}, \quad \mathbf{B}_{ai} = \frac{\partial x_a}{\partial \theta'_i}, \quad \mathbf{C}_{ia} = -\frac{\partial \theta_i}{\partial x'_a}, \quad \mathbf{D}_{ij} = \frac{\partial \theta_i}{\partial \theta'_j}.$$

It is convenient to change variables in two steps:

(i) One first passes from (θ, x) to (θ, x') . This step generates the Jacobian

$$J_1 = \det \left. \frac{\partial x_a}{\partial x'_b} \right|_{\theta} = \det (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) . \quad (1.61)$$

(ii) One then goes from (θ, x') to (θ', x') . The second step just gives, as explained previously, the Jacobian

$$J_2 = (\det \mathbf{D})^{-1} . \quad (1.62)$$

The complete Jacobian J , also called the *Berezinian* of the matrix of partial derivatives, is thus

$$J = \frac{D(x, \theta)}{D(x', \theta')} = J_1 J_2 = \text{Ber } \mathbf{M} \equiv \det (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) (\det \mathbf{D})^{-1} . \quad (1.63)$$

For the Jacobian to be non-singular, the matrices \mathbf{A} and \mathbf{D} must be invertible (and, therefore, their contributions of degree 0 in θ' also).

Trace of mixed matrices. In the case of the integration over ordinary commuting variables, if one performs a change of variables infinitesimally close to the identity,

$$x_a = x'_a + \varepsilon f_a(x'),$$

from the general identity $\ln \det = \text{tr} \ln$ (see Section A2.1), one infers that the Jacobian, expanded at first order in ε , becomes (one can also use the result (1.69) and Wick's theorem (1.78))

$$J = \det \frac{\partial x_a}{\partial x'_b} = 1 + \varepsilon \text{tr} \frac{\partial f_a}{\partial x'_b} + O(\varepsilon^2) = 1 + \varepsilon \sum_a \frac{\partial f_a}{\partial x'_a} + O(\varepsilon^2) .$$

In the mixed case,

$$x_a = x'_a + \varepsilon f_a(x', \theta'), \quad \theta_i = \theta'_i + \varepsilon \varphi_i(x', \theta'), \quad (1.64)$$

we use the identity (1.63). We set

$$\mathbf{M} = 1 + \varepsilon \mathbf{M}_1 + O(\varepsilon^2), \quad \mathbf{M}_1 = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ \mathbf{C}_1 & \mathbf{D}_1 \end{pmatrix},$$

and obtain

$$J = 1 + \varepsilon (\text{tr } \mathbf{A}_1 - \text{tr } \mathbf{D}_1) + O(\varepsilon^2), \quad \text{tr } \mathbf{A}_1 - \text{tr } \mathbf{D}_1 = \sum_a \frac{\partial f_a}{\partial x_a} - \sum_i \frac{\partial \varphi_i}{\partial \theta_i} . \quad (1.65)$$

Therefore, to maintain the connection between Jacobian and trace, one is led to define the supertrace of a mixed matrix, for which the notation Str will be used, as the difference of traces

$$\text{Str } \mathbf{M}_1 = \text{tr } \mathbf{A}_1 - \text{tr } \mathbf{D}_1 . \quad (1.66)$$

It is simple to verify that the supertrace, like the usual trace, has a cyclic property,

$$\text{Str } \mathbf{M}_1 \mathbf{M}_2 = \text{Str } \mathbf{M}_2 \mathbf{M}_1 .$$

1.7 Gaussian integrals with Grassmann variables

We consider now a Grassmann algebra \mathfrak{B} in which the generators are separated into two conjugated sets. We denote by θ_i and $\bar{\theta}_i$, $i = 1, \dots, n$, the corresponding generators.

Complex conjugation. In many situations, a complex conjugation can be defined, which exchanges θ_i and $\bar{\theta}_i$. It has the properties of the Hermitian conjugation of operators: it acts by complex conjugation on complex numbers and

$$\theta_i^\dagger = \bar{\theta}_i, \quad \bar{\theta}_i^\dagger = \theta_i, \quad (A_1 A_2)^\dagger = A_2^\dagger A_1^\dagger, \quad \forall A_1, A_2 \in \mathfrak{B}. \quad (1.67)$$

An invariant quantity, like the integration measure $d\theta_i d\bar{\theta}_i$, is called *formally real*. It is possible to verify that if, in an integral, the integrand is formally real and one integrates over a pair of conjugated variables θ_i and $\bar{\theta}_i$, the result is still formally real.

The Gaussian integral. We calculate Gaussian integrals with, again, the motivation: a reduction of more general integrals to a finite or formal infinite sum of Gaussian integrals. We first consider the integral

$$\mathcal{Z}(\mathbf{S}) = \int d\theta_1 d\bar{\theta}_1 d\theta_2 d\bar{\theta}_2 \cdots d\theta_n d\bar{\theta}_n \exp\left(\sum_{i,j=1}^n \bar{\theta}_i S_{ij} \theta_j\right), \quad (1.68)$$

with S_{ij} complex. The quadratic form and, thus, the integrand is formally real if

$$\sum_{i,j=1}^{\nu} (\bar{\theta}_i S_{ij} \theta_j)^\dagger = \sum_{i,j=1}^{\nu} \bar{\theta}_j S_{ij}^* \theta_i = \sum_{i,j=1}^{\nu} \bar{\theta}_i S_{ij}^\dagger \theta_j,$$

that is, if the matrix \mathbf{S} is Hermitian.

According to the rules of Grassmann integration, the result is simply the coefficient of the product $\bar{\theta}_n \theta_n \cdots \bar{\theta}_1 \theta_1$ in the expansion of the integrand. The integrand can be rewritten as

$$\exp\left(\sum_{i,j=1}^n \bar{\theta}_i S_{ij} \theta_j\right) = \prod_{i=1}^n \exp\left(\bar{\theta}_i \sum_{j=1}^n S_{ij} \theta_j\right) = \prod_{i=1}^n \left(1 + \bar{\theta}_i \sum_{j=1}^n S_{ij} \theta_j\right).$$

In each factor, only the term proportional to $\bar{\theta}_i$ contributes to the integral. In the expansion of the product, the terms that give non-zero contributions to the integral thus are of the form

$$\sum_{\substack{\text{permutations} \\ \{j_1, \dots, j_n\}}} S_{n j_n} S_{n-1 j_{n-1}} \cdots S_{1 j_1} \bar{\theta}_n \theta_{j_n} \cdots \bar{\theta}_1 \theta_{j_1}.$$

A commutation of the generators to put them in the standard order $\bar{\theta}_n \theta_n \cdots \bar{\theta}_1 \theta_1$, generates a sign, which is the signature of the permutation. One then recognizes the coefficient as the determinant of S_{ij} . Thus,

$$\mathcal{Z}(\mathbf{S}) = \det \mathbf{S}. \quad (1.69)$$

The result is the inverse of the one obtained with complex commuting variables. In perturbative expansions, this leads to a sign $(-1)^L$ in front of the Feynman diagrams with L fermion loops.

For $\det \mathbf{S} \neq 0$, a simpler calculation relies on the change of variables $\theta_i \mapsto \theta'_i$,

$$\sum_j S_{ij} \theta_j = \theta'_i. \quad (1.70)$$

Then, using the relations (1.58) and (1.59) for the Jacobian, one finds

$$\begin{aligned}\mathcal{Z}(\mathbf{S}) &= \det \mathbf{S} \int d\theta'_1 d\bar{\theta}_1 \cdots d\theta'_n d\bar{\theta}_n \exp\left(\sum_{i=1}^n \bar{\theta}_i \theta'_i\right) \\ &= \det \mathbf{S} \int \prod_{i=1}^n d\theta'_i d\bar{\theta}_i (1 + \bar{\theta}_i \theta'_i) = \det \mathbf{S}.\end{aligned}$$

1.7.1 General Gaussian integrals

We introduce two other sets of Grassmann generators $\{\eta_i\}$ and $\{\bar{\eta}_i\}$, $i = 1, \dots, n$, and consider the larger Grassmann algebra \mathfrak{C} generated by the set of the $4n$ generators $\{\theta_i, \bar{\theta}_i, \eta_i, \bar{\eta}_i\}$. Following the strategy of Section 1.1, we first calculate the integral

$$\mathcal{Z}_G(\eta, \bar{\eta}) = \int \prod_i d\theta_i d\bar{\theta}_i \exp\left[\sum_{i,j=1}^n S_{ij} \bar{\theta}_i \theta_j + \sum_{i=1}^n (\bar{\eta}_i \theta_i + \bar{\theta}_i \eta_i)\right], \quad (1.71)$$

in which the integrand is an element of \mathfrak{C} , and S_{ij} is a complex matrix with $\det \mathbf{S} \neq 0$.

Again we define $\Delta = \mathbf{S}^{-1}$. The calculation, as before, relies on a change of variables $\theta \mapsto \theta'$, $\bar{\theta} \mapsto \bar{\theta}'$, with

$$\theta_i = \theta'_i - \sum_j \Delta_{ij} \eta_j, \quad \bar{\theta}_i = \bar{\theta}'_i - \sum_j \bar{\eta}_j \Delta_{ji},$$

and leads to the result

$$\mathcal{Z}_G(\eta, \bar{\eta}) = \det \mathbf{S} \exp\left(-\sum_{i,j=1}^n \bar{\eta}_i \Delta_{ij} \eta_j\right). \quad (1.72)$$

Using the notation $\langle \bullet \rangle$ for expectation values with respect to the Gaussian weight of equation (1.71), with our definition of \mathcal{Z}_G , one finds (note the sign in equation (1.74)),

$$\frac{\partial}{\partial \bar{\eta}_i} \mathcal{Z}_G = \det \mathbf{S} \langle \theta_i \rangle, \quad (1.73)$$

$$\frac{\partial}{\partial \eta_i} \mathcal{Z}_G = \det \mathbf{S} \langle -\bar{\theta}_i \rangle. \quad (1.74)$$

Wick's theorem for Grassmann integrals. Gaussian expectation values are defined by

$$\begin{aligned}\langle \bar{\theta}_{i_1} \theta_{j_1} \bar{\theta}_{i_2} \theta_{j_2} \cdots \bar{\theta}_{i_n} \theta_{j_n} \rangle \\ = \frac{1}{\det \mathbf{S}} \int \left(\prod_i d\theta_i d\bar{\theta}_i \right) \bar{\theta}_{i_1} \theta_{j_1} \cdots \bar{\theta}_{i_n} \theta_{j_n} \exp\left(\sum_{i,j=1}^n \bar{\theta}_i S_{ij} \theta_j\right),\end{aligned} \quad (1.75)$$

From equations (1.73) and (1.74), it follows that

$$\det \mathbf{S} \langle \bar{\theta}_{i_1} \theta_{j_1} \bar{\theta}_{i_2} \theta_{j_2} \cdots \bar{\theta}_{i_n} \theta_{j_n} \rangle = \left[\frac{\partial}{\partial \bar{\eta}_{j_1}} \frac{\partial}{\partial \eta_{i_1}} \cdots \frac{\partial}{\partial \bar{\eta}_{j_n}} \frac{\partial}{\partial \eta_{i_n}} \mathcal{Z}_G(\eta, \bar{\eta}) \right] \Big|_{\eta=\bar{\eta}=0}, \quad (1.76)$$

and, using the result (1.72),

$$\begin{aligned}\langle \bar{\theta}_{i_1} \theta_{j_1} \bar{\theta}_{i_2} \theta_{j_2} \cdots \bar{\theta}_{i_n} \theta_{j_n} \rangle \\ = \left\{ \frac{\partial}{\partial \bar{\eta}_{j_1}} \frac{\partial}{\partial \eta_{i_1}} \cdots \frac{\partial}{\partial \bar{\eta}_{j_n}} \frac{\partial}{\partial \eta_{i_n}} \exp\left[-\sum_{i,j=1}^n \bar{\eta}_j \Delta_{ji} \eta_i\right] \right\} \Big|_{\eta=\bar{\eta}=0}.\end{aligned} \quad (1.77)$$

After explicit differentiation (which is the same as integration), one obtains

$$\begin{aligned} \langle \bar{\theta}_{i_1} \theta_{j_1} \cdots \bar{\theta}_{i_n} \theta_{j_n} \rangle &= \det \Delta_{j_i i_k} = \det \langle \bar{\theta}_{i_k} \theta_{j_i} \rangle \\ &= \sum_{\substack{\text{permutations } P \\ \text{of } \{j_1, \dots, j_n\}}} \operatorname{sgn}(P) \Delta_{j_{P_1} i_1} \Delta_{j_{P_2} i_2} \cdots \Delta_{j_{P_n} i_n}, \end{aligned} \quad (1.78)$$

where $\operatorname{sgn}(P)$ is the signature of the permutation P , which is Wick's theorem for Grassmann algebras. The result differs from the expression (1.30), obtained in the case of complex commuting variables, only by the factor $\operatorname{sgn}(P)$.

Perturbative expansion. Expressions (1.73) and (1.74) form the basis of perturbation theory. To calculate the integral

$$\mathcal{Z}(\eta, \bar{\eta}) = \int \prod_i d\bar{\theta}_i d\theta_i \exp \left[\sum_{i,j=1}^n S_{ij} \bar{\theta}_i \theta_j + V(\bar{\theta}, \theta) + \sum_{i=1}^n (\bar{\eta}_i \theta_i + \bar{\theta}_i \eta_i) \right], \quad (1.79)$$

one can formally expand in a power series in V and integrate term by term. One then finds

$$\mathcal{Z}(\eta, \bar{\eta}) = \exp \left[V \left(-\frac{\partial}{\partial \eta}, \frac{\partial}{\partial \bar{\eta}} \right) \right] \mathcal{Z}_G(\eta, \bar{\eta}). \quad (1.80)$$

1.7.2 Pfaffian and determinant

One can also calculate Gaussian integrals of the form

$$\mathcal{Z}(\mathbf{A}) = \int d\theta_{2n} \cdots d\theta_2 d\theta_1 \exp \left(\frac{1}{2} \sum_{i,j=1}^{2n} \theta_i A_{ij} \theta_j \right), \quad (1.81)$$

where, since the product $\theta_i \theta_j$ is antisymmetric in (ij) , the matrix \mathbf{A} can be chosen to be antisymmetric:

$$A_{ij} + A_{ji} = 0. \quad (1.82)$$

Expanding the exponential in a power series, one notes that only the term of order n that contains all products of degree $2n$ in θ ,

$$\mathcal{Z}(\mathbf{A}) = \frac{1}{2^n n!} \int d\theta_{2n} \cdots d\theta_1 \left(\sum_{i,j} \theta_i A_{ij} \theta_j \right)^n, \quad (1.83)$$

gives a non-zero contribution. In the expansion of the product, only the terms containing a permutation of $\theta_1 \cdots \theta_{2n}$ do not vanish. Then, ordering all terms to factorize the product $\theta_1 \theta_2 \cdots \theta_{2n}$, one finds

$$\mathcal{Z}(\mathbf{A}) = \frac{1}{2^n n!} \sum_{\substack{\text{permutations } P \\ \text{of } \{i_1, \dots, i_{2n}\}}} \operatorname{sgn}(P) A_{i_1 i_2} A_{i_3 i_4} \cdots A_{i_{2n-1} i_{2n}}, \quad (1.84)$$

where $\operatorname{sgn}(P) = \pm 1$ is the signature of the permutation P . The quantity in the right-hand side is called the *Pfaffian* of the antisymmetric matrix \mathbf{A} :

$$\mathcal{Z}(\mathbf{A}) = \operatorname{Pf}(\mathbf{A}). \quad (1.85)$$

With Grassmann integral techniques, one can prove the classical algebraic identity,

$$\text{Pf}^2(\mathbf{A}) = \det \mathbf{A}. \quad (1.86)$$

Indeed, $\mathcal{Z}^2(\mathbf{A})$ can be written as

$$\mathcal{Z}^2(\mathbf{A}) = \int d\theta_{2n} \cdots d\theta_1 d\theta'_{2n} \cdots d\theta'_1 \exp \left[\frac{1}{2} \sum_{i,j} (\theta_i A_{ij} \theta_j + \theta'_i A_{ij} \theta'_j) \right]. \quad (1.87)$$

We change variables, setting

$$\eta_k = \frac{1}{\sqrt{2}} (\theta_k + i\theta'_k), \quad \bar{\eta}_k = \frac{1}{\sqrt{2}} (\theta_k - i\theta'_k).$$

The Jacobian is $(-1)^n$. Also

$$\theta_i \theta_j + \theta'_i \theta'_j = \bar{\eta}_i \eta_j - \bar{\eta}_j \eta_i, \quad (1.88)$$

$$d\eta_{2n} \cdots d\eta_1 d\bar{\eta}_{2n} \cdots d\bar{\eta}_1 = (-1)^{n^2} \prod_i d\eta_i d\bar{\eta}_i. \quad (1.89)$$

Using the antisymmetry of the matrix \mathbf{A} , one then finds

$$\text{Pf}^2(\mathbf{A}) = \int d\eta_1 d\bar{\eta}_1 \cdots d\eta_{2n} d\bar{\eta}_{2n} \exp \left(\sum_{i,j} \bar{\eta}_i A_{ij} \eta_j \right) = \det \mathbf{A}.$$

Wick's theorem. One can prove another version of Wick's theorem for expectation values with the weight $\exp[\sum_{ij} \theta_i A_{ij} \theta_j / 2]$. One finds

$$\langle \theta_{i_1} \theta_{i_2} \cdots \theta_{i_{2p}} \rangle = \sum_{\substack{\text{all possible pairings} \\ P \text{ of } \{i_1, \dots, i_{2p}\}}} \text{sgn}(P) \langle \theta_{i_{P_1}} \theta_{i_{P_2}} \rangle \cdots \langle \theta_{i_{P_{2p-1}}} \theta_{i_{P_{2p}}} \rangle, \quad (1.90)$$

where $\text{sgn}(P)$ is the signature of the permutation P .

1.8 Legendre transformation

The Legendre transformation relates Hamiltonian and Lagrangian in classical mechanics, the free energy and the thermodynamic potential in statistical physics, and the generating functionals of connected correlation and vertex functions (to which contribute only one-line-irreducible Feynman diagrams). We recall here the definition (assuming for simplicity a finite number of variables) and explain a few basic properties.

We consider a real function $W(\mathbf{h})$ of n variables h_i , which has first- and second-order partial derivatives, and such that the matrix of elements $\partial^2 W / \partial h_i \partial h_j$ is positive:

$$\forall \mathbf{v} \in \mathbb{R}^n \text{ and } |\mathbf{v}| > 0, \quad \sum_{i,j} v_i \frac{\partial^2 W}{\partial h_i \partial h_j} v_j > 0.$$

The Legendre transform of the function $W(\mathbf{h})$ is a function $\Gamma(\mathbf{m})$ defined by the two equations,

$$W(\mathbf{h}) + \Gamma(\mathbf{m}) = \sum_i h_i m_i , \quad (1.91a)$$

$$m_i = \frac{\partial W}{\partial h_i} . \quad (1.91b)$$

The positivity of the matrix of second derivatives ensures that, when equation (1.91b) can be inverted, it has a unique solution $\mathbf{h}(\mathbf{m})$.

The Legendre transformation is involutive. Indeed, a differentiation of equation (1.91a) with respect to \mathbf{m} yields

$$\frac{\partial \Gamma}{\partial m_i} - h_i + \sum_j \frac{\partial h_j}{\partial m_i} \frac{\partial}{\partial h_j} \left(W(\mathbf{h}) - \sum_i h_i m_i \right) \Big|_{\mathbf{m} \text{ fixed}} = 0 ,$$

and, thus,

$$h_i = \frac{\partial \Gamma}{\partial m_i} . \quad (1.92)$$

Then, comparing the derivatives of equation (1.92) with respect to \mathbf{m} with the derivatives of (1.91b) with respect to \mathbf{h} , one finds that the matrix $\partial^2 \Gamma / \partial m_i \partial m_j$ is the inverse of the matrix $\partial^2 W / \partial h_i \partial h_j$:

$$\sum_k \frac{\partial^2 W}{\partial h_i \partial h_k} \frac{\partial^2 \Gamma}{\partial m_k \partial m_j} = \delta_{ij} , \quad (1.93)$$

and, thus, the matrix $\partial^2 \Gamma / \partial m_i \partial m_j$ is also positive.

Stationarity of $W + \Gamma$. If $W(\mathbf{h})$ depends on one additional parameter μ , one finds

$$\frac{\partial \Gamma}{\partial \mu} \Big|_{\mathbf{m} \text{ fixed}} + \frac{\partial W}{\partial \mu} \Big|_{\mathbf{h} \text{ fixed}} + \sum_i \frac{\partial h_i}{\partial \mu} \frac{\partial}{\partial h_i} \Big|_{\mathbf{m} \text{ fixed}} \left(W(\mathbf{h}) - \sum_i h_i m_i \right) = 0 ,$$

and, thus,

$$\frac{\partial \Gamma}{\partial \mu} \Big|_{\mathbf{m} \text{ fixed}} + \frac{\partial W}{\partial \mu} \Big|_{\mathbf{h} \text{ fixed}} = 0 . \quad (1.94)$$

Legendre transformation and real steepest descent method. We consider the generating function $W(\mathbf{h})$ of the cumulants of a distribution e^{-S} :

$$e^{W(\mathbf{h})} = \int d^n x e^{-S(\mathbf{x}) + \mathbf{h} \cdot \mathbf{x}} . \quad (1.95)$$

We recall that, in the perturbative sense, $W(\mathbf{b})$ is the sum of connected contributions (equation (1.18)). We calculate W by the steepest descent method. The saddle point equation is

$$-\frac{\partial S(\mathbf{x})}{\partial x_i} + h_i = 0 , \quad (1.96)$$

and, in the leading order approximation,

$$W(\mathbf{h}) + S(\mathbf{x}) - \mathbf{h} \cdot \mathbf{x} = 0 . \quad (1.97)$$

Moreover, the matrix of second derivatives must be positive when the saddle point is real. Therefore, at leading order, $W(\mathbf{h})$ and $S(\mathbf{x})$ are related by a Legendre transformation and $\Gamma(\mathbf{m}) = S(\mathbf{m})$. Up to a trivial constant shift, the relation is exact in the case of Gaussian distributions.

In the sense of power series, the Legendre transformation generalizes to complex and Grassmann variables.

2 Euclidean path integrals and quantum mechanics (QM)

In most of this work, we study QM and quantum field theory (QFT) in *Euclidean formulation*. This means that we mainly discuss matrix elements of the *quantum statistical operator* $e^{-\beta H}$, where H is the quantum Hamiltonian and β the inverse temperature, rather than those of the *quantum evolution operator* $e^{-iHt/\hbar}$.

The statistical operator, which is proportional to the density matrix at thermal equilibrium $e^{-\beta H}/Z$, where $Z(\beta) = \text{tr } e^{-\beta H}$ is the quantum partition function, describes ‘evolution’ in imaginary time, and, in this sense, most of its algebraic properties are the same as those of the real time evolution operator, explicit expressions being obtained by analytic continuation $\beta \mapsto it/\hbar$ [3]. Therefore, in this chapter, to keep track of the \hbar factors of real-time evolution, we first set $\beta = t/\hbar$.

Our basic tools to study first QM, and then QFT, are *functional integrals* [4], that is, *path* [5, 6, 7] and *field integrals*.

The path integral formulation of QM is well-suited to the study of systems with an arbitrary number of degrees of freedom. It makes a smooth transition between non-relativistic QM and QFT possible.

Another property plays an essential role in this work: the Euclidean field-integral formulation emphasizes the deep connection between QFT and the statistical physics of macroscopic systems [8], in particular, systems with short-range interactions near a continuous phase transition.

The operator $e^{-\beta H}$ has a useful property: it provides a tool to study the structure of the quantum ground state. For example, if H is bounded from below, the ground state energy E_0 is given by

$$E_0 = - \lim_{\beta \rightarrow \infty} \frac{1}{\beta} \ln \text{tr } e^{-\beta H}. \quad (2.1)$$

In addition, if the ground state is unique and isolated, $e^{-\beta H}$ projects, for β large, onto the ground state vector $|0\rangle$ (in the quantum bra-ket notation):

$$e^{-\beta H} \underset{\beta \rightarrow \infty}{\sim} e^{-\beta E_0} |0\rangle \langle 0|. \quad (2.2)$$

Therefore, the corresponding Euclidean functional integral often leads to a simple and intuitive understanding of the structure of the ground state of systems with a large number of degrees of freedom. Moreover, it gives a natural interpretation to barrier penetration effects in the semi-classical approximation.

Also, it is generally easier to define mathematically the path integral representing the operator $e^{-\beta H}$ (the Feynman–Kac formula) than $e^{-iHt/\hbar}$.

The main disadvantage of the Euclidean presentation of QM is that classical expressions have a somewhat unusual form, because time is imaginary. We shall speak of *Euclidean action*, *Euclidean Lagrangian*, and *Euclidean time*.

In this chapter, first we derive the path integral representation of the matrix elements of the quantum statistical operator for Hamiltonians of the simple separable form $p^2/2m + V(q)$. The path integral makes it possible to define a functional measure and, correspondingly, expectation values called correlation functions, which are generalized moments and related to quantum observables, after an analytic continuation in time.

Note that, in the Euclidean formulation, when the potential becomes time dependent, the direct quantum interpretation is lost (going back to real time, one finds a potential depending on an imaginary time) while the path integral may still have an interpretation in the framework of one-dimensional classical statistical physics.

We calculate explicitly the path integral corresponding to the Euclidean action of harmonic oscillators to which is added a time-dependent external force. The result can then be used to generate Gaussian correlation functions, but also to reduce the evaluation of path integrals, in the case of analytic potentials, to perturbation theory.

We show, in a first example, that path integrals are especially well-suited to the study of the classical limit, by relating quantum and classical partition functions. We use the semi-classical approximation of the partition function to derive Bohr–Sommerfeld’s quantization condition.

In the Appendix, we describe a useful representation of the two-point function.

2.1 Markovian evolution and locality

Markovian evolution. Let $U(t, t')$, $t > t'$, be a bounded continuous operator in Hilbert space, which describes evolution between two times t' and t , and satisfies the Markov property,

$$U(t, t'')U(t'', t') = U(t, t'), \quad t \geq t'' \geq t', \quad \text{with} \quad U(t', t') = \mathbf{1}. \quad (2.3)$$

This property is also characteristic of the kind of stochastic processes we examine in Chapter 34, and implies that the evolution between t'' and time t depends only on the state of a system at time t'' , but not on the details of preceding evolution.

We further assume that $U(t, t')$ is differentiable with a continuous derivative and set

$$\left. \frac{\partial U(t, t')}{\partial t} \right|_{t=t'} = -K(t)/\hbar.$$

Differentiating equation (2.3) with respect to t , in the limit $t'' = t$, one obtains,

$$\hbar \frac{\partial U}{\partial t}(t, t') = -K(t)U(t, t'). \quad (2.4)$$

The Markov property (2.3) makes it possible to express the operator $U(t'', t')$ as a product of operators corresponding to arbitrarily small time intervals $\varepsilon = (t'' - t')/n$. Indeed,

$$U(t'', t') = \prod_{m=1}^n U[t' + m\varepsilon, t' + (m-1)\varepsilon], \quad \text{with} \quad n\varepsilon = t'' - t', \quad (2.5)$$

where the product is time ordered according to the rule (2.3).

The operator K . So far, the arguments are rather general. We now specialize the operator K . If K is anti-Hermitian, of the form $K = iH$, where H is a quantum Hamiltonian, then U is unitary and has the form of a quantum evolution operator. Moreover, when the operator K is time independent, it is the generator of time translations, and the formal solution of equation (2.4) is $U(t, t') = e^{-i(t-t')H/\hbar}$.

By contrast, in this chapter, we mostly choose K Hermitian and positive. Moreover, if K is time independent, and of the form of a quantum Hamiltonian H , and if we set $t - t' = \hbar\beta$, then U is a quantum statistical operator of the form $e^{-\beta H}$, proportional to the density matrix at thermal equilibrium at temperature $T = 1/\beta$.

In this case, we still call the variable t time (or Euclidean time), although from the point of view of quantum evolution it is an imaginary time. Indeed, the continuation $t \mapsto it$ formally transforms the statistical into the evolution operator.

When K is time dependent, the direct connection with quantum physics is lost. However, the formalism can still be applied to classical statistical physics in the continuum, where t is a position and the operator $U(t + \varepsilon, t)$ the analogue of the transfer matrix.

Nevertheless, the algebraic part of the calculations that follow applies to all situations.

Note that, in Chapter 34, other operators (Fokker–Planck Hamiltonians) appear that, in general are neither Hermitian nor anti-Hermitian.

Position operator and matrix elements. We introduce a distinguished basis, the basis in which the quantum position operator \hat{q} is diagonal. Using the standard QM bra–ket notation, we denote by $|q\rangle$ the eigenvector of the quantum operator \hat{q} with eigenvalue q .

In terms of matrix elements, equation (2.5) becomes

$$\langle q'' | U(t'', t') | q' \rangle = \int \prod_{k=1}^{n-1} dq_k \prod_{k=1}^n \langle q_k | U(t_k, t_{k-1}) | q_{k-1} \rangle, \quad (2.6)$$

with the conventions

$$t_k = t' + k\varepsilon, \quad q_0 = q', \quad q_n = q''.$$

In the expression, we take the limit $n \rightarrow \infty$ at $n\varepsilon$ fixed, reducing the calculation of $\langle q'' | U(t'', t') | q' \rangle$ to the asymptotic evaluation of the matrix elements $\langle q | U(t + \varepsilon, t) | q' \rangle$ for $\varepsilon \rightarrow 0$.

Locality of short-time evolution. If the operator K is local in the basis in which the position operator \hat{q} is diagonal, which means that its matrix elements $\langle q_1 | K(t) | q_2 \rangle$ have a support restricted to $q_1 = q_2$, then, for $\varepsilon \rightarrow 0$, only the matrix elements of $\langle q' | U | q \rangle$ with $|q - q'|$ small contribute significantly to expression (2.6), and one can define a *path integral representation* of $\langle q'' | U(t'', t') | q' \rangle$.

In particular, this property holds for operators of the form $K(t) \equiv K(\hat{p}, \hat{q}; t)$, where \hat{p} and \hat{q} are the momentum and position quantum operators with the commutation relation $[\hat{q}, \hat{p}] = i\hbar$, such that the classical quantity $K(p, q; t)$ is polynomial in p and analytic in q .

2.2 Statistical operator: Path integral representation

We apply the preceding considerations to the special class of local Hamiltonians of the separable form:

$$H = \hat{\mathbf{p}}^2/2m + V(\hat{\mathbf{q}}, t), \quad [\hat{q}_\alpha, \hat{p}_\beta] = i\hbar\delta_{\alpha\beta}, \quad (2.7)$$

where $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ are d -component quantum operators, and m is the mass of a particle. We specialize to the separable form, because the quantization of the corresponding classical Hamiltonian,

$$H_{\text{cl.}} = \mathbf{p}^2/2m + V(\mathbf{q}, t), \quad (2.8)$$

does not involve the problem of products of non-commuting operators. More general operators are discussed starting with Chapter 3.

We have seen that, to calculate the matrix elements of the evolution operator, we need to evaluate them only for short time intervals, with enough precision, and then use equation (2.6).

2.2.1 Short-time evolution

In the $|\mathbf{q}\rangle$ basis, equation (2.4), expressed in terms of the matrix elements $\langle \mathbf{q}|U|\mathbf{q}'\rangle$, takes the form of a Schrödinger equation in imaginary time:

$$-\hbar \frac{\partial}{\partial t} \langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle = \left[-\frac{\hbar^2}{2m} \nabla_{\mathbf{q}}^2 + V(\mathbf{q},t) \right] \langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle, \quad (2.9)$$

($\nabla_{\mathbf{q}} \equiv (\partial/\partial q_1, \dots, \partial/\partial q_d)$) with the boundary condition ($\delta^{(d)}$ is the d -dimensional Dirac function),

$$\langle \mathbf{q}|U(t',t')|\mathbf{q}'\rangle = \delta^{(d)}(\mathbf{q} - \mathbf{q}').$$

When the potential V vanishes, $\langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle$ is obtained by a Fourier transformation:

$$\begin{aligned} \langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle &= \int \frac{d^d p}{(2\pi\hbar)^d} \exp \left[\frac{1}{\hbar} (i(\mathbf{q} - \mathbf{q}') \cdot \mathbf{p} - (t - t')\mathbf{p}^2/2m) \right] \\ &= \left(\frac{m}{2\pi\hbar(t - t')} \right)^{d/2} \exp \left[-\frac{1}{\hbar} \frac{m(\mathbf{q} - \mathbf{q}')^2}{2(t - t')} \right]. \end{aligned} \quad (2.10)$$

To solve equation (2.9) in the small $\varepsilon = t - t'$ limit, we set

$$\langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle = \exp[-\sigma(\mathbf{q}, \mathbf{q}'; t, t')/\hbar].$$

From equation (2.9), we infer

$$\frac{\partial \sigma}{\partial t} = -\frac{1}{2m} (\nabla_{\mathbf{q}} \sigma)^2 + V(\mathbf{q}, t) + \frac{\hbar}{2m} \nabla_{\mathbf{q}}^2 \sigma.$$

The function σ is dominated by the free contribution. We thus expand it for ε small as,

$$\sigma(\mathbf{q}, \mathbf{q}'; t, t') = m \frac{(\mathbf{q} - \mathbf{q}')^2}{2\varepsilon} + \frac{d}{2} \hbar \ln(2\pi\hbar\varepsilon/m) + \sigma_1(\mathbf{q}, \mathbf{q}'; t, t') + o(\varepsilon), \quad (2.11)$$

where $\sigma_1 = O(\varepsilon)$. Neglecting higher-order terms, we obtain the equation ($\partial_t \equiv \partial/\partial t$),

$$[(t - t')\partial_t + (\mathbf{q} - \mathbf{q}') \cdot \nabla_{\mathbf{q}}] \sigma_1 = (t - t')V(\mathbf{q}, t). \quad (2.12)$$

We introduce the linear trajectory

$$\mathbf{q}(\tau) = \mathbf{q}' + \frac{\tau - t'}{t - t'} (\mathbf{q} - \mathbf{q}'), \quad (2.13)$$

which goes from \mathbf{q}' to \mathbf{q} at constant velocity and note that it is a solution of the homogeneous part of equation (2.12) as a function of q and t at τ fixed. The solution of equation (2.12) can then be written as

$$\sigma_1(\mathbf{q}, \mathbf{q}'; t, t') = \int_{t'}^t d\tau V(\mathbf{q}(\tau), \tau). \quad (2.14)$$

The free contribution can also be expressed in terms of the trajectory (2.13) as

$$\frac{1}{2} m(\mathbf{q} - \mathbf{q}')^2 / (t - t') = \frac{1}{2} \int_{t'}^t d\tau m \dot{\mathbf{q}}^2(\tau)$$

($\dot{q} \equiv dq/dt$). This leads to the expression

$$\langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle = \left(\frac{m}{2\pi\hbar\varepsilon} \right)^{d/2} \exp \left\{ -\frac{1}{\hbar} \int_{t'}^t d\tau \left[\frac{1}{2} m \dot{\mathbf{q}}^2(\tau) + V(\mathbf{q}(\tau), \tau) \right] + o(\varepsilon) \right\}. \quad (2.15)$$

It can be verified that the normalization is such that, for $\varepsilon \rightarrow 0$, $\langle \mathbf{q}|U(t,t')|\mathbf{q}'\rangle$ has $\delta^{(d)}(\mathbf{q} - \mathbf{q}')$ as a limit.

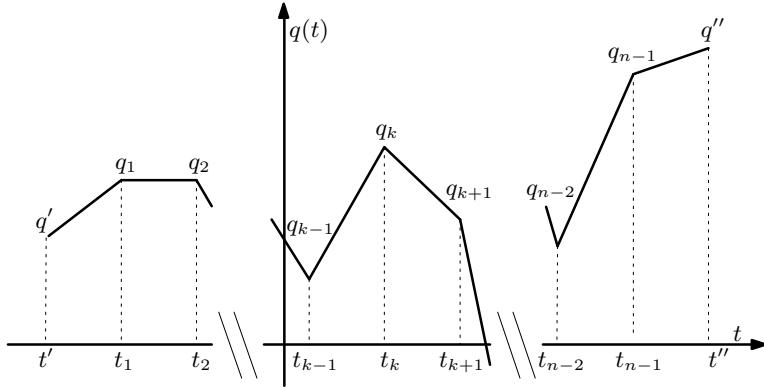


Fig. 2.1 A continuous, piecewise, linear path contributing to the integral (2.16)

2.2.2 The path integral

From equation (2.6), one derives

$$\langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle = \lim_{n \rightarrow \infty} \left(\frac{m}{2\pi\hbar\varepsilon} \right)^{dn/2} \int \prod_{k=1}^{n-1} d^d q_k \exp [-S(\mathbf{q}, \varepsilon)/\hbar], \quad (2.16)$$

with

$$S(\mathbf{q}, \varepsilon) = \int_{t'}^{t''} dt \left[\frac{1}{2} m \dot{\mathbf{q}}^2(t) + V(\mathbf{q}(t), t) \right] + no(\varepsilon), \quad (2.17)$$

where $\mathbf{q}(t)$ now is the piecewise, linear, continuous trajectory defined by (Fig. 2.1),

$$\mathbf{q}(t) = \mathbf{q}_k + \frac{t - t_k}{t_{k+1} - t_k} (\mathbf{q}_{k+1} - \mathbf{q}_k), \quad \text{for } t_k \leq t \leq t_{k+1}.$$

In terms of the function $\mathbf{q}(t)$, which interpolates in time the variables $\mathbf{q}_k \equiv \mathbf{q}(t_k)$, the integral over the variables \mathbf{q}_k is also the integral over points of the path $\mathbf{q}(t)$.

One can then verify that higher orders in ε in equation (2.17) give vanishing contributions in the small ε , $n\varepsilon$ fixed, limit and, therefore,

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ \text{at } n\varepsilon \text{ fixed}}} S(\mathbf{q}, \varepsilon) = S(\mathbf{q}) \equiv \int_{t'}^{t''} dt \left[\frac{1}{2} m \dot{\mathbf{q}}^2(t) + V(\mathbf{q}(t), t) \right]. \quad (2.18)$$

We call $S(\mathbf{q})$ *Euclidean action*. It is expressed as a time integral of the *Euclidean Lagrangian* (the Lagrangian in imaginary time when $V(q)$ does not depend on time explicitly), Legendre transforms of the Hamiltonian (2.8) (see Section 1.8).

The continuum limit of expression (2.16) can thus, formally, be written as

$$\langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle = \int_{\mathbf{q}(t')=\mathbf{q}'}^{\mathbf{q}(t'')=\mathbf{q}''} [d\mathbf{q}(t)] \exp [-S(\mathbf{q})/\hbar]. \quad (2.19)$$

The right-hand side is called *path integral*, because the integral involves a summation over all paths satisfying the prescribed boundary conditions, with a weight $\exp [-S/\hbar]$.

Notation. Note that we shall always write the integration measure $[d\mathbf{q}(t)]$ with brackets to distinguish path integrals from ordinary integrals.

Moreover, in the symbol $[d\mathbf{q}(t)]$ is buried an infinite normalization factor,

$$\mathcal{N} = \left(\frac{m}{2\pi\hbar\varepsilon} \right)^{dn/2}. \quad (2.20)$$

Therefore, we calculate only *expectation values*, which amounts to normalizing a path integral by dividing it by a reference path integral with the same kinetic term, the free motion $V \equiv 0$, or the harmonic oscillator, for example.

Brownian trajectories. The most singular term in σ (equation (2.11)) for $\varepsilon \rightarrow 0$ is $m(q - q')^2/2\varepsilon$ (independently of the potential). The support of the matrix element $\langle \mathbf{q}' | U(t, t') | \mathbf{q}' \rangle$ is thus restricted to $|\mathbf{q}' - \mathbf{q}| = O(\sqrt{\varepsilon})$. For $|\mathbf{q}' - \mathbf{q}| = O(\sqrt{\varepsilon})$,

$$\sigma_1(\mathbf{q}, \mathbf{q}'; t') = \varepsilon V((\mathbf{q} + \mathbf{q}')/2, t') + O(\varepsilon^2) = \frac{1}{2}\varepsilon(V(\mathbf{q}) + V(\mathbf{q}')) + O(\varepsilon^2) = \varepsilon V(\mathbf{q}, t) + O(\varepsilon^{3/2}),$$

because the potential is assumed to be analytic (we could generalize to piecewise-analytic potentials). Hence, a replacement of σ_1 in the expression (2.14), for instance, by $\varepsilon V(\mathbf{q}, t)$, modifies σ by a contribution of order $\varepsilon^{3/2}$, which is negligible.

This support property shows that in the action (2.18) the two terms play quite different roles. The kinetic term $\int dt \dot{\mathbf{q}}^2$ selects the class of paths contributing to the path integral, those for which $[\mathbf{q}(t + \varepsilon) - \mathbf{q}(t)]^2/\varepsilon$ remains finite when ε goes to 0. More precisely, one finds that the expectation value of $[\mathbf{q}(t + \varepsilon) - \mathbf{q}(t)]^2$ is proportional to $|\varepsilon|\hbar/m$ for $\varepsilon \rightarrow 0$ (see Section A2.2 for a proof in the continuum limit). The kinetic term really is a part of the functional measure. It is essential to the very existence of the path integral.

By contrast, the potential weights paths according to the value of $\mathbf{q}(t)$ at each time, and determines the physical properties of the theory.

Brownian paths and leading contributions. The paths contributing to the path integral are typical of the *Brownian motion* (for details see Chapter 34); in particular, they are continuous but not differentiable, in contrast with what the formal expression (2.18) suggests. Still, the notation (2.18) is useful because the paths that give the largest contributions to the path integral (2.19) are in the neighbourhood of the paths that minimize the action (2.18) and, thus, solutions of the classical, Euclidean equations of motion. These paths are such that (functional derivatives are defined in Section 2.5.3)

$$\frac{\delta\mathcal{S}}{\delta q_i(t)} = 0,$$

and the operator associated with the kernel $\delta^2\mathcal{S}/\delta q_i(t_1)\delta q_j(t_2)$ is positive. These paths are classical, differentiable Euclidean paths. This observation is at the basis of *semi-classical approximations*.

Quantum evolution and path integral. When the potential is time independent, the path integral representing the matrix elements of the time-evolution operator can be recovered by an analytic continuation to real time $t \mapsto it$. One then formally obtains,

$$\langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle = \int_{\mathbf{q}(t')=\mathbf{q}'}^{\mathbf{q}(t'')=\mathbf{q}''} [d\mathbf{q}(t)] \exp(i\mathcal{A}(\mathbf{q})/\hbar), \quad (2.21)$$

where \mathcal{A} is the classical action, time integral of the Lagrangian,

$$\mathcal{A}(\mathbf{q}) \equiv \int_{t'}^{t''} dt \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \quad \text{with} \quad \mathcal{L}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = \frac{1}{2}m\dot{\mathbf{q}}^2(t) - V(\mathbf{q}(t)). \quad (2.22)$$

2.3 Explicit evaluation of a path integral: The harmonic oscillator

This work will exhibit later numerous examples of situations where functional integrals play an essential role. Here, we begin with a simple problem, the calculation of the path integral associated with the Hamiltonian of the one-dimensional quantum harmonic oscillator. This example illustrates that Gaussian path integrals can be calculated explicitly in the continuum, without returning to the limiting process involving discrete time intervals.

We consider the quantum Hamiltonian ($\omega > 0$ constant),

$$H = \frac{1}{2m}\hat{p}^2 + \frac{1}{2}m\omega^2\hat{q}^2. \quad (2.23)$$

The corresponding Euclidean classical action is

$$\mathcal{S}_0(q) = \int_{t'}^{t''} \left[\frac{1}{2}m\dot{q}^2(t) + \frac{1}{2}m\omega^2q^2(t) \right] dt, \quad (2.24)$$

which is a quadratic form in $q(t)$. When inserted into equation (2.19), the expression leads to the Gaussian integral,

$$\langle q'' | U_0(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [dq(t)] \exp [-\mathcal{S}_0(q)/\hbar]. \quad (2.25)$$

First, to determine the dependence on the boundary conditions q', q'' , we change variables $q(t) \mapsto r(t)$ (a constant shift at fixed t) setting,

$$q(t) = q_c(t) + r(t),$$

in which q_c is the solution of the classical (Euclidean) equation of motion,

$$m\ddot{q}_c(t) - m\omega^2q_c(t) = 0, \quad \text{with } q_c(t') = q', \quad q_c(t'') = q''. \quad (2.26)$$

The new path $r(t)$ then satisfies the boundary conditions

$$r(t') = r(t'') = 0. \quad (2.27)$$

The action becomes

$$\mathcal{S}_0(q) = \mathcal{S}_0(q_c) + \mathcal{S}_0(r) + m \int_{t'}^{t''} [\dot{q}_c(t)\dot{r}(t) + \omega^2q_c(t)r(t)] dt.$$

Integrating by parts, $\int \dot{q}_c \dot{r} = r\dot{q}_c - \int r\ddot{q}_c$, taking into account equation (2.26) and the boundary conditions (2.27), one notes that the terms linear in r cancel and the action (2.24) reduces to

$$\mathcal{S}_0(q) = \mathcal{S}_0(q_c) + \mathcal{S}_0(r).$$

One obtains

$$\langle q'' | U_0(t'', t') | q' \rangle = \mathcal{N}(\omega, \tau) e^{-\mathcal{S}_0(q_c)/\hbar}, \quad (2.28)$$

where

$$\mathcal{N}(\omega, \tau) = \int [dr(t)] \exp \left[-\frac{m}{2\hbar} \int_{t'}^{t''} dt (\dot{r}^2(t) + \omega^2 r^2(t)) \right], \quad (2.29)$$

with $r(t') = r(t'') = 0$.

Evaluation of the classical action. Setting $\tau = t'' - t'$, one can write the solution of equation (2.26) as

$$q_c(t) = \frac{1}{\sinh(\omega\tau)} [q' \sinh(\omega(t'' - t)) + q'' \sinh(\omega(t - t'))]. \quad (2.30)$$

To calculate $\mathcal{S}_0(q_c)$ one can again integrate by parts, $\int \dot{q}^2 dt = q\dot{q} - \int q\ddot{q} dt$, and use the equation of motion (2.26). The result is

$$\mathcal{S}_0(q_c) = \frac{m\omega}{2 \sinh \omega\tau} [(q'^2 + q''^2) \cosh \omega\tau - 2q'q'']. \quad (2.31)$$

The remaining path integral. To complete the calculation, one still must evaluate the last Gaussian integral over $r(t)$. The integral no longer depends on q', q'' and yields a normalization factor, function only of ω and $t'' - t'$. Since this involves an infinite normalization factor, we postpone the calculation until Section 2.6.1. The final result is

$$\begin{aligned} \langle q'' | U_0(t'', t) | q' \rangle &= \left(\frac{m\omega}{2\pi\hbar \sinh \omega\tau} \right)^{1/2} \\ &\times \exp \left\{ -\frac{m\omega}{2\hbar \sinh \omega\tau} [(q'^2 + q''^2) \cosh \omega\tau - 2q'q''] \right\}. \end{aligned} \quad (2.32)$$

2.4 Partition function: Classical and quantum statistical physics

In this section, Hamiltonians are assumed to be *time independent* and to have a discrete spectrum.

2.4.1 The quantum partition function

The quantum partition function $\mathcal{Z}(\beta) = \text{tr } e^{-\beta H}$ has a path integral representation that is immediately inferred from the representation of the statistical operator. One finds

$$\begin{aligned} \mathcal{Z}(\beta) &= \text{tr } e^{-\beta H} \equiv \text{tr } U(\hbar\beta/2, -\hbar\beta/2) = \int dq \langle q | U(\hbar\beta/2, -\hbar\beta/2) | q \rangle \\ &= \int [dq(t)] \exp [-\mathcal{S}(q)/\hbar], \end{aligned} \quad (2.33)$$

where the paths now satisfy periodic boundary conditions: $q(-\hbar\beta/2) = q(\hbar\beta/2)$, and one integrates over all values of $q(\hbar\beta/2)$. It is actually convenient to rescale time $t \mapsto t/\hbar$. The action then reads

$$\mathcal{S}(q)/\hbar = \int_{-\beta/2}^{\beta/2} dt \left[\frac{1}{2} m \dot{q}^2(t)/\hbar^2 + V(\mathbf{q}(t)) \right]. \quad (2.34)$$

The harmonic oscillator. Equation (2.28) makes it possible to relate the normalization of the path integral to the partition function $\mathcal{Z}_0(\beta)$ of the harmonic oscillator. Taking the trace of $U_0(\hbar\beta/2, -\hbar\beta/2)$, one finds

$$\begin{aligned} \mathcal{Z}_0(\beta) &= \text{tr } U_0(\hbar\beta/2, -\hbar\beta/2) \equiv \int dq \langle q | U_0(\hbar\beta/2, -\hbar\beta/2) | q \rangle \\ &= \mathcal{N}(\omega, \beta) \left(\frac{\pi\hbar}{m\omega \tanh(\beta\hbar\omega/2)} \right)^{1/2}. \end{aligned} \quad (2.35)$$

The large β limit. One can also set the boundary conditions at $t = 0$ and $t = \beta$. Since the action is time-translation invariant the result is the same. However, in the formal large β limit (relevant for the ground-state energy), in the first case one obtains an explicitly time-translation invariant formalism on the whole real line while in the second case, one is led to integrate over paths on the positive real line with the boundary condition $q(0) = 0$. The first formalism is clearly simpler.

2.4.2 Classical and quantum statistical physics

We consider a *classical* one-dimensional lattice model, where the real variable q_k characterizes the configuration at site k . We choose a lattice with n sites and periodic boundary conditions ($q_n = q_0$). The partition function is defined by

$$\mathcal{Z}(n, \varepsilon) = \int \left(\prod_{k=1}^n e^{-\varepsilon V(q_k)} dq_k \right) \exp [-\mathcal{E}(q, \varepsilon)], \quad (2.36)$$

where the configuration energy, divided by the temperature, is given by

$$\mathcal{E}(q, \varepsilon) = \frac{1}{2\varepsilon} \sum_{k=1}^n (q_k - q_{k-1})^2. \quad (2.37)$$

The partition function can also be considered as a time-discretized form of the path integral (2.33). The parameter ε plays, somewhat, the role of the temperature, and $V(q)$ determines the distribution of q on each site.

We introduce the function

$$S(q, q') = \frac{1}{2\varepsilon} (q - q')^2 + \frac{1}{2} \varepsilon V(q) + \frac{1}{2} \varepsilon V(q'), \quad (2.38)$$

and the kernel

$$\mathcal{T}(q, q') = e^{-S(q, q')}.$$

Using the bra-ket notation of QM, one can express the kernel in terms of elements of a *transfer matrix* \mathbf{T} , which has also the form of a quantum operator,

$$\mathcal{T}(q, q') = \langle q' | \mathbf{T} | q \rangle \quad (2.39)$$

In terms of \mathbf{T} , the partition function (2.36) can be written as

$$\mathcal{Z}(n, \varepsilon) = \text{tr } \mathbf{T}^n. \quad (2.40)$$

In one dimension, one expects a non-trivial collective behaviour only at low temperatures, where $\varepsilon \rightarrow 0$. Then, one finds

$$\mathbf{T} \underset{\varepsilon \rightarrow 0}{\sim} \sqrt{2\pi\varepsilon} e^{-\varepsilon \mathbf{H}}, \quad (2.41)$$

where \mathbf{H} is a quantum Hamiltonian, which in terms of the quantum position and momentum operators (with $\hbar = 1$ and $[\hat{q}, \hat{p}] = i$), takes the form

$$\mathbf{H} = \frac{1}{2} \hat{p}^2 + V(\hat{q}).$$

For $\varepsilon \rightarrow 0$, the eigenvalues and eigenvectors of the transfer matrix are related in a simple way to those of \mathbf{H} . In the thermodynamic limit $n \rightarrow \infty$, the transfer matrix is dominated by its largest eigenvalue, which corresponds to the ground state energy E_0 of \mathbf{H} . The free energy $\mathcal{W} = \ln \mathcal{Z}$ per unit length is

$$\frac{1}{n} \mathcal{W} = \frac{1}{n} \ln \mathcal{Z}(n, \varepsilon) \sim -\varepsilon E_0 + \frac{1}{2} \ln(2\pi\varepsilon).$$

On the other hand, combining equations (2.40) and (2.41), one infers that, when $\varepsilon \rightarrow 0$ with $n\varepsilon = \beta$ fixed, $\mathcal{Z}(n, \varepsilon)$ becomes proportional to the quantum partition function

$$\mathcal{Z}(\beta) = \text{tr } e^{-\beta \mathbf{H}} = \lim_{\varepsilon \rightarrow 0} (2\pi\varepsilon)^{n/2} \mathcal{Z}(n, \varepsilon).$$

We have found an interesting relation between quantum statistical physics in zero dimension (one particle) and classical statistical physics in one dimension. This relation extends to higher space dimensions: d -dimensional classical statistical physics and $(d-1)$ -dimensional quantum statistical physics. Finally, we observe that the large β limit, which is the zero-temperature limit of the quantum model, is also the thermodynamic limit of the classical model.

2.5 Correlation functions. Generating functional

In the classical statistical model (2.36), the m -point correlation function is given by

$$\langle q_{i_1} q_{i_2} \cdots q_{i_m} \rangle_n = \mathcal{Z}^{-1}(n, \varepsilon) \int \left(\prod_{k=1}^n dq_k \mathcal{T}(q_{k-1}, q_k) \right) q_{i_1} q_{i_2} \cdots q_{i_m}. \quad (2.42)$$

We assume the order $0 < i_1 \leq i_2 \leq \cdots \leq i_m < n$, and we express the m -point correlation function in terms of the transfer matrix (2.39). We then integrate over all q variables except $q_{i_1}, q_{i_2}, \dots, q_{i_m}$ and find,

$$\begin{aligned} \langle q_{i_1} q_{i_2} \cdots q_{i_m} \rangle_n &= \mathcal{Z}^{-1}(n, \varepsilon) \int \prod_{s=1}^m dq_{i_s} \langle q_{i_1} | \mathbf{T}^{n-i_m+i_1} | q_{i_m} \rangle q_{i_m} \\ &\quad \times \langle q_{i_m} | \mathbf{T}^{i_m-i_{m-1}} | q_{i_{m-1}} \rangle q_{i_{m-1}} \cdots \langle q_{i_2} | \mathbf{T}^{i_2-i_1} | q_{i_1} \rangle q_{i_1}. \end{aligned}$$

The position operator \hat{q} is diagonal in the $|q\rangle$ basis and, therefore,

$$\langle q' | \mathbf{T}^r | q \rangle q = \langle q' | \mathbf{T}^r \hat{q} | q \rangle.$$

We conclude the n -point function can be rewritten as,

$$\langle q_{i_1} q_{i_2} \cdots q_{i_m} \rangle_n = \mathcal{Z}^{-1}(n, \varepsilon) \text{tr } \mathbf{T}^{n-i_m+i_1} \hat{q} \mathbf{T}^{i_m-i_{m-1}} \hat{q} \cdots \mathbf{T}^{i_2-i_1} \hat{q}.$$

For $\varepsilon \rightarrow 0$, we can introduce the Hamiltonian \mathbf{H} . It is also convenient to change notation, associating discrete values of a continuous variable t to site positions by $t_k = \varepsilon i_k$ (and $\beta = n\varepsilon$). We then obtain

$$\begin{aligned} Z^{(m)}(t_1, t_2, \dots, t_m) &\equiv \langle q(t_1) q(t_2) \cdots q(t_m) \rangle_\beta \\ &\stackrel{\varepsilon \rightarrow 0}{\sim} \mathcal{Z}^{-1}(\beta) \text{tr} \left[e^{-(\beta-t_n+t_1)\mathbf{H}} \hat{q} e^{-(t_m-t_{m-1})\mathbf{H}} \hat{q} \cdots e^{-(t_2-t_1)\mathbf{H}} \hat{q} \right]. \end{aligned} \quad (2.43)$$

From the direct definition (2.42), it follows that correlation functions have, for $\varepsilon \rightarrow 0$, $n\varepsilon = \beta$, $\varepsilon i_k = t_k$ fixed, the path integral representation

$$\langle q(t_1) q(t_2) \cdots q(t_m) \rangle_\beta = \mathcal{Z}^{-1}(\beta) \int [dq] q(t_1) \cdots q(t_m) e^{-\mathcal{S}(q)}, \quad (2.44)$$

with

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} dt \left[\frac{1}{2} \dot{q}^2(t) + V(q(t)) \right].$$

Note here that the analogy between classical and quantum statistical physics is not complete. Indeed, the classical correlation functions have no direct quantum analogue. Only the expectation values $\langle q^m(t) \rangle$ are both quantum and classical observables. However, this does not prevent the analogy between quantum and classical systems to extend to a less trivial subset of correlation functions in higher space dimensions, in the case of systems isotropic in space.

Moreover, the analytic continuations of the classical correlation functions $t_k \mapsto it_k$ are the dynamic correlation functions of the quantum theory.

2.5.1 Thermodynamic limit

In the classical statistical model, the limit $\beta \rightarrow \infty$ (zero temperature of the quantum model) is the thermodynamic limit. Assuming that the ground state is unique (this is always true in QM with a finite number of degrees of freedom and a regular potential), one finds

$$\langle q(t_1)q(t_2)\cdots q(t_m) \rangle = \langle 0 | \hat{q} e^{-(t_m-t_{m-1})(\mathbf{H}-E_0)} \hat{q} \cdots e^{-(t_2-t_1)(\mathbf{H}-E_0)} \hat{q} | 0 \rangle. \quad (2.45)$$

In particular, the expectation value of $q(t)$ then is

$$\langle q(t) \rangle = \langle 0 | \hat{q} | 0 \rangle.$$

The connected two-point correlation function (the generalized second cumulant) is defined by

$$W^{(2)}(t_1, t_2) = \langle (q(t_1) - \langle q \rangle)(q(t_2) - \langle q \rangle) \rangle = \langle q(t_1)q(t_2) \rangle - \langle q \rangle^2. \quad (2.46)$$

Then,

$$W^{(2)}(t_1, t_2) = \langle 0 | \hat{q} e^{-(t_2-t_1)(\mathbf{H}-E_0)} \hat{q} | 0 \rangle - \langle 0 | \hat{q} | 0 \rangle^2.$$

Denoting by E_k , $E_0 < E_1 < E_2 \cdots$ the successive eigenvalues of \mathbf{H} , and $|k\rangle$ the corresponding eigenvectors, one finds for large separations $|t_2 - t_1| \rightarrow \infty$,

$$W^{(2)}(t_1, t_2) = e^{-|t_2-t_1|(E_1-E_0)} (\langle 0 | \hat{q} | 1 \rangle)^2 + O(e^{-|t_2-t_1|(E_2-E_0)}).$$

The connected two-point function decreases exponentially with a rate which, in lattice units, is

$$W_{i_1 i_2}^{(2)} \propto e^{-\varepsilon(E_1-E_0)|i_2-i_1|}.$$

The decay of the two-point function is traditionally characterized by the *correlation length* ξ . Here, one finds

$$\xi = \frac{1}{E_1 - E_0} \frac{1}{\varepsilon}.$$

In the continuum limit $\varepsilon \rightarrow 0$, the correlation length in lattice units diverges. Keeping the variables t_1, t_2, \dots, β fixed when $\varepsilon \rightarrow 0$, corresponds to measuring distances on the lattice in correlation length units, that is, in macroscopic units.

The existence of a non-trivial large-distance physics is the direct consequence of the divergence of the correlation length. A continuum limit can thus be defined, which is somewhat *universal* in the sense that it does not depend on the initial lattice structure, and on the precise way time is discretized. This is the first example of a situation we encounter again in the context of second order phase transitions.

Ground state degeneracy. With the conditions we have assumed, the ground state is unique. The existence and some properties of phase transitions, which we start discussing in Chapter 14, are related to a possible ground state degeneracy.

2.5.2 Generating functional of correlation functions

We consider a path integral defined in terms of the Euclidean action $\mathcal{S}(q)$, to which we add a coupling to an external force [9],

$$\mathcal{S}(q, b) = \mathcal{S}(q) - \int dt b(t)q(t),$$

and satisfying periodic boundary conditions, $q(-\beta/2) = q(\beta/2)$. Then,

$$\mathcal{Z}(b, \beta) \equiv \text{tr } U(\hbar\beta/2, -\hbar\beta/2) = \int [dq(t)] e^{-\mathcal{S}(q,b)/\hbar}. \quad (2.47)$$

The quantity $\mathcal{Z}(b, \beta)$ is the *generating functional* of correlation functions: if one expands the path integral in powers of $b(t)$, one finds

$$\mathcal{Z}(b, \beta) = \mathcal{Z}(0, \beta) \sum_{n=0}^{\infty} \frac{1}{n!} \int dt_1 \cdots dt_n Z^{(n)}(t_1, \dots, t_n) b(t_1) \cdots b(t_n), \quad (2.48)$$

where

$$Z^{(n)}(t_1, \dots, t_n) \equiv \langle q(t_1) \cdots q(t_n) \rangle = \frac{1}{\mathcal{Z}(0, \beta)} \int [dq(t)] e^{-\mathcal{S}(q,0)/\hbar} q(t_1) \cdots q(t_n). \quad (2.49)$$

2.5.3 Functional differentiation and correlation functions

In this chapter, and in many places in the work, we use functional derivatives. We define them here.

To recover the correlation functions $Z^{(n)}$ (equation (2.48)) from the generating functional $\mathcal{Z}(b)$, we can use the *functional differentiation* operator $\delta/\delta b(t)$. Functional differentiation obeys the standard algebraic rules (linearity and Leibniz's rule):

$$\begin{aligned} \frac{\delta}{\delta(t)} [\mathcal{Z}_1(b) + \mathcal{Z}_2(b)] &= \frac{\delta}{\delta b(t)} \mathcal{Z}_1(b) + \frac{\delta}{\delta b(t)} \mathcal{Z}_2(b), \\ \frac{\delta}{\delta b(t)} [\mathcal{Z}_1(b) \mathcal{Z}_2(b)] &= \mathcal{Z}_1(b) \frac{\delta}{\delta b(t)} \mathcal{Z}_2(b) + \mathcal{Z}_2(b) \frac{\delta}{\delta b(t)} \mathcal{Z}_1(b), \end{aligned} \quad (2.50)$$

and, in addition,

$$\frac{\delta}{\delta b(t)} b(u) = \delta(t - u), \quad (2.51)$$

where $\delta(x)$ is Dirac's δ -function.

For example, a functional differentiation with respect to $b(t)$ of $\mathcal{Z}(b, \beta)$ (equation (2.47)), yields

$$\hbar \frac{\delta}{\delta b(t_1)} \mathcal{Z}(b, \beta) = \int [dq] q(t_1) \exp [-\mathcal{S}(q, b)/\hbar].$$

Therefore, by differentiating p times with respect to $b(t)$, one can generate any product $q(t_1) \cdots q(t_p)$. Taking the $b \equiv 0$ limit, one obtains correlation functions corresponding to the Euclidean action $\mathcal{S}(q)$,

$$\hbar^p \prod_{j=1}^p \frac{\delta}{\delta b(t_j)} \mathcal{Z}(b, \beta) \Big|_{b \equiv 0} = \int [dq] \prod_{j=1}^p q(t_j) \exp [-\mathcal{S}(q)/\hbar] \quad (2.52a)$$

$$\equiv \mathcal{Z}(b \equiv 0, \beta) \langle q(t_1) q(t_2) \cdots q(t_p) \rangle. \quad (2.52b)$$

More generally, a differential operator $\mathcal{F}(\hbar\delta/\delta b(t))$ generates in the right-hand side the expectation value of the functional $\mathcal{F}(q)$.

2.6 Harmonic oscillator. Correlation functions and Wick's theorem

We consider the Hamiltonian (harmonic oscillator coupled to an external force)

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2q^2 - qb(t). \quad (2.53)$$

The Hamiltonian corresponds to the Euclidean action

$$\mathcal{S}_G(q, b) = \int_{-\tau/2}^{\tau/2} dt \left[\frac{1}{2}m\dot{q}^2(t) + \frac{1}{2}m\omega^2q^2(t) - b(t)q(t) \right]. \quad (2.54)$$

We now calculate explicitly the Gaussian path integral corresponding to the action (2.54) (see equation (2.19)), with *periodic boundary conditions*,

$$\mathcal{Z}_G(b, \beta) \equiv \text{tr } U(\tau/2, -\tau/2) = \int_{q(\tau/2)=q(-\tau/2)} [dq(t)] \exp[-\mathcal{S}_G(q, b)/\hbar]. \quad (2.55)$$

This quantity is also the generating functional of statistical correlation functions (equation (2.47)) of the Gaussian model discussed in Section 2.3.

To eliminate the linear term from the action, we adapt the method followed in Section 1.1. We change variables $q(t) \mapsto r(t)$ setting

$$q(t) = q_c(t) + r(t), \quad q_c(\tau/2) = q_c(-\tau/2) \Rightarrow r(\tau/2) = r(-\tau/2). \quad (2.56)$$

The action becomes

$$\mathcal{S}_G(q, b) = \mathcal{S}_0(r) + \mathcal{S}_G(q_c, b) + \int_{-\tau/2}^{\tau/2} dt \left[m\dot{r}(t)\dot{q}_c(t) + m\omega^2r(t)q_c(t) - b(t)r(t) \right].$$

In the term linear in r , we integrate by parts, and use the periodic boundary conditions (2.56). Then,

$$\begin{aligned} \int_{-\tau/2}^{\tau/2} dt \dot{r}(t)\dot{q}_c(t) &= r(\tau/2)\dot{q}_c(\tau/2) - r(-\tau/2)\dot{q}_c(-\tau/2) - \int_{-\tau/2}^{\tau/2} dt r(t)\ddot{q}_c(t) \\ &= r(\tau/2)(\dot{q}_c(\tau/2) - \dot{q}_c(-\tau/2)) - \int_{-\tau/2}^{\tau/2} dt r(t)\ddot{q}_c(t). \end{aligned}$$

The term linear in r vanishes if the function $q_c(t)$ satisfies the classical equation

$$-\ddot{q}_c(t) + \omega^2q_c(t) = b(t)/m,$$

with the boundary condition $\dot{q}_c(\tau/2) = \dot{q}_c(-\tau/2)$. The solution can be written as

$$q_c(t) = \frac{1}{m} \int_{-\tau/2}^{\tau/2} \Delta(t-u)b(u)du,$$

where the function Δ is the solution of the equation

$$-\ddot{\Delta}(t) + \omega^2\Delta(t) = \delta(t),$$

with the periodic boundary conditions $\Delta(\tau/2) = \Delta(-\tau/2)$, $\dot{\Delta}(\tau/2) = \dot{\Delta}(-\tau/2)$.

One finds

$$\Delta(t) = \frac{1}{2\omega \sinh(\omega\tau/2)} \cosh(\omega(\tau/2 - |t|)). \quad (2.57)$$

In the limit $\tau \rightarrow \infty$, the function reduces to

$$\Delta(t) = \frac{1}{2\omega} e^{-\omega|t|}. \quad (2.58)$$

The classical action becomes

$$\begin{aligned} S_G(q_c, b) &= \int_{-\tau/2}^{\tau/2} dt [\frac{1}{2}m\dot{q}_c^2(t) + \frac{1}{2}m\omega^2 q_c^2(t) - b(t)q_c(t)] = -\frac{1}{2} \int_{-\tau/2}^{\tau/2} dt q_c(t)b(t) \\ &= -\frac{1}{2m} \int_{-\tau/2}^{\tau/2} dt du b(t)\Delta(t-u)b(u). \end{aligned}$$

The remaining integral over $r(t)$ just yields $\text{tr } U_0(\tau/2, -\tau/2)$ (equation (2.32)). After setting $\tau = \hbar\beta$, one finds (equation (2.35)),

$$\text{tr } U_0(\tau/2, -\tau/2) = \mathcal{Z}_0(\beta),$$

where $\mathcal{Z}_0(\beta)$ is the partition function of the harmonic oscillator. Therefore,

$$\frac{\mathcal{Z}_G(b, \beta)}{\mathcal{Z}_0(\beta)} = e^{-S_G(q_c, b)/\hbar} = \exp \left[\frac{1}{2m\hbar} \int_{-\hbar\beta/2}^{\hbar\beta/2} du dv \Delta(v-u)b(v)b(u) \right]. \quad (2.59)$$

2.6.1 Correlation functions, Wick's theorem

Replacing $\mathcal{Z}_G(b, \beta)$ by the explicit expression (2.59), and differentiating twice, one obtains the two-point correlation function

$$\langle q(t)q(u) \rangle_0 = \mathcal{Z}_0^{-1}(\beta)\hbar^2 \frac{\delta^2}{\delta b(t)\delta b(u)} \mathcal{Z}_G(b, \beta) \Big|_{b=0} = \frac{\hbar}{m} \Delta(t-u). \quad (2.60)$$

More generally, the arguments of Section 1.1 apply: as a characteristic property of the Gaussian measure, all correlation functions can be expressed in terms of the two-point function as stated by Wick's theorem (1.14) (ℓ must be even):

$$\langle q(t_1)q(t_2) \cdots q(t_\ell) \rangle_0 = \sum_{\substack{\text{all possible pairings} \\ P \text{ of } \{1, 2, \dots, \ell\}}} \langle q(t_{P_1})q(t_{P_2}) \rangle_0 \cdots \langle q(t_{P_{\ell-1}})q(t_{P_\ell}) \rangle_0. \quad (2.61)$$

Harmonic oscillator: The partition function. We can now determine the dependence of the partition function $\mathcal{Z}_0(\beta)$ on the parameter ω . Indeed, differentiating the path integral, one obtains

$$\frac{\partial}{\partial \omega} \ln \mathcal{Z}_0(\beta) = -\frac{m\omega}{\hbar} \int_{-\hbar\beta/2}^{\hbar\beta/2} dt \langle q^2(t) \rangle_0 = -\frac{\hbar\beta}{2} \frac{\cosh(\omega\hbar\beta/2)}{\sinh(\omega\hbar\beta/2)}. \quad (2.62)$$

Hence,

$$\mathcal{Z}_0(\beta) = \mathcal{N}' \frac{1}{\sinh(\beta\hbar\omega/2)}.$$

For dimensional reasons, \mathcal{N}' is a pure number. It can be obtained by taking the limit $\beta \rightarrow \infty$, where one should find $e^{-\beta E_0}$. The complete result, thus, is

$$\mathcal{Z}_0(\beta) = \frac{1}{2 \sinh(\beta\hbar\omega/2)} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}}, \quad (2.63)$$

which, indeed, is the partition function of the harmonic oscillator. This also completes the calculation of the normalization in expression (2.32), as a consequence of the relation (2.35).

2.6.2 Harmonic oscillator: Paths and square integrable functions

Sometimes, mainly in semi-classical calculations (solitons, instantons), an alternative method of calculation is useful, which we explain here in the example of the harmonic oscillator. Since we expect problems with infinite normalizations, we first work with the discretized form and then describe it in the continuum. We set here $\hbar = m = 1$. Then,

$$\mathcal{S}_0(q) = \sum_{k=1}^n \left[\frac{(q_k - q_{k-1})^2}{2\varepsilon} + \frac{1}{2}\varepsilon\omega^2 q_k^2 \right], \text{ with } q_0 = q_n.$$

Due to the periodic boundary conditions, the system is translation invariant, and thus the quadratic form in the variables q_k is simple to diagonalize. We introduce a discrete Fourier representation, setting

$$q_k = \frac{1}{\sqrt{n}} \sum_{\ell=0}^{n-1} e^{2i\pi k\ell/n} c_\ell, \quad (2.64)$$

with the reality conditions

$$c_0 = \bar{c}_0, \quad \bar{c}_{n-\ell} = c_\ell. \quad (2.65)$$

Then,

$$\mathcal{S}_0(q) = \sum_{\ell=0}^{n-1} \bar{c}_\ell \left[(1 - \cos(2\pi\ell/n))/\varepsilon + \frac{1}{2}\omega^2\varepsilon \right] c_\ell, \quad (2.66)$$

where the orthogonality relations

$$\frac{1}{n} \sum_{k=0}^{n-1} e^{2i\pi k\ell/n} = \begin{cases} 1 & \text{for } \ell = 0 \pmod{n}, \\ 0 & \text{otherwise,} \end{cases}$$

have been used. These relations also show that the transformation is unitary, and the Jacobian of the change of variables is a phase factor.

The integral now has the form (1.29), but the relation (2.65) implies that only about half of the complex variables are independent. One finds

$$\mathcal{Z}_0 = (2\varepsilon)^{-n/2} \left[\prod_{\ell=0}^{n-1} (1 - \cos(2\pi\ell/n))/\varepsilon + \frac{1}{2}\omega^2\varepsilon \right]^{-1/2}. \quad (2.67)$$

The product can be calculated explicitly. Setting

$$\cosh \theta = 1 + \omega^2\varepsilon^2/2,$$

one obtains

$$\prod_{\ell=0}^{n-1} \left[(1 - \cos(2\pi\ell/n))/\varepsilon + \frac{1}{2}\omega^2\varepsilon \right] = \frac{2}{(2\varepsilon)^n} (\cosh n\theta - 1).$$

In the $\varepsilon \rightarrow 0$ limit, with $n\varepsilon = \beta$ fixed, $n\theta \rightarrow \beta\omega$, and one recovers the partition function of the harmonic oscillator,

$$\mathcal{Z}_0(\beta) = \frac{e^{-\beta\omega/2}}{1 - e^{-\beta\omega}}.$$

Continuum calculation. The calculation with discrete variables suggests how to perform a continuum calculation. In the continuum limit, the change of variables (2.64) becomes the expansion of $q(t)$ on normalized periodic *square integrable functions*:

$$q(t) = \frac{1}{\sqrt{\beta}} \sum_{\ell} c_{\ell} e^{2i\pi\ell t/\beta}, \quad \text{with} \quad c_{-\ell} = \bar{c}_{\ell}.$$

The Jacobian is unity and the measure simply

$$[dq(t)] \mapsto dc_0 \prod_{\ell>0} dc_{\ell} d\bar{c}_{\ell}. \quad (2.68)$$

The function \mathcal{S}_0 becomes

$$\mathcal{S}_0 = \frac{1}{2}\omega^2 c_0^2 + \sum_{\ell \geq 1} \bar{c}_{\ell} (\omega^2 + 4\pi^2 \ell^2 / \beta^2) c_{\ell}.$$

The integration then is straightforward:

$$\mathcal{Z}_0(\beta) \propto \frac{1}{\omega} \prod_{\ell \geq 1} [(\omega^2 + 4\pi^2 \ell^2 / \beta^2)]^{-1}. \quad (2.69)$$

The infinite product diverges and must be normalized. The free Hamiltonian is not available because the partition function does not exist. However, it is possible to compare different values of ω , or calculate the derivative

$$\frac{\partial}{\partial \omega} \ln \mathcal{Z}_0(\beta) = -\frac{1}{\omega} - \sum_{\ell > 0} \frac{2\omega}{\omega^2 + 4\pi^2 \ell^2 / \beta^2} = -\frac{\beta}{2 \tanh(\omega\beta/2)},$$

which is the result (2.62).

2.7 Perturbed harmonic oscillator

We now consider the Hamiltonian (in this section we set $\hbar = 1$)

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2 + V_I(q), \quad (2.70)$$

where we assume only that $V_I(q)$ is expandable in powers of q :

$$V_I(q) = \sum_n v_n q^n.$$

The corresponding partition function is given by

$$\mathcal{Z}(\beta) = \int_{q(-\beta/2)=q(\beta/2)} [dq] \exp \left\{ - \int_{-\beta/2}^{\beta/2} [\frac{1}{2}\dot{q}^2(t) + \frac{1}{2}\omega^2 q^2(t) + V_I(q(t))] dt \right\}. \quad (2.71)$$

To evaluate the path integral (2.71), one can generalize the identities (1.15) and (1.17). We apply identity (2.52a) to the path integral (2.55):

$$\mathcal{Z}(\beta) = \left\{ \exp \left[- \int_{-\beta/2}^{\beta/2} dt V_I \left(\frac{\delta}{\delta b(t)} \right) \right] \mathcal{Z}_G(b, \beta) \right\} \Big|_{b=0}. \quad (2.72)$$

We then replace the partition function $\mathcal{Z}_G(b, \beta)$ by its explicit expression (2.59), calculated in Section 2.6 (equation (2.59)),

$$\frac{\mathcal{Z}(\beta)}{\mathcal{Z}_0(\beta)} = \exp \left[- \int_{-\beta/2}^{\beta/2} dt V_I \left(\frac{\delta}{\delta b(t)} \right) \right] \exp \left[\frac{1}{2} \int du dv b(u) \Delta(u-v) b(v) \right] \Big|_{b=0}. \quad (2.73)$$

If $V_I(q)$ is a polynomial, a more explicit form is obtained by expanding (2.73) in powers of $V_I(q)$. Perturbation theory is then reduced to calculating Gaussian expectation values:

$$\mathcal{Z}(\beta)/\mathcal{Z}_0(\beta) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int dt_1 dt_2 \cdots dt_k \langle V_I(q(t_1)) \cdots V_I(q(t_k)) \rangle_0,$$

where $\langle \bullet \rangle_0$ means the expectation value with respect to the Gaussian measure. The arguments given in Section 1.1 immediately apply here also, and the successive terms in the expansion can be calculated using Wick's theorem (1.14), in the form (2.61). This is the basis of perturbation theory.

Equal-time correlation functions. The perturbative expansion involves the Gaussian expectation values of equal-time products, which, with the help of Wick's theorem, can be expressed in terms of $\langle q^2(t) \rangle_0$. Because the paths contributing to the path integrals are continuous, this expectation value is defined. However, equal-time expectation values containing time derivatives are not, because the paths are not differentiable. An example is provided by a particle in a magnetic field (see Section 3.3.1).

Example. The quartic anharmonic oscillator. We consider the Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2q^2 + \lambda q^4, \quad (2.74)$$

and expand the partition function to second order in λ :

$$\mathcal{Z}(\beta)/\mathcal{Z}_0(\beta) = 1 - \lambda \int_{-\beta/2}^{\beta/2} dt \langle q^4(t) \rangle_0 + \frac{1}{2} \lambda^2 \int_{-\beta/2}^{\beta/2} dt_1 dt_2 \langle q^4(t_1) q^4(t_2) \rangle_0 + O(\lambda^3).$$

Then, using Wick's theorem (2.61), we infer

$$\langle q^4(t) \rangle_0 = 3 (\langle q^2(t) \rangle_0)^2 = 3\Delta^2(0),$$

and

$$\begin{aligned} \langle q^4(t_1) q^4(t_2) \rangle_0 &= 9 (\langle q^2(t_1) \rangle_0)^2 (\langle q^2(t_2) \rangle_0)^2 + 72 \langle q^2(t_1) \rangle_0 (\langle q(t_1) q(t_2) \rangle_0)^2 \langle q^2(t_2) \rangle_0 \\ &\quad + 24 (\langle q(t_1) q(t_2) \rangle_0)^4 \\ &= 9\Delta^4(0) + 72\Delta^2(0)\Delta^2(t_1 - t_2) + 24\Delta^4(t_1 - t_2). \end{aligned}$$

Combining these expressions and using the periodicity of $\Delta(t)$, we find

$$\begin{aligned} \mathcal{Z}(\beta)/\mathcal{Z}_0(\beta) &= 1 - 3\lambda\beta\Delta^2(0) + \frac{9}{2}\lambda^2\beta^2\Delta^4(0) + 36\beta\lambda^2\Delta^2(0) \int_{-\beta/2}^{\beta/2} dt \Delta^2(t) \\ &\quad + 12\lambda^2\beta \int_{-\beta/2}^{\beta/2} dt \Delta^4(t) + O(\lambda^3). \end{aligned} \quad (2.75)$$

In particular, the first three terms exponentiate, in agreement with the general result (1.18).

The classical limit and perturbation theory. A perturbative expansion can be generated for any decomposition of the potential into the sum of a quadratic term and a remainder, as in equation (2.70). However, if one wants to associate the perturbative expansion with a formal expansion in powers of \hbar , then one concludes from expression (2.19) that the action, and thus the potential, have to be expanded around a minimum. Denoting by q_0 a minimum of the potential, we then write

$$V(q) = V(q_0) + \frac{1}{2}V''(q_0)(q - q_0)^2 + V_1(q - q_0).$$

The expansion in powers of the coefficients of V_1 can then be organized as an expansion in powers of \hbar , called a *loop expansion*. Some problems associated with a possible degeneracy of the classical minimum are examined in Chapter 39.

Correlation functions and perturbation theory. We have explained how to calculate, in the form of a perturbative expansion, the path integral for any Hamiltonian of the form $p^2/2m + V(q)$ in terms of the path integral (2.55). The method can immediately be generalized to the corresponding correlation functions.

2.8 Semi-classical expansion

In the formal limit $\hbar \rightarrow 0$, one expects the quantum partition function to converge towards the classical partition function. We verify this property here by calculating the leading order and the first correction of the semi-classical expansion of the partition function (2.33). In particular, since \hbar has a dimension, the expansion parameter must take the form of \hbar divided by an action. The calculation exhibits the explicit expansion parameter.

We then use this result to generate semi-classical Wentzel–Kramers–Brillouin (WKB)-like approximations for the spectrum.

2.8.1 Quantum partition function

For $\hbar \rightarrow 0$, the leading term in the action (2.34) is the kinetic term and the dominant contributions come from constant paths. It is thus convenient to first calculate the matrix elements

$$\langle q_0 | e^{-\beta H} | q_0 \rangle = \int_{q(-\beta/2)=q_0}^{q(\beta/2)=q_0} [dq(t)] \exp [-\mathcal{S}(q)/\hbar]. \quad (2.76)$$

In the path integral, we change variables $q(t) \mapsto q(t) + q_0$ in such a way that the integral becomes

$$\langle q_0 | e^{-\beta H} | q_0 \rangle = \int_{q(-\beta/2)=0}^{q(\beta/2)=0} [dq(t)] \exp [-\Sigma(q)], \quad (2.77)$$

with

$$\Sigma(q) = \Sigma_0(q) + \int_{-\beta/2}^{\beta/2} dt V(q_0 + q(t)), \text{ and} \quad (2.78)$$

$$\Sigma_0(q) = \int_{-\beta/2}^{\beta/2} dt \frac{1}{2} m \dot{q}^2(t) / \hbar^2.$$

With these new boundary conditions, $q(t)$ is of order \hbar and the potential can be expanded in powers of q as

$$V(q_0 + q(t)) = V(q_0) + V'(q_0)q(t) + \frac{1}{2}V''(q_0)q^2(t) + O(\hbar^3).$$

Then, we expand the integrand,

$$\langle q_0 | e^{-\beta H} | q_0 \rangle = \mathcal{N}(\beta) e^{-\beta V(q_0)} \left[1 - V'(q_0) \int_{-\beta/2}^{\beta/2} dt \langle q(t) \rangle \right. \\ \left. + \frac{1}{2}(V'(q_0))^2 \int_{-\beta/2}^{\beta/2} dt du \langle q(t)q(u) \rangle - \frac{1}{2}V''(q_0) \int_{-\beta/2}^{\beta/2} dt \langle q^2(t) \rangle + O(\hbar^3) \right],$$

and calculate the successive terms ($\langle \bullet \rangle$ is a notation for the expectation value with respect to $e^{-\Sigma_0}$). First, we note that $\langle q(t) \rangle = 0$. The two-point function corresponding to Σ_0 , with the proper boundary conditions, is proportional to $\Delta(t, u) = \Delta(u, t)$, the solution of

$$-\ddot{\Delta}(t, u) = \delta(t - u), \quad \text{with } \Delta(-\beta/2, u) = \Delta(\beta/2, u) = 0.$$

It follows that

$$\frac{m}{\hbar^2} \langle q(t)q(u) \rangle = \Delta(t, u) = -\frac{1}{2}|t - u| + \frac{1}{4}\beta - ut/\beta, \quad (2.79)$$

a form that can be substituted in the expansion.

The normalization $\mathcal{N}(\beta)$ is given by

$$\mathcal{N}(\beta) = \langle q = 0 | e^{-\beta p^2/2m} | q = 0 \rangle,$$

and is obtained from the free expression (2.10), where $t - t'$ is replaced by $\hbar\beta$ (and $d = 1$). One finds

$$\mathcal{N}(\beta) = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}}. \quad (2.80)$$

The complete result takes the form of the simple integral (after an integration by parts),

$$\mathcal{Z}(\beta) = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}} \int dq \exp \left[-\beta V(q) - \beta^2 \hbar^2 V''(q)/24m + O(\hbar^4) \right]. \quad (2.81)$$

Discussion.

(i) For $\hbar \rightarrow 0$, one recovers the classical partition function with a Boltzmann weight obtained by integrating $e^{-\beta H(p, q)}$ over p , where H is the classical Hamiltonian, $H = p^2/2m + V(q)$:

$$\mathcal{Z}_{\text{cl.}}(\beta) = \int \frac{dp dq}{2\pi\hbar} e^{-\beta H(p, q)} = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}} \int dq e^{-\beta V(q)}.$$

(ii) Defining the *thermal wavelength* by

$$\lambda_{\text{th.}} = \hbar \sqrt{\beta/m},$$

and a length scale typical of the variations of the potential, and which increases when $\beta \rightarrow 0$ (high temperature),

$$l_{\text{pot.}} \propto \sqrt{|\langle V(q) \rangle / \langle V''(q) \rangle|},$$

we note that the ratio between the classical term and the first quantum correction is of order $\lambda_{\text{th.}}/l_{\text{pot.}}$. At high temperatures (*i.e.*, β small), the thermal wavelength is small, and statistical systems have a classical behaviour. On the contrary, at low temperature, quantum effects eventually dominate. Note that this analysis applies to regular potentials, at this order twice differentiable, but not to other potentials like the idealized potentials often used in QM examples (square well, and so on).

(iii) Formally, the classical limit leads to a kind of *dimensional reduction*: the quantum partition function corresponds to a path integral, that is, an integral over one-dimensional objects. By contrast, the corresponding classical function is given by an integral over the zero mode, which is just one point, that is, of dimension 0.

2.8.2 WKB spectrum

The spectrum in the WKB limit, that is, a limit in which $\hbar \rightarrow 0$ at fixed energy (in contrast with perturbation theory where the energies are also of order \hbar), can be obtained from the semi-classical expansion of the partition function described in Section 2.8.1.

In the case of a Hamiltonian with a discrete spectrum, the partition function can be written as

$$\mathcal{Z}(\beta) = \text{tr } e^{-\beta H} = \sum_{n=0} e^{-\beta E_n}.$$

The Laplace transform of $\mathcal{Z}(\beta)$ is then

$$G(E) = \int_0^\infty d\beta e^{\beta E} \mathcal{Z}(\beta) = \sum_n \frac{1}{E_n - E},$$

where the result in the right-hand side is obtained by analytic continuation from sufficiently negative values of the energy variable E .

After continuation to complex values of E , one obtains the eigenvalue distribution

$$\frac{1}{2i\pi} (G(E + i0) - G(E - i0)) = \sum_n \delta(E - E_n).$$

For what follows, it is convenient to consider the integrated distribution ($\theta(s)$ is the step function, $\theta(s) = 0$ for $s < 0$, $\theta(s) = 1$ for $s > 0$),

$$\int_{-\infty}^E dE' \frac{1}{2i\pi} (G(E' + i0) - G(E' - i0)) = \sum_n \theta(E - E_n).$$

In particular, if we choose $E = E_k$, a value of the spectrum, then

$$\int_{-\infty}^{E_k} dE' \frac{1}{2i\pi} (G(E' + i0) - G(E' - i0)) = k + 1/2, \quad (2.82)$$

where we have set $\theta(0) = 1/2$, a prescription motivated by a more careful analysis.

Approximating $\mathcal{Z}(\beta)$ by the leading term of the semi-classical result (2.81), from equation (2.82) we infer

$$G_{\text{cl.}}(E) = \frac{1}{\hbar} \sqrt{m/2} \int dq [V(q) - E]^{-1/2}.$$

We take the discontinuity of $G_{\text{cl.}}(E)$ on the cut. The exact discrete distribution is replaced by a continuous distribution, a result that is not surprising for a classical approximation. Nevertheless, from the integrated distribution, we can extract an approximation for the eigenvalues valid for large quantum numbers, because our approximation is also a high temperature approximation, where physical quantities are dominated by eigenvalues with large quantum numbers. We find

$$\frac{1}{2i} \int_{-\infty}^E dE' (G_{\text{cl.}}(E' + i0) - G_{\text{cl.}}(E' - i0)) = \frac{1}{\hbar} \int dq \theta(E - V(q)) \sqrt{2m[E - V(q)]}.$$

Equation (2.82) then yields the Bohr–Sommerfeld quantization condition:

$$\int dq \theta(E_k - V(q)) \sqrt{2m[E_k - V(q)]} = \hbar\pi(k + \tfrac{1}{2}).$$

The left-hand side is finite for $\hbar \rightarrow 0$: one verifies that, indeed, this approximation is an approximation for large quantum numbers $E_k = O(1)$, $k\hbar = O(1)$, while the distance between eigenvalues goes to zero with \hbar .

The successive terms in the expansion of the partition function yield corrections to this leading order result.

A2 Additional remarks

A2.1 A useful relation between determinant and trace

In Section 1.6.5, we have used the identity $\ln \det = \text{tr} \ln$. For *general complex matrices*, the proof of the identity is elementary, and is based on reducing the matrix to a *triangular form*, a transformation that is always possible.

To indicate how one can prove algebraically the result for operators, we consider the Gaussian path integral

$$\mathcal{Z}(\lambda) = \int [dq(t)] \exp \left[-\frac{1}{2} \int dt_1 dt_2 q(t_1) K(t_1, t_2; \lambda) q(t_2) \right], \quad (A2.1)$$

where λ is a parameter, the complex kernel K representing the operator \mathbf{K} is a symmetric function of (t_1, t_2) , and $\text{Re } \mathbf{K}$ is a strictly positive operator. The Gaussian integration yields,

$$\mathcal{Z}(\lambda) \propto [\det(\mathbf{K})]^{-1/2}. \quad (A2.2)$$

Using the definition (A2.1), we calculate the logarithmic derivative,

$$\frac{d \ln \mathcal{Z}}{d \lambda} = -\frac{1}{2} \left\langle \int dt_1 dt_2 q(t_1) \frac{dK(t_1, t_2; \lambda)}{d \lambda} q(t_2) \right\rangle,$$

where the expectation value $\langle \bullet \rangle$ refers to the Gaussian measure in expression (A2.1). Then,

$$\langle q(t_1) q(t_2) \rangle = \mathbf{K}^{-1}(t_1, t_2) \Rightarrow \frac{d \ln \mathcal{Z}}{d \lambda} = -\frac{1}{2} \text{tr} \frac{d \mathbf{K}}{d \lambda} \mathbf{K}^{-1}.$$

Therefore, comparing with the derivative of expression (A2.2), one obtains

$$\frac{d \ln \mathcal{Z}}{d \lambda} = -\frac{1}{2} \frac{d}{d \lambda} \ln \det(\mathbf{K}) = -\frac{1}{2} \text{tr} \frac{d \mathbf{K}}{d \lambda} \mathbf{K}^{-1}.$$

For an invertible, differentiable operator $\mathbf{M}(\lambda)$,

$$\frac{d}{d \lambda} \text{tr} \ln \mathbf{M}(\lambda) = \text{tr} \mathbf{M}^{-1}(\lambda) \frac{d}{d \lambda} \mathbf{M}(\lambda). \quad (A2.3)$$

It follows that

$$\frac{d}{d \lambda} [\ln \det(\mathbf{K}) - \text{tr} \ln(\mathbf{K})] = 0.$$

Integrated, the equation implies, but only for any symmetric operator \mathbf{K} , the identity

$$\ln \det \mathbf{K} = \text{tr} \ln \mathbf{K}. \quad (A2.4)$$

The result can be generalized by replacing the path integral by field integrals, and also integrating over complex fields to remove the restriction to symmetric operators.

A2.2 The two-point function: An integral representation

The two-point correlation function plays a special role in QFT and statistical physics. It has a useful representation that reveals some of its properties.

We assume that the quantum Hamiltonian H is Hermitian, bounded from below and, to simplify the notation, has a discrete spectrum with a unique ground state.

The connected two-point function (equation (2.46)) can be written as

$$W^{(2)}(t) \equiv \langle q(0)q(t) \rangle = \langle 0 | \hat{q} e^{-|t|H} \hat{q} e^{|t|H} | 0 \rangle - (\langle 0 | \hat{q} | 0 \rangle)^2.$$

Introducing the basis in which H is diagonal, we can expand it as

$$W^{(2)}(t) = \sum_{n>0} |\langle 0 | \hat{q} | n \rangle|^2 e^{-(\varepsilon_n - \varepsilon_0)|t|}, \text{ with } \varepsilon_0 < \varepsilon_1 \leq \varepsilon_2 \cdots \leq \varepsilon_n \cdots, \quad (A2.5)$$

where the state vectors $|n\rangle$ and the energies ε_n are, respectively, the eigenfunctions and eigenvalues of H .

As a consequence of the Hermiticity of H , the eigenvalues are real and the exponentials in the sum of the right-hand side have positive coefficients. The expansion (A2.5) leads to the representation of the Fourier transform $\tilde{Z}^{(2)}(\omega)$ of the connected two-point function,

$$\tilde{W}^{(2)}(\omega) = \int dt W^{(2)}(t) e^{i\omega t} = 2 \sum_{n>0} \frac{(\varepsilon_n - \varepsilon_0) |\langle 0 | \hat{q} | n \rangle|^2}{\omega^2 + (\varepsilon_n - \varepsilon_0)^2}. \quad (A2.6)$$

Two properties of $\tilde{W}^{(2)}(\omega)$ follow: it is a meromorphic function of ω^2 with poles only on the real negative axis. Moreover, the pole residues are all positive; it follows that $\tilde{Z}^{(2)}(\omega)$ cannot decrease faster than $1/\omega^2$ for ω^2 large. More precisely, the limit of the derivative of $W^{(2)}(t)$, when $t \rightarrow 0_+$, is given by

$$\lim_{t \rightarrow 0_+} \frac{d}{dt} Z^{(2)}(t) = \langle 0 | \hat{q} [\hat{q}, H] | 0 \rangle.$$

Since the left-hand side is real, one can replace the operator in the right-hand side by its Hermitian part:

$$\lim_{t \rightarrow 0_+} \frac{d}{dt} Z^{(2)}(t) = \frac{1}{2} \langle 0 | [\hat{q} [\hat{q}, H]] | 0 \rangle.$$

For a Hamiltonian quadratic in the momentum variable of the form $H = \frac{1}{2m}\hat{p}^2 + O(\hat{p})$, the commutators can be evaluated explicitly, and one obtains

$$[\hat{q} [\hat{q}, H]] = -\frac{1}{m} \Rightarrow \lim_{t \rightarrow 0_+} \frac{d}{dt} Z^{(2)}(t) = -\frac{1}{2m}.$$

Combining this result with the representation (A2.5), one finds

$$\frac{1}{2m} = \sum_{n \geq 0} |\langle 0 | \hat{q} | n \rangle|^2 (\varepsilon_n - \varepsilon_0),$$

and, therefore, from (A2.6),

$$\tilde{Z}^{(2)}(\omega) \underset{\omega \rightarrow \infty}{\sim} \frac{1}{m\omega^2}.$$

The result is not surprising. The behaviour for ω large is related to short-time evolution and we have shown that the most singular part of the matrix elements of $e^{-\beta H}$ is then determined by the free part $p^2/2m$ of the Hamiltonian.

Finally, when the spectrum of H has a continuous part, the sum in (A2.6) is replaced by an integral, the poles are replaced by a cut with a positive discontinuity, but the other conclusions remain unchanged. The relativistic generalization of representation (A2.6) is called the Källen–Lehmann representation [10].

Thermal correlation functions. The preceding calculation can easily be generalized to the finite temperature, or finite length from the classical point of view:

$$Z^{(2)}(t) = \mathcal{Z}^{-1}(\beta) \operatorname{tr} e^{-(\beta - |t|)H} \hat{q} e^{-|t|H} \hat{q}, \quad \mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta H}.$$

One verifies that the result is the same:

$$\lim_{t \rightarrow 0+} \frac{d}{dt} Z^{(2)}(t) = -\frac{1}{2m}.$$

This implies, in particular,

$$\left\langle (q(t + \varepsilon) - q(t))^2 \right\rangle \underset{\varepsilon \rightarrow 0}{\sim} |\varepsilon| \frac{1}{m},$$

confirming that the generic paths contributing to the path integral are Brownian paths.

A2.3 Time-ordered products of operators

We express all expressions in the *real-time formalism* and use the framework of simple QM, because the generalization to QFT is just a matter of changing notation. We start from expression (2.44), after continuation to real time, and want to directly recover a form analogous to the operator form (2.43). We also immediately take the thermodynamic limit. We define the n -point correlation function as

$$Z^{(n)}(t_1, t_2, \dots, t_n) = \int [dq] q(t_1) \cdots q(t_n) e^{i\mathcal{A}(q)/\hbar}. \quad (\text{A2.7})$$

We then order times,

$$t_1 \leq t_2 \leq \cdots \leq t_n, \quad (\text{A2.8})$$

and decompose the time interval into $(n+1)$ subintervals $(-\infty, t_1), (t_1, t_2), \dots, (t_n, +\infty)$. The total action is the sum of the corresponding contributions:

$$\mathcal{A}(q) = \sum_{i=1}^{n+1} \int_{t_{i-1}}^{t_i} \left[\frac{1}{2} m \dot{q}^2(t) - V(q(t)) \right] dt, \quad \text{with } t_0 = -\infty, \quad t_{n+1} = +\infty. \quad (\text{A2.9})$$

We rewrite the path integral (A2.7) with the help of the identity

$$\prod_{i=1}^n q(t_i) = \int \prod_{i=1}^n dq_i \delta[q(t_i) - q_i] q_i.$$

The path integral then factorizes into a product of path integrals corresponding to the different subintervals.

Returning to the very definition of the path integral (equations (2.18) and (2.19)), we note that the numerator in expression (A2.7) is exactly (recalling the ordering (2.3))

$$\langle 0 | e^{it_n H} \hat{q} e^{-i(t_n - t_{n-1})H} \hat{q} \cdots e^{-i(t_2 - t_1)H} \hat{q} e^{-it_1 H} | 0 \rangle.$$

Introducing the operator $Q(t)$, the Heisenberg representation of the operator \hat{q} ,

$$Q(t) = e^{itH} \hat{q} e^{-itH}, \quad (A2.10)$$

we can write

$$Z^{(n)}(t_1, t_2, \dots, t_n) = \langle 0 | Q(t_n) \cdots Q(t_1) | 0 \rangle. \quad (A2.11)$$

The order of the operators on the right-hand side reflects the time ordering (A2.8).

We introduce a time-ordering operator T , which, to a set of time-dependent operators $A_1(t_1), \dots, A_l(t_l)$, associates the time-ordered product (T-product) of these operators. For example, for $l = 2$,

$$T [A_1(t_1) A_2(t_2)] = A_1(t_1) A_2(t_2) \theta(t_1 - t_2) + A_2(t_2) A_1(t_1) \theta(t_2 - t_1).$$

We can then rewrite expression (A2.11), irrespective now of the order between the times t_1, \dots, t_n ,

$$Z^{(n)}(t_1, t_2, \dots, t_n) = \langle 0 | T [Q(t_1) Q(t_2) \cdots Q(t_n)] | 0 \rangle. \quad (A2.12)$$

The n -point function (A2.7) can be expressed as the vacuum expectation value of the time-ordered product of Heisenberg operators. These time-ordered products are the analytic continuation of the imaginary time correlation functions. They generate Green's functions from which one can derive scattering amplitudes (see Section 6.3.2).

More generally, at a finite temperature $1/\beta$, one finds the time-dependent correlation functions of quantum statistical physics

$$Z^{(n)}(t_1, t_2, \dots, t_n) = \mathcal{Z}^{-1}(\beta) \text{tr} \{ e^{-\beta H} T [Q(t_1) Q(t_2) \cdots Q(t_n)] \}. \quad (A2.13)$$

3 Quantum mechanics (QM): Path integrals in phase space

In Chapter 2, we have defined a path integral representation of the matrix elements of the quantum statistical operator $e^{-\beta H}$ in the case of Hamiltonians H of the separable form $p^2/2m + V(q)$. We will now extend the construction to Hamiltonians that are more general functions of phase-space variables [11–13]. This results in integrals over paths in phase space involving the action expressed in terms of the classical Hamiltonian $H(p, q)$. However, in the case of general Hamiltonians, the path integral is not completely defined, and this reflects the problem that the classical Hamiltonian does not completely specify the quantum Hamiltonian due to the problem of ordering quantum operators in products.

When the Hamiltonian is a quadratic function of the momentum variables, the integral over momenta is Gaussian and can be performed. In the separable example, the path integral of Chapter 2 is recovered. In the case of the charged particle in a magnetic field a more general form is found. It is ambiguous, since a problem of operator ordering arises, and the ambiguity must be fixed. Hamiltonians that are general quadratic functions provide other important examples and, therefore, we analyse them thoroughly. Such Hamiltonians appear in the quantization of the motion on Riemannian manifolds. We find that the problem of ambiguities is even more severe. We illustrate the analysis by the quantization of the free motion on the sphere S_{N-1} .

In the Appendix, we discuss various additional issues related to quantization problems, including the path integral quantization of systems for which the action generating the classical equation of motion cannot be globally defined. The problem arises when phase space, as in the quantization of spin degrees of freedom, or ordinary space, as in the example of the magnetic monopole, have non-trivial topological properties.

3.1 General Hamiltonians: Phase-space path integral

We first recall how the *classical* equations of motion, in the Hamiltonian formalism and in real time, can be derived by a variational principle from an action function of phase-space variables, position, and conjugate momentum, which we denote by q, p . We then construct the corresponding path integral representation *first in real time*, then in imaginary time.

3.1.1 Hamiltonian and Lagrangian

The classical action corresponding to a Hamiltonian $\mathcal{H}(p, q; t)$ reads ($\dot{q} \equiv dq/dt$)

$$\mathcal{A}(p, q) = \int_{t_1}^{t_2} dt [p(t)\dot{q}(t) - \mathcal{H}(p(t), q(t); t)], \quad (3.1)$$

with boundary conditions $q(t_1) = q_1, q(t_2) = q_2$. Indeed, varying the trajectory in phase space,

$$\frac{\delta \mathcal{A}(p, q)}{\delta p(t)} = \dot{q}(t) - \frac{\partial \mathcal{H}}{\partial p} = 0, \quad \frac{\delta \mathcal{A}(p, q)}{\delta q(t)} = -\dot{p}(t) - \frac{\partial \mathcal{H}}{\partial q} = 0, \quad (3.2)$$

one recovers the classical equations of motion in the Hamiltonian formulation.

Legendre transformation. Hamiltonian and Lagrangian are related by a Legendre transformation (see Section 1.8). The *Legendre transformation* is defined here by (\dot{q} and q are considered as *independent quantities*)

$$\mathcal{H}(p, q; t) + \mathcal{L}(\dot{q}, q; t) = p(t)\dot{q}(t), \quad \dot{q}(t) = \frac{\partial \mathcal{H}}{\partial p}.$$

One verifies that these equations imply

$$p(t) = \frac{\partial \mathcal{L}}{\partial \dot{q}}, \quad \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial \mathcal{H}}{\partial q} = 0,$$

and equations (3.2) then are equivalent to the Euler–Lagrange equations derived from the action (3.1) expressed in terms of the Lagrangian,

$$\mathcal{A}(q) = \int_{t_1}^{t_2} dt \mathcal{L}(\dot{q}(t), q(t); t).$$

3.1.2 QM: Path integral for time evolution

In classical mechanics, the Lagrangian is more intuitive, because it is expressed in terms of physical observables, unlike the Hamiltonian. However, QM is based on a Hamiltonian formalism, the Hamiltonian governing unitary time evolution.

For a general quantum Hamiltonian \hat{H} function of position \hat{q} and conjugate momentum \hat{p} operators, with canonical commutation relation $[\hat{q}, \hat{p}] = i\hbar$, one can use the method of Section 2.1, based on evaluating the matrix elements of the evolution operator solution of

$$i\hbar \frac{\partial U}{\partial t}(t, t') = \hat{H}(t)U(t, t'), \quad U(t', t') = \mathbf{1}, \quad (3.3)$$

for short time intervals. Rather, we postulate as an ansatz, a representation by a path integral over trajectories in phase space of the form

$$\langle q'' | U(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [dp(t)dq(t)] \exp[i\mathcal{A}(p, q)/\hbar], \quad (3.4)$$

where $\mathcal{A}(p, q)$ is the action (3.1), and \mathcal{H} a smooth function, classical limit of the quantum Hamiltonian \hat{H} . Expression (3.4) is especially aesthetic, since it involves only the invariant Liouville measure on phase space and the classical action. In particular, it is formally invariant under canonical transformations: transformations in phase space preserving the Poisson brackets,

$$(p, q) \mapsto (P(p, q), Q(p, q)), \text{ with } \{p, q\} = \{P, Q\}.$$

Again, the extension to several degrees of freedom is simple: one substitutes in the path integral representation (3.4), the corresponding classical action, and Liouville measure.

Verification. We evaluate the path integral for a short time interval $\tau = t'' - t'$ limit. The action becomes

$$\mathcal{A}(p, q) = p(q'' - q') - \tau \mathcal{H}(p, q_{\text{av.}}),$$

where $q_{\text{av.}}$ is some average value between q' and q'' . In a time-discretized form, the measure in phase space is normalized with respect to $2\pi\hbar$. Then,

$$\begin{aligned} \langle q'' | U(t' + \tau, t') | q' \rangle &\approx \int dp e^{ip(q'' - q')/\hbar} \left[1 - \frac{i\tau}{\hbar} \mathcal{H}(p, q_{\text{av.}}) \right] \\ &= \left[1 - \frac{i\tau}{\hbar} \mathcal{H}(-i\hbar\partial/\partial q, q_{\text{av.}}) \right] \delta(q'' - q') = \langle q'' | (\mathbf{1} - i\tau \mathbf{H}/\hbar) | q' \rangle, \end{aligned}$$

where \mathbf{H} , in fact, can be any quantum operator that has the correct classical limit. Therefore, the path integral (3.4) is, formally, the correct expression, but it is ambiguous.

Remarks

(i) The preceding discussion illustrates the problem that, since the path integral involves only the classical Hamiltonian, in the replacement of a quantum Hamiltonian by its classical limit, the order in products of non-commuting operators is lost and thus, the path integral, beyond its formal expression, is not defined.

A time-discretized form of the path integral involves a precise quantum Hamiltonian, but then this information is lost in the continuum limit. Therefore, the path integral must be supplemented with some additional information or limiting procedure.

(ii) The canonical invariance can be true only for a very restricted class of transformations. We will show elsewhere (Section A3.2) that in the case of a one-dimensional, one degree of freedom Hamiltonian \mathcal{H} , one can always find a canonical transformation that maps \mathcal{H} onto a free Hamiltonian:

$$p\dot{q} - \mathcal{H} \longmapsto P\dot{Q} - \frac{1}{2m}P^2.$$

One could then naively conclude that semi-classical approximations are always exact. It is easy to produce counter examples.

Phase space path integral for the statistical operator. The statistical operator is obtained by changing to imaginary time $t \mapsto -it$ in expression (3.4). One finds

$$\langle q'' | U(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [dp(t)dq(t)] \exp [-\mathcal{S}(p, q) / \hbar], \quad (3.5)$$

where $\mathcal{S}(p, q)$ is the Euclidean action in the Hamiltonian formalism,

$$\mathcal{S}(p, q) = \int_{t'}^{t''} dt [-ip(t)\dot{q}(t) + \mathcal{H}(p(t), q(t); t)]. \quad (3.6)$$

3.1.3 Separable Hamiltonians: Equivalence

First, we verify that in the case of a simple Hamiltonian like

$$\mathcal{H} = p^2/2m + V(q),$$

where no commutation problem is involved, the real-time continuation of the path integral (2.19) is recovered after integration over $p(t)$ in expression (3.4).

The classical action is

$$\mathcal{A}(p, q) = \int_{t'}^{t''} dt [p(t)\dot{q}(t) - p^2(t)/2m - V(q(t))]. \quad (3.7)$$

In the path integral (3.4), the integral over the momentum variables $p(t)$ is Gaussian. Following the strategy explained in Section 2.6, we change variables, $p(t) \mapsto r(t)$, with

$$p(t) = m\dot{q}(t) + r(t). \quad (3.8)$$

The action becomes

$$\mathcal{A}(q, r) = \int_{t''}^{t'} dt \left(-\frac{1}{2m}r^2(t) + \frac{1}{2}m\dot{q}^2(t) - V(q(t)) \right).$$

The path integral then factorizes into an integral over $r(t)$,

$$\mathcal{N}(t', t'') = \int [dr(t)] \exp \left(-\frac{i}{\hbar} \int_{t'}^{t''} \frac{r^2(t)}{2m} dt \right),$$

that does not depend on the potential $V(q)$, and yields only a normalization factor \mathcal{N} function of t' and t'' , and an integral over $q(t)$. One finds,

$$\langle q'' | U(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [dq(t)] \exp \left(\frac{i}{\hbar} \mathcal{A}(q) \right),$$

with

$$\mathcal{A}(q) = \int_{t'}^{t''} dt \left[\frac{1}{2} m \dot{q}^2(t) - V(q(t)) \right]. \quad (3.9)$$

Therefore, we have verified explicitly the equivalence between the representations (3.4) and (2.21).

3.2 The harmonic oscillator. Perturbative expansion

We first consider the quantum harmonic oscillator coupled linearly to external sources. We then use the results to define a perturbative expansion for general Hamiltonians.

3.2.1 The quantum harmonic oscillator

As an example, we consider the Euclidean action

$$\mathcal{S}(p, q) = \int_{t'}^{t''} dt \left[-ip(t)\dot{q}(t) + \mathcal{H}(p(t), q(t)) - a(t)p(t) - b(t)q(t) \right], \quad (3.10)$$

where \mathcal{H} is the Hamiltonian of the harmonic oscillator ($\omega > 0$ constant),

$$\mathcal{H}(p, q) = \frac{1}{2} p^2 + \frac{1}{2} \omega^2 q^2,$$

and $a(t)$, $b(t)$ are external sources. We calculate the trace of the statistical operator at zero temperature, which corresponds to a path integral (3.5) with vanishing boundary conditions at infinite time. As we have done in Section 3.1.3, we first integrate over $p(t)$, which amounts to substituting

$$p(t) = i\dot{q}(t) + a(t),$$

and, after an integration by parts, obtain

$$\mathcal{S}(q) = \int dt \left[\frac{1}{2} \dot{q}^2(t) + \frac{1}{2} \omega^2 q^2(t) - (b(t) - ia(t))q(t) - \frac{1}{2} a^2(t) \right].$$

We then use the results (2.32) and (2.59) and find

$$\text{tr } U(+\infty, -\infty) = \text{tr } U_0 e^{-\mathcal{S}(a,b)/\hbar},$$

with (equation (2.58))

$$\mathcal{S}(a, b) = -\frac{1}{2} \int_{t'}^{t''} dt du (b(t) - ia(t)) \frac{1}{2\omega} e^{-\omega|t-u|} (b(u) - ia(u)) - \frac{1}{2} \int_{t'}^{t''} dt a^2(t). \quad (3.11)$$

Differentiating twice $e^{-S(a,b)}$ with respect to $a(t)$ and $b(t)$, and then setting $a(t), b(t) \equiv 0$, we obtain the Gaussian two-point functions,

$$\begin{cases} \langle p(t)p(t') \rangle = \frac{1}{2}\omega e^{-\omega|t-t'|}, \\ \langle p(t)q(t') \rangle = -\frac{1}{2}i \operatorname{sgn}(t-t') e^{-\omega|t-t'|}, \\ \langle q(t)q(t') \rangle = \frac{1}{2\omega} e^{-\omega|t-t'|}, \end{cases} \quad (3.12)$$

where $\operatorname{sgn}(t)$ is the sign function: $\operatorname{sgn}(t) = 1$ for $t > 0$, $\operatorname{sgn}(-t) = -\operatorname{sgn}(t)$.

The free action limit. The $\omega = 0$ limit with boundaries at infinity does not exist (in another context, this is called an infrared divergence). Finite time boundary conditions are required. For simplicity, we choose conditions relevant for the partition function, $q(-\beta/2) = q(\beta/2) = 0$ (see Section 2.8.1). Then,

$$\Delta(t, t') = -\frac{1}{2}|t-t'| + \frac{1}{4}\beta - tt'/\beta, \quad (3.13)$$

and

$$\begin{cases} \langle p(t)p(t') \rangle = 1/\beta, \\ \langle p(t)q(t') \rangle = -\frac{1}{2}i \operatorname{sgn}(t-t') - it'/\beta, \\ \langle q(t)q(t') \rangle = \frac{1}{4}\beta - \frac{1}{2}|t-t'| - tt'/\beta. \end{cases} \quad (3.14)$$

For more generic boundary conditions, $q(-\beta/2) = q'$, $q(\beta/2) = q''$,

$$\langle q(t) \rangle = q' + (t + \beta/2)(q'' - q')/\beta, \quad \langle p(t) \rangle = (q'' - q')/\beta,$$

and the *connected* two-point correlation functions are still given by equations (3.14).

3.2.2 Phase space path integral: Perturbative definition

We decompose a general Hamiltonian into the sum ($\omega > 0$ constant)

$$\mathcal{H}(p, q) = \frac{1}{2}p^2 + \frac{1}{2}\omega^2q^2 + \mathcal{H}_{\text{Int.}}(p, q),$$

where $\mathcal{H}_{\text{Int.}}$ can be expanded as a series in p, q .

Perturbative expansion. We expand the path integral in powers of $\mathcal{H}_{\text{Int.}}$. The first term is given by the Gaussian expectation value $\langle H_{\text{Int.}} \rangle$. Wick's theorem implies that general Gaussian expectation values can be expressed in terms of the various Gaussian two-point functions.

If $\mathcal{H}_{\text{Int.}}$ is separable, that is,

$$\mathcal{H}_{\text{Int.}}(p, q) = \mathcal{H}_1(p) + \mathcal{H}_2(q), \quad (3.15)$$

only the pp and qq two-point functions taken at $t = t'$ are involved, and expectation values are defined. This corresponds to the case where no problem of ordering in products of quantum operators arises. However, when $\mathcal{H}_{\text{Int.}}$ contains products pq , the expectation value involves $\operatorname{sgn}(0)$ (equations (3.12)), which is not defined. This reflects the problem of operator ordering. From the path integration viewpoint, with the Gaussian measure associated with the harmonic oscillator, the paths in phase space that contribute are not regular enough. Finally, only the choice $\operatorname{sgn}(0) = 0$ corresponds to a Hermitian Hamiltonian.

At higher orders, Wick's theorem again implies that all expectation values can be expressed in terms of Gaussian two-point functions. However, the singular terms correspond to equal-time contributions, *which are already present at leading order*.

3.3 Hamiltonians quadratic in momentum variables

We have outlined a few problems one faces when trying to define a general path integral in phase space. To show that expression (3.6) has, nevertheless, at least some heuristic value, we now discuss the example of more general Hamiltonians quadratic in momenta.

3.3.1 Quantization in a static magnetic field

The classical Hamiltonian of a particle (in \mathbb{R}^d) in a potential and a static magnetic field has the form,

$$\mathcal{H}_{\text{mag.}} = \frac{1}{2m} [\mathbf{p} + e\mathbf{A}(\mathbf{q})]^2 + V(\mathbf{q}), \quad (3.16)$$

where \mathbf{A} is the vector potential and e the electric charge. The Hamiltonian is *gauge invariant* (see Chapter 21), that is, invariant in the changes ($\nabla \equiv (\partial/\partial q_1, \dots, \partial/\partial q_d)$)

$$\mathbf{A}(\mathbf{q}) \mapsto \mathbf{A}(\mathbf{q}) + \nabla \Lambda(\mathbf{q}), \quad \mathbf{p} \mapsto \mathbf{p} - e\nabla \Lambda(\mathbf{q}), \quad (3.17)$$

which do not affect the magnetic field.

The quantization of the Hamiltonian leads to a problem of order of quantum operators in the product $\mathbf{p} \cdot \mathbf{A}(\mathbf{q})$ [14]. The order in the quantum Hamiltonian (\hat{X} denotes the quantum operator associated with the classical variable X),

$$\hat{H} = \frac{1}{2m} [\hat{\mathbf{p}}^2 + e\mathbf{A}(\hat{\mathbf{q}}) \cdot \hat{\mathbf{p}} + e\hat{\mathbf{p}} \cdot \mathbf{A}(\hat{\mathbf{q}}) + e^2 \mathbf{A}^2(\hat{\mathbf{q}})] + V(\hat{\mathbf{q}}), \quad (3.18)$$

is determined by the *condition of Hermiticity*. Indeed, a change in the order of operators is equivalent to the addition of an imaginary potential proportional to

$$\mathbf{A}(\hat{\mathbf{q}}) \cdot \hat{\mathbf{p}} - \hat{\mathbf{p}} \cdot \mathbf{A}(\hat{\mathbf{q}}) = i\hbar \nabla \cdot \mathbf{A}(\hat{\mathbf{q}}).$$

In the presence of a magnetic field, the phase of wave functions $\psi(\mathbf{q})$ can be changed at each point of space independently (a gauge symmetry). The *U(1) gauge transformation*,

$$\psi(\mathbf{q}) \mapsto \psi(\mathbf{q}) e^{-ie\Lambda(\mathbf{q})/\hbar}, \quad (3.19)$$

can be cancelled by adding the gradient term (3.17) to the vector potential.

Path integral representation. The Euclidean action (3.6) becomes,

$$\mathcal{S}(\mathbf{p}, \mathbf{q}) = \int_{t'}^{t''} dt [-i\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) + \mathcal{H}_{\text{mag.}}(\mathbf{p}(t), \mathbf{q}(t))]. \quad (3.20)$$

In the path integral,

$$\langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle = \int_{\mathbf{q}(t') = \mathbf{q}'}^{\mathbf{q}(t'') = \mathbf{q}''} [d\mathbf{p}(t)d\mathbf{q}(t)] \exp [-\mathcal{S}(\mathbf{p}, \mathbf{q})/\hbar], \quad (3.21)$$

the integral over momentum is still Gaussian. In the action, the terms linear in \mathbf{p} can be eliminated by the change of variables, $\mathbf{p}(t) \mapsto \mathbf{r}(t)$, with

$$\mathbf{p}(t) = im\dot{\mathbf{q}}(t) - e\mathbf{A}(\mathbf{q}(t)) + \mathbf{r}(t).$$

After integration over $\mathbf{r}(t)$, one obtains an integral over the path $\mathbf{q}(t)$ with the action

$$\mathcal{S}(\mathbf{q}) = \int_{t'}^{t''} dt [\frac{1}{2}m\dot{\mathbf{q}}^2(t) + ie\mathbf{A}(\mathbf{q}(t)) \cdot \dot{\mathbf{q}}(t) + V(\mathbf{q}(t))], \quad (3.22)$$

in which one recognizes the Euclidean (or imaginary time) classical action integral of the Lagrangian, corresponding to the Hamiltonian (3.16).

Note that, in this example, the Euclidean action is not real, and thus does not define a positive measure. Actually, the imaginary contribution ensures consistency of the path integral with gauge invariance: the transformation (3.17) adds a total derivative to the Lagrangian and the variation of the action, $\delta\mathcal{S}$ is

$$\delta\mathcal{S} = ie [\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}')].$$

It follows that the matrix elements $\langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle$ are multiplied by the phase factor $\exp[-ie(\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}'))/\hbar]$, consistently with the gauge transformation (3.19).

Hermiticity and the sgn(0) problem. From the analysis of Section 3.2.2, one expects a problem related to operator ordering, because the Hamiltonian is not separable (the analysis is inspired by section 5.2 of Ref. [6]).

To understand where the problem shows up, we expand the path integral corresponding to the action

$$\mathcal{S}(\mathbf{q}) = \int dt \left[\frac{1}{2} \dot{\mathbf{q}}^2(t) + ie \mathbf{A}(\mathbf{q}(t)) \cdot \dot{\mathbf{q}}(t) + \frac{1}{2} \omega^2 \mathbf{q}^2(t) \right], \quad \omega > 0,$$

in powers of the charge e . At first order, we find a contribution which, using Wick's theorem, can be written as ($\langle \bullet \rangle$ is a notation for expectation value with weight $e^{-\mathcal{S}}$)

$$-\frac{ie}{\hbar} \int dt \sum_{i,j} \langle \dot{q}_i(t) q_j(t) \rangle \left\langle \frac{\partial A_i}{\partial q_j} \right\rangle.$$

From equations (2.58, 2.60), we derive

$$\langle \dot{q}_i(t_1) q_j(t_2) \rangle = -\frac{1}{2} \hbar \delta_{ij} \operatorname{sgn}(t_1 - t_2) e^{-\omega|t_1 - t_2|}.$$

Therefore, formally

$$-\frac{ie}{\hbar} \int dt \sum_{i,j} \langle \dot{q}_i(t) q_j(t) \rangle \left\langle \frac{\partial A_i}{\partial q_j} \right\rangle = \frac{1}{2} ie \operatorname{sgn}(0) \int dt \langle \nabla \cdot \mathbf{A}(\mathbf{q}(t)) \rangle.$$

As expected, the result involves $\operatorname{sgn}(0)$, because the Hamiltonian is not separable and, thus, the path integral is not defined. This difficulty also reflects the property that, since the Brownian paths are not differentiable, a potential involving $\dot{\mathbf{q}}$ is not defined and a regularization is required (see Chapter 8). For example, one can add a term like $\nu^2(\ddot{q})^2$ in the action and take then the $\nu = 0$ limit.

Note that the assignment $\operatorname{sgn}(0) = 0$ preserves invariance under time reversal. Moreover, equation (2.58) implies

$$\frac{d}{dt} \langle q_i(t) q_j(t) \rangle = 0.$$

Therefore, only the choice $\operatorname{sgn}(0) = 0$ is *consistent with the commutation between time derivative and averaging*.

Another choice would be equivalent to the addition of a term $-\frac{1}{2} ie \operatorname{sgn}(0) \hbar \nabla \cdot \mathbf{A}(\mathbf{q})$ to the action, proportional to the commutator $[\hat{\mathbf{p}}, \mathbf{A}(\hat{\mathbf{q}})]$, showing the relation between this ambiguity and the problem of ordering operators. Such a term could be cancelled explicitly by modifying the action, resulting in a more complicate formalism.

Finally, the operator $e^{-\beta \hat{H}}$ is Hermitian when the Hamiltonian is Hermitian. From the point of view of the path integral, the hermiticity condition implies formally the invariance of the path integral under the simultaneous changes $\mathcal{S} \rightarrow \mathcal{S}^*$ (complex conjugation) and $t \rightarrow -t$ (transposition, and thus exchange of boundary conditions). By preserving this symmetry throughout the calculation, and, therefore, setting $\operatorname{sgn}(0) = 0$, one ensures consistency with the (Hermitian) choice of quantization (3.18).

We meet this problem again in Sections 4.2, 4.6, and 37.6.1, and in Chapter 34, and it is implicit in field theories with derivative couplings, like gauge theories.

3.3.2 General quadratic Hamiltonians

A general Hamiltonian quadratic in \mathbf{p} can be derived from a general Lagrangian quadratic in the velocities. Because, in all examples we encounter, the quantization problem is initially formulated in terms of a classical Lagrangian, we consider the problem of quantizing a real-time Lagrangian in n -dimensional space of the form

$$\mathcal{L}(\dot{q}, q) = \frac{1}{2} \sum_{\alpha, \beta} \dot{q}^\alpha g_{\alpha\beta}(q) \dot{q}^\beta, \quad 1 \leq \alpha, \beta \leq n, \quad (3.23)$$

where $g_{\alpha\beta}(q)$ is a positive matrix. Interesting examples correspond to random motion on Riemannian manifolds, as we shall briefly discuss in Section 34.9. The tensor $g_{\alpha\beta}(q)$ then is the metric tensor (for details see Chapter 28). Field theoretical generalizations are studied in Chapters 19 and 29.

The corresponding classical Hamiltonian is obtained by a *Legendre transformation*. The conjugate momenta are

$$p_\alpha = \frac{\partial \mathcal{L}(\dot{q}, q)}{\partial \dot{q}^\alpha} = \sum_\beta g_{\alpha\beta}(q) \dot{q}^\beta$$

and, therefore, the Hamiltonian is

$$H(p, q) = \sum_\alpha p_\alpha \dot{q}^\alpha - \mathcal{L}(\dot{q}, q) = \frac{1}{2} \sum_{\alpha, \beta} p_\alpha g^{\alpha\beta}(q) p_\beta, \quad (3.24)$$

where the traditional notation $g^{\alpha\beta}$ for the matrix inverse of $g_{\alpha\beta}$ has been used:

$$\sum_\gamma g_{\alpha\gamma}(q) g^{\gamma\beta}(q) = \delta_\alpha^\beta.$$

The Euclidean action reads

$$\mathcal{S}(p, q) = \int dt \left[-i \sum_\alpha p_\alpha(t) \dot{q}_\alpha(t) + \frac{1}{2} \sum_{\alpha, \beta} p_\alpha(t) g^{\alpha\beta}(q(t)) p_\beta(t) \right]. \quad (3.25)$$

Again the integration over $p(t)$ is Gaussian and can be performed. Formally, one finds,

$$\text{tr } U = \int_t \prod \frac{dq(t)}{\sqrt{\det g^{\alpha\beta}(q(t))}} e^{-\mathcal{S}(q)},$$

with

$$\mathcal{S}(q) = \frac{1}{2} \int dt \sum_{\alpha, \beta} \dot{q}^\alpha(t) g_{\alpha\beta}(q(t)) \dot{q}^\beta(t). \quad (3.26)$$

In $\mathcal{S}(q)$ we recognize the Euclidean action integral of the classical Lagrangian (3.23) after continuation to imaginary time. This is not surprising, since the integration over p_α is equivalent to a Legendre transformation.

However, in contrast with the preceding example, the Gaussian integration has generated a non-trivial $q(t)$ dependent normalization factor,

$$\mathcal{N}(q) \propto \frac{1}{\sqrt{\det g^{\alpha\beta}(q(t))}}, \quad (3.27)$$

which, when $g_{\alpha\beta}$ is the metric tensor on a Riemannian manifold, formally reconstructs the *covariant measure* on the manifold (see Section 28.3.1).

However, a difficulty appears with the evaluation of the normalization $\mathcal{N}(q)$. We rewrite the p integral with the action (3.25) as

$$\int [dp(t)] \exp \left[-\frac{1}{2} \int dt dt' \sum_{\alpha,\beta} p_\alpha(t) K_{\alpha\beta}(t,t') p_\beta(t') \right],$$

with

$$K_{\alpha\beta}(t,t') \equiv \delta(t-t') g^{\alpha\beta}(q(t)).$$

Using the general identity (A2.4), $\ln \det \mathbf{K} = \text{tr} \ln \mathbf{K}$, one finds formally

$$\mathcal{N}(\mathbf{q}) \propto e^{\text{tr} \ln \mathbf{K}} = \exp \left[\frac{1}{2} \delta(0)' \int \text{tr} \ln \mathbf{g}[q(t)] dt \right]. \quad (3.28)$$

Due to the factor $\delta(0)$, the expression is not defined (δ being the Dirac δ -function). This difficulty is again related to the problem of ordering operators in products. If one performs a small \hbar (semi-classical) expansion of the path integral, one finds a divergent quantum correction (see Chapters 19, 29). This divergence is cancelled by the leading contribution coming from (3.28), as the calculation of Section 3.3.3 shows. However, the remaining finite part is not defined in the formal continuum time limit. It is possible to use a time-discretized form, which reflects a choice of quantization, to calculate it. Another direct way of understanding this difficulty is to notice that, since in expression (3.27) the difference $|q - q'|$ is generically of order $\sqrt{|t - t'|}$, a replacement of $\mathbf{g}(\mathbf{q}_{\text{av}})$ by any other symmetric function of q and q' , which has the same $q = q'$ limit, changes this quantity at order $|t - t'|$, in general. The modification of $\mathcal{N}(\mathbf{q})$ then generates a finite quantum correction to the classical action, typical of a commutation of momentum and position operators. Some more details about this problem can be found in Section A37.2.

From the point of view of a statistical model corresponding to the discretized path integral, this means that the *continuum limit* is less universal than in the simpler flat $\mathbf{g} = \mathbf{1}$ case. It depends on a number of additional parameters related to the choice of quantization. However, in many examples, quantization is constrained by symmetry properties. The same symmetries then also determine the corresponding parameters.

3.3.3 Phase-space formalism: The $\delta(0)$ problem

If one expands around flat space, setting

$$g^{\alpha\beta} = \delta^{\alpha\beta} + h^{\alpha\beta},$$

adds to the action (3.25) either a harmonic oscillator term or finite time boundaries, it is possible to expand the path integral in powers of \hbar . The propagators are given by equations (3.12) or (3.14) multiplied by $\delta^{\alpha\beta}$. Using Wick's theorem, one finds that the first-order correction has the form

$$\begin{aligned} -\frac{1}{2} \int dt \sum_{\alpha,\beta} \langle p_\alpha(t) h^{\alpha\beta}(q(t)) p_\beta(t) \rangle &= -\frac{1}{2n} \int dt \langle \mathbf{p}^2(t) \rangle \langle \text{tr } \mathbf{h}(q(t)) \rangle \\ &\quad - \frac{1}{2n^2} \sum_{\alpha,\beta} \langle \nabla_\alpha \nabla_\beta h^{\alpha\beta}(q(t)) \rangle. \end{aligned}$$

We note that no $\delta(0)$ contribution is generated, but only the unavoidable less singular $\text{sgn}(0)$, resulting from the ambiguity in operator ordering appears in the perturbative expansion. The divergence thus results from the momentum integration. This exercise exhibits a situation in which the order of integrations in path integrals matters.

Finally, the first order is the most singular, because it contains the equal-time contributions. When the ambiguity is fixed, the full perturbative expansion is defined.

3.4 The spectrum of the $O(2)$ -symmetric rigid rotator

To illustrate the discussion of Hamiltonians quadratic in momentum variables, we calculate the spectrum of the $O(N)$ rigid rotator, using the path integral representation (this model is also the one-dimensional $O(N)$ non-linear σ -model, see Chapter 19). We first examine the $N = 2$ case, which is simpler and can be treated exactly. The $O(2)$ rotator actually provides an example of the peculiarities of the path integral when position space has non-trivial topological properties.

In the case of a general parametrization of the circle, the Hamiltonian is a quadratic function of momenta with a position-dependent coefficient, which leads to quantization problems. However, if the circle is parametrized by an angle θ , the Hamiltonian of the $O(2)$ rotator becomes simply

$$H = -\frac{1}{2} \frac{\partial^2}{(\partial\theta)^2}, \quad (3.29)$$

and would be a free Hamiltonian if θ were not an angular variable. As a consequence, the spectrum, instead of being continuous, is discrete. The eigenvalues are

$$E_l = \frac{1}{2}l^2, \quad (3.30)$$

where l is an integer, the angular momentum.

The path integral representation of the matrix elements of the statistical operator $e^{-\beta H}$ is analogous to the path integral ($\dot{\theta} \equiv d\theta/dt$) of the free motion:

$$\langle \theta'' | e^{-\beta H} | \theta' \rangle = \int_{\theta(0)=\theta'}^{\theta(\beta)=\theta''} [d\theta(t)] \exp \left[-\frac{1}{2} \int_0^\beta (\dot{\theta}(t))^2 dt \right]. \quad (3.31)$$

However, the cyclic character of the variable also modifies the evaluation of the path integral. The matrix elements $\langle \theta'' | e^{-\beta H} | \theta' \rangle$ are invariant under a translation of θ' and θ'' , and periodic in both angles. Therefore, the operator $e^{-\beta H}$ can be diagonalized in the basis of eigenfunctions $\psi_l = e^{il\theta}$:

$$e^{-\beta H_l} = \frac{1}{(2\pi)^2} \int_0^{2\pi} d\theta' d\theta'' e^{il(\theta'-\theta'')} \langle \theta'' | e^{-\beta H} | \theta' \rangle, \text{ with } l \text{ integer.}$$

We now introduce the topological charge

$$Q = \frac{1}{2\pi} \int_0^\beta dt \dot{q}(t).$$

The charge is topological, because it depends only on the end points of a trajectory. It characterizes the different classes of mappings of the circle onto itself. Indeed, if a trajectory goes from θ' to θ'' by winding q times around the circle, then

$$Q = q + \frac{\theta'' - \theta'}{2\pi},$$

where the integer q is called the *winding number*.

We then consider the path integral

$$e^{-\beta H_l} = \int_{\theta(0)=\theta'}^{\theta(\beta)=\theta''} \exp \left\{ - \int_0^\beta dt \left[\frac{1}{2} (\dot{\theta}(t))^2 + il\dot{\theta}(t) \right] \right\}, \quad (3.32)$$

where the trajectories go from θ' to θ'' while winding an arbitrary number of times around the circle. Since l is an integer, in the path integral the additional contribution is independent of the winding number.

Evaluation of the path integral. We first solve the classical equation of motion. We now find the family of trajectories which go from θ' to θ'' ,

$$\theta_c(t) = \theta' + t(\theta'' - \theta' + 2\pi q)/\beta, \quad \text{with } q \in \mathbb{Z}. \quad (3.33)$$

Trajectories corresponding to different winding numbers are topologically distinct, that is, they cannot be related by continuous deformation. Hence, they have all to be taken into account, because Brownian paths are continuous and, thus, fluctuations around one trajectory do not include any other one.

We then shift $\theta(t)$, setting $\theta(t) = \theta_c(t) + u(t)$. The path integral (3.31) becomes a sum of contributions of the form,

$$e^{-\beta H_l} = \sum_{q=-\infty}^{+\infty} \mathcal{N}(\beta) \int_0^{2\pi} \frac{d\theta' d\theta''}{(2\pi)^2} \exp \left[-\frac{1}{2\beta} (\theta'' - \theta' + 2\pi q)^2 + il(\theta' - \theta'') \right]. \quad (3.34)$$

The normalization $\mathcal{N}(\beta)$ is given by a path integral which is independent of θ' , θ'' and q . Since the integration over $u(t)$ sums fluctuations around the classical trajectory, it is expected that the angular character of $u(t)$ is irrelevant and, therefore,

$$\mathcal{N}(\beta) = \sqrt{2\pi/\beta}.$$

We set

$$\theta'' = \theta' - 2\pi q + \varphi.$$

Then,

$$\exp[-\beta H_l] = \frac{1}{\sqrt{2\pi\beta}} \sum_{q=-\infty}^{+\infty} \int_{2\pi q}^{2\pi(q+1)} d\varphi e^{-il\varphi} \exp \left(-\frac{1}{2\beta} \varphi^2 \right).$$

The sum yields the Gaussian integral

$$\exp(-\beta H_l) = \frac{1}{\sqrt{2\pi\beta}} \int_{-\infty}^{+\infty} d\varphi \exp[-(il\varphi + \varphi^2/2\beta)] = e^{-\beta l^2/2}$$

which is the exact spectrum (3.30).

It has been possible to perform an exact calculation because the $O(2)$ group is Abelian, and the group manifold flat. The discussion of the general $O(N)$ group is more involved, as we now show, because the group manifold has a curvature.

3.5 The spectrum of the $O(N)$ -symmetric rigid rotator

The Hamiltonian of the $O(N)$ rigid rotator can be written as

$$H = \frac{1}{2} \mathbf{L}^2, \quad (3.35)$$

where the vector \mathbf{L} , the angular momentum, represents the set of generators of the Lie algebra of the $O(N)$ group. If the sphere is parametrized in terms of coordinates q^i , the Hamiltonian takes the form

$$H = \frac{1}{2} \sum_{i,j} g^{ij}(\mathbf{q}) p_i p_j, \quad (3.36)$$

where $g^{ij}(\mathbf{q})$ is the inverse of the metric tensor on the sphere (see also Section 34.9).

For example, if the sphere is parametrized locally by a vector \mathbf{r} in \mathbb{R}^N of unit length and components $(\mathbf{q}, (1 - \mathbf{q}^2)^{1/2})$, the inverse metric tensor reads

$$g^{ij}(\mathbf{q}) = \delta^{ij} - q^i q^j. \quad (3.37)$$

According to the discussion of Section 3.3.2, the corresponding path integral representation of $e^{-\beta H}$ is then

$$\langle \mathbf{q}'' | e^{-\beta H} | \mathbf{q}' \rangle = \int \left[\sqrt{g(\mathbf{q}(t))} d\mathbf{q}(t) \right] \exp \left[-\frac{1}{2} \int_0^\beta dt \sum_{i,j} g_{ij}(\mathbf{q}(t)) \dot{q}^i(t) \dot{q}^j(t) \right], \quad (3.38)$$

in which $g(\mathbf{q})$ is the determinant of the matrix g_{ij} . The contribution to the measure coming from the Gaussian integration over the momenta p_i has formally generated the invariant measure on the sphere. The path integral (3.38) can also be expressed in terms of a vector \mathbf{r} of unit length as

$$\langle \mathbf{r}'' | e^{-\beta H} | \mathbf{r}' \rangle = \int_{\mathbf{r}(0)=\mathbf{r}'}^{\mathbf{r}(\beta)=\mathbf{r}''} [d\mathbf{r}(t) \delta(1 - \mathbf{r}^2(t))] \exp \left[-\frac{1}{2} \int_0^\beta dt \dot{\mathbf{r}}^2(t) \right]. \quad (3.39)$$

Because, for $N > 2$, the sphere has curvature, it cannot be mapped onto a flat space and, thus, the Hamiltonian can no longer be mapped onto a free Hamiltonian. On the other hand, all loops on the sphere are contractible and there are no longer different topological classes.

High temperature expansion. We call θ the angle between \mathbf{r}' and \mathbf{r}'' :

$$\cos \theta = \mathbf{r}' \cdot \mathbf{r}'', \quad 0 \leq \theta \leq \pi.$$

We introduce a matrix $\mathbf{R}(t)$, which acts on $\mathbf{r}(t)$ and rotates \mathbf{r}' onto \mathbf{r}'' in the plane $(\mathbf{r}', \mathbf{r}'')$ in a time β . Its restriction to the two-dimensional $(\mathbf{r}', \mathbf{r}'')$ plane has the form

$$\begin{bmatrix} \cos(\theta t/\beta) & \sin(\theta t/\beta) \\ -\sin(\theta t/\beta) & \cos(\theta t/\beta) \end{bmatrix}.$$

It is the identity in the subspace orthogonal to the $(\mathbf{r}', \mathbf{r}'')$ plane.

We then change variables, $\mathbf{r}(t) \mapsto \boldsymbol{\rho}(t)$, setting

$$\mathbf{r}(t) = \mathbf{R}(t)\boldsymbol{\rho}(t).$$

We call u and v the two components of $\boldsymbol{\rho}$ in the $(\mathbf{r}', \mathbf{r}'')$ plane, u being the component along \mathbf{r}' and $\boldsymbol{\rho}_T$ the component in the orthogonal subspace. With this notation, we find

$$\langle \mathbf{r}'' | e^{-\beta H} | \mathbf{r}' \rangle = \int_{\boldsymbol{\rho}(0)=\mathbf{r}'}^{\boldsymbol{\rho}(\beta)=\mathbf{r}''} [d\boldsymbol{\rho}(t) \delta(1 - \boldsymbol{\rho}^2(t))] \exp [-\mathcal{S}(\boldsymbol{\rho})], \quad (3.40)$$

with

$$\begin{aligned} \mathcal{S}(\boldsymbol{\rho}) = & \frac{1}{2} \int_0^\beta dt \left[\dot{\boldsymbol{\rho}}_T^2(t) + \dot{u}^2(t) + \dot{v}^2(t) + \frac{\theta^2}{\beta^2} (u^2(t) + v^2(t)) \right. \\ & \left. + 2 \frac{\theta}{\beta} (\dot{v}(t)u(t) - \dot{u}(t)v(t)) \right], \end{aligned} \quad (3.41)$$

and the constraint

$$u^2 + v^2 + \boldsymbol{\rho}_T^2 = 1. \quad (3.42)$$

In contrast with the Abelian example, where an exact calculation is possible, we can here perform only a small β (large temperature) expansion, corresponding to the Wentzel–Kramers–Brillouin (WKB) or semi-classical limit, and valid for large quantum numbers. We take into account only fluctuations around the classical solution $u = 1$, $\rho_T = 0$, $v = 0$, neglecting exponentially small contributions in β^{-1} .

We eliminate the variable u from the action (3.41) by solving the constraint (3.42):

$$u = (1 - v^2 - \rho_T^2)^{1/2},$$

and expand the action in powers of ρ_T and v . The leading order is

$$\mathcal{S}(\rho) = \theta^2/2\beta,$$

a result that shows that the calculation is valid for $\theta = O(\sqrt{\beta})$. The next order is given by the Gaussian integration and requires the quadratic terms in ρ_T and v ,

$$\frac{1}{2} \int_0^\beta dt [\dot{\rho}_T^2(t) - (\theta^2/\beta^2)\rho_T^2(t) + \dot{v}^2(t)].$$

The integral over $v(t)$ is independent of θ and can be absorbed into the normalization. The integrals over the components of ρ_T factorize and give identical results: the integral over ρ_T is the integral over one component to the power $(N - 2)$. Since each component satisfies the conditions

$$\rho_i(0) = \rho_i(\beta) = 0,$$

we expand the functions $\rho_i(t)$ on the appropriate orthonormal basis:

$$\rho_i(t) = \sqrt{\frac{2}{\beta}} \sum_{n>0} \rho_{in} \sin(n\pi t/\beta).$$

The Gaussian integral over the variables ρ_{in} then yields

$$\langle \mathbf{r}'' | e^{-\beta H} | \mathbf{r}' \rangle \sim K(\beta) e^{-\theta^2/2\beta} \left[\prod_{n>0} \left(1 - \frac{\theta^2}{n^2 \pi^2} \right) \right]^{-(N-2)/2}.$$

The normalization constant $K(\beta) = (2\pi\beta)^{-(N-1)/2}$ is independent of θ . The infinite product can be calculated:

$$\prod_{n>0} \left(1 - \frac{\theta^2}{n^2 \pi^2} \right) = \frac{\sin \theta}{\theta},$$

and, therefore,

$$\langle \mathbf{r}'' | e^{-\beta H} | \mathbf{r}' \rangle \sim K(\beta) \left(\frac{\theta}{\sin \theta} \right)^{(N-2)/2} e^{-\theta^2/2\beta}. \quad (3.43)$$

To extract the eigenvalues of H , we project the expression onto the orthogonal polynomials $P_l^N(\cos \theta)$ associated with the $O(N)$ group, which satisfy

$$\int_0^\pi d\theta (\sin \theta)^{N-2} P_l^N(\cos \theta) P_{l'}^N(\cos \theta) = \delta_{ll'}, \quad (3.44)$$

and which are proportional to the Gegenbauer polynomials $C_l^{(N-2)/2}$.

For β small, we need only the small θ expansion of these polynomials:

$$P_l^N(\cos \theta) = P_l^N(1) \left(1 - \frac{l(l+N-2)}{2(N-1)} \theta^2 + O(\theta^4) \right).$$

If we assume that, to each value of l , corresponds only one eigenvalue E_l of H , then the following expansion holds,

$$\begin{aligned} e^{-\beta E_l} &\propto K(\beta) \int_0^\pi d\theta P_l^N(\cos \theta) (\theta \sin \theta)^{(N-2)/2} e^{-\theta^2/(2\beta)} \\ &= e^{-\beta E_0} \left(1 - \frac{1}{2} l(l+N-2)\beta + O(\beta^2) \right), \end{aligned} \quad (3.45)$$

and, therefore,

$$E_l = E_0 + \frac{1}{2} l(l+N-2) + O(\beta). \quad (3.46)$$

Since E_l is independent of β , we can infer from this calculation the exact result, up to an additive constant E_0 .

Ambiguities and symmetry. Concerning this calculation, a comment is in order: we have explained in Section 3.3 that the path integrals (3.38, 3.39) are not defined, because the measure gives formally divergent contributions. We have stated that these divergences are cancelled by divergences in perturbation theory. As a consequence, the resulting expressions are ambiguous, and these ambiguities reflect the problem of operator ordering in the quantization of a classical Hamiltonian. Still, we have obtained here some non-trivial results. The reason is that at every stage of the calculation explicit $O(N)$ invariance has been maintained. This chooses implicitly among all possible definitions of the path integral a subclass that corresponds to an $O(N)$ -symmetric quantized Hamiltonian.

We see later, when discussing the non-linear σ -model (Chapter 19), that such a Hamiltonian is fully determined up to an additive constant. The ambiguities of the quantization here are entirely contained in E_0 .

A3 Quantization. Topological actions: Quantum spins, magnetic monopoles

Convention. In the Appendix, in order to recover the usual expressions of classical mechanics, we work in *real time*.

A3.1 Symplectic form and quantization: General remarks

In the Hamiltonian formulation of classical mechanics, the action has the form (3.6),

$$\mathcal{A}(p, q) = \int_{t'}^{t''} dt [p(t)\dot{q}(t) - H(p(t), q(t), t)] . \quad (A3.1)$$

The term $\int p\dot{q} dt$ also represents the area in phase space between the classical trajectory C and the axis $p = 0$. In the notation of exterior calculus, it can be written as

$$\int_{\partial D} p(t)\dot{q}(t) dt = \int_D dp \wedge dq , \quad (A3.2)$$

in which ∂D , the boundary of the domain D , contains the classical trajectory C and a fixed reference curve. If we now parametrize phase space differently, introducing new coordinates u_α , the right-hand side of the equation becomes

$$\int_D dp \wedge dq = \int_D \sum_{\alpha, \beta} \omega_{\alpha\beta}(u) du_\alpha \wedge du_\beta .$$

In the terminology of forms $\omega = \sum_{\alpha, \beta} \omega_{\alpha\beta} du_\alpha \wedge du_\beta$ is a 2-form (see Section 1.5), which, by construction, is obtained by differentiating a 1-form; here,

$$\omega = d\omega' , \quad (A3.3)$$

$$\omega' = p(u) \sum_\alpha \frac{dq}{u_\alpha} du_\alpha = \sum_\alpha \omega_\alpha du_\alpha , \quad (A3.4)$$

and is called the *symplectic form*.

Since the square of the operator d acting on differential forms vanishes, the form ω is *closed*, that is, it satisfies

$$d\omega = 0 . \quad (A3.5)$$

Example. In Section 4.1, we discuss the holomorphic formalism. We introduce a complex parametrization of phase space (equation (4.39)) of the form

$$p - iq = -i\sqrt{2}z , \quad p + iq = i\sqrt{2}\bar{z} .$$

In terms of z, \bar{z} , the symplectic form becomes

$$dp \wedge dq = \frac{1}{i} dz \wedge d\bar{z} .$$

Preceding considerations immediately generalize to several degrees of freedom. Let u_α be $2n$ variables parametrizing a phase space for n degrees of freedom. The action \mathcal{A} in the Hamiltonian formulation can be written as

$$\mathcal{A}(u) = \int_D \omega - \int_{\partial D} dt H(u(t), t), \quad (A3.6)$$

where ω , a symplectic form, is a *closed 2-form*,

$$\omega(u) = \sum_{\alpha, \beta} \omega_{\alpha\beta} du_\alpha \wedge du_\beta, \quad \text{with } d\omega = 0.$$

The condition $d\omega = 0$ is sufficient to ensure that the classical equations of motion depend only on the boundary ∂D of the domain D , but not on the interior, as was obvious for the initial action (A3.1). The equations of motion then take the form

$$\sum_{\beta} \omega_{\alpha\beta}(u) \dot{u}_\beta(t) = \frac{\partial H}{\partial u_\alpha(t)}. \quad (A3.7)$$

Locally, equation (A3.5) can be integrated in the same way as $dp \wedge dq$ can be integrated into $p dq$. However, if phase space has non-trivial topological properties, it cannot always be integrated globally, that is, the symplectic form is not *exact*. This property has peculiar consequences in QM, since the path integral involves the action explicitly in the form $e^{i\mathcal{A}/\hbar}$. For the path integral to make sense, this phase factor must be independent of the choice of the action \mathcal{A} .

This problem shows up in the example of the quantization of spin degrees of freedom discussed in Section A3.3.1.

Canonical invariance of the symplectic form. We consider the symplectic form

$$\Omega = \sum_i dp_i \wedge dq_i.$$

We introduce a smooth function $\Sigma(Q, q)$ and the transformation $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{P}, \mathbf{Q})$ defined by,

$$p_i = \partial\Sigma/\partial q_i, \quad P_i = -\partial\Sigma/\partial Q_i. \quad (A3.8)$$

First, we implement the transformation $\mathbf{p} \mapsto \mathbf{Q}$. The symplectic form becomes

$$\Omega = \sum_{i,j} \frac{\partial^2 \Sigma}{\partial q_i \partial Q_j} dQ_j \wedge dq_i.$$

Then, we implement the transformation $\mathbf{q} \mapsto \mathbf{P}$:

$$dP_j = \sum_i \frac{\partial^2 \Sigma}{\partial Q_j \partial q_i} dq_i,$$

and, therefore,

$$\Omega = \sum_j dP_j \wedge dQ_j.$$

A3.2 Classical equations of motion and quantization

We consider a general Hamiltonian $H(\mathbf{p}, \mathbf{q}, t)$ corresponding to the classical action,

$$\mathcal{A}(\mathbf{p}, \mathbf{q}) = \int_{T'}^T dt [\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - H(\mathbf{p}(t), \mathbf{q}(t); t)], \quad (A3.9)$$

with $q_i(T') = x'_i$ and $q_i(T) = x_i$.

We specialize to Hamiltonian systems (called integrable) for which the classical equations of motion can be solved with arbitrary boundary conditions. In particular, this implies that one can calculate explicitly the action for a classical trajectory as a function of initial time T' and position \mathbf{x}' , and final time T and position \mathbf{x} .

The action A_c corresponding to a trajectory $\{\mathbf{p}_c(t), \mathbf{q}_c(t)\}$ in phase space, the solution of the classical of motions is a function of the initial and final positions, and time:

$$A_c(\mathbf{x}', \mathbf{x}; T) = \mathcal{A}(\mathbf{p}_c, \mathbf{q}_c). \quad (A3.10)$$

The analytic continuation of all expressions to imaginary time is simple.

It is well-known, and verified again in the next section (see equation (A3.19)), that the classical action A_c satisfies the Hamilton–Jacobi equations.

A3.2.1 Preliminary remarks

We first recall a few classical results. We denote by $S(\mathbf{Q}, \mathbf{q}; t)$ a function satisfying the Hamilton–Jacobi equations:

$$\frac{\partial S}{\partial t} = -H\left(\frac{\partial S}{\partial \mathbf{q}}, \mathbf{q}; t\right), \quad (A3.11)$$

with $S(\mathbf{Q}, \mathbf{Q}; T') = 0$, and with additional implicit boundary conditions at $t = T'$, which will be explained. We use S to generate a time-dependent canonical transformation in phase space, transforming the set (\mathbf{p}, \mathbf{q}) into (\mathbf{P}, \mathbf{Q}) :

$$p_i = \partial S / \partial q_i, \quad P_i = -\partial S / \partial Q_i. \quad (A3.12)$$

The implicit boundary conditions follow from the conditions

$$t = T' \Rightarrow P_i = p_i \quad \text{and} \quad Q_i = q_i. \quad (A3.13)$$

We use the relation (A3.11) and apply the first transformation (A3.12) to the action (A3.9). We obtain

$$\mathcal{A}(\mathbf{p}, \mathbf{q}) = \int_{T'}^T \left(\sum_i \frac{\partial S}{\partial q_i} \dot{q}_i(t) + \frac{\partial S}{\partial t} \right) dt. \quad (A3.14)$$

The quantities \mathbf{q} and \mathbf{Q} are time dependent. Expression (A3.14) can be rewritten as

$$\mathcal{A}(\mathbf{p}, \mathbf{q}) = \int_{T'}^T \frac{d}{dt} S(\mathbf{q}(t), \mathbf{Q}(t); t) dt - \int_{T'}^T \sum_i \frac{\partial S}{\partial Q_i} \dot{Q}_i(t) dt. \quad (A3.15)$$

Finally, using the second equation (A3.12), one obtains

$$\mathcal{A}(\mathbf{p}, \mathbf{q}) = S(\mathbf{x}, \mathbf{Q}(T); T) + \int_{T'}^T P_i(t) \dot{Q}_i(t) dt. \quad (A3.16)$$

The equations of motion are now trivial,

$$\dot{Q}_i = 0 \implies Q_i(t) = Q_i(T'). \quad (\text{A3.17})$$

The conditions (A3.13) determine the solution

$$Q_i(t) = x'_i. \quad (\text{A3.18})$$

Therefore, we have shown that

$$A_c(\mathbf{x}', \mathbf{x}; T) \equiv S(\mathbf{x}, \mathbf{x}'; T), \quad (\text{A3.19})$$

and found a canonical transformation which maps the initial Hamiltonian system onto a trivial one with a vanishing Hamiltonian. By performing an additional inverse transformation based on a standard Hamiltonian, like a free Hamiltonian of the form

$$H = \frac{1}{2} \sum p_i^2, \quad (\text{A3.20})$$

or a harmonic oscillator,

$$H = \frac{1}{2} \sum_i (p_i^2 + q_i^2), \quad (\text{A3.21})$$

we can map the original Hamiltonian system onto any convenient Hamiltonian.

A3.2.2 Liouville measure

The transformation (A3.12) leaves the Liouville measure in phase-space invariant. To prove it, we perform the transformation in two steps (a proof quite similar to the proof of canonical invariance given in Section A3.1). First, we go from p_i to Q_i :

$$\prod_i (dp_i dq_i) = \prod_i (dq_i dQ_i) \det \left(\frac{\partial^2 S}{\partial q_k \partial Q_l} \right). \quad (\text{A3.22})$$

We then eliminate q_i in favour of P_i :

$$\frac{\partial q_i}{\partial P_j} = \left(\frac{\partial P}{\partial q} \right)_{ij}^{(-1)} = - \left(\frac{\partial^2 S}{\partial q \partial Q} \right)_{ij}^{(-1)}. \quad (\text{A3.23})$$

Therefore, the second Jacobian cancels the first one (the argument also shows the invariance of the symplectic form $\sum_i dp_i \wedge dq_i$ (see also Section A3.1)).

The analysis suggests that we can perform two transformations (A3.12) on the path integral representation of the evolution operator, to reduce it to a path integral corresponding to a standard system for which the evolution operator is exactly known:

$$\langle \mathbf{x}' | U(T', T) | \mathbf{x} \rangle = \int \prod_i [dq_i(t) dp_i(t)] \exp [i\mathcal{A}(\mathbf{p}, \mathbf{q})], \quad (\text{A3.24})$$

with $q_i(T') = x'_i, \quad q_i(T) = x_i.$

In this way, it would seem that we are able to calculate exactly the evolution operator of any system for which we know how to solve the classical equations of motion with arbitrary boundary conditions. In particular, this would apply to systems with one degree of freedom, with a Hamiltonian H of the form

$$H = p^2/2m + V(q). \quad (\text{A3.25})$$

Unfortunately, it is easy to verify that the result is wrong at least in one-dimensional QM. Actually, the whole procedure is somewhat ill-defined. Indeed, changes of variables on the path integral on phase space are even more ambiguous than transformations on ordinary path integrals in configuration space. We have given some indications of this problem in Section 3.1. Let us just mention that if we discretize time, we discover that the transformation is not really canonical, because to a variable $q(t_k)$ corresponds a momentum variable $p(t_k + \Delta t)$ of a slightly displaced time. This effect, invisible in the naive continuum limit, completely changes the result. It is thus necessary to work with the discretized form of the phase-space path integral, but this is cumbersome. On the other hand, the situation seems to be more favourable in quantum field theory. The problem we have described above comes mainly from commutation of quantum operators. We emphasize in Chapter 7 that the commutators are infinite in field theory and disappear in the renormalization. Therefore, one may expect, as this has been verified on a few examples with enough symmetries, that the semi-classical approximations of classically integrable field theories reproduce features of the exact solution.

A3.3 Topological actions

We now indicate how non-trivial topological properties of phase space, irrelevant from the point of view of classical mechanics, affect quantization.

In Section 3.4, we have evaluated the path integral in an example where space has non-trivial topological properties. We now discuss two other examples where topology plays an essential role, in the sense that the action which generates the classical equations of motion cannot be globally defined. We first quantize angular momentum operators in the path integral formalism, in a fixed representation. One of the peculiarities of this system is that phase space itself has non-trivial topological properties. The second example is provided by the magnetic monopole which gives a non-trivial topological structure to ordinary space. In both examples, one calls the action topological. Because a topological action cannot be globally defined, its amplitude is quantized, a property specific to QM. Indeed, in classical mechanics, a multiplication of the action by a constant does not affect the equations of motion.

A3.3.1 Spin dynamics and quantization

Classical spin dynamics. We consider a vector \mathbf{S} in three dimensions of fixed length s : $\mathbf{S}^2 = s^2$. We give \mathbf{S} the simple dynamics,

$$\frac{d\mathbf{S}}{dt} = \mathbf{H} \times \mathbf{S}, \quad (A3.26)$$

in which \mathbf{H} is a constant vector. This equation is first order in time and involves two independent variables corresponding to a point on the sphere. These variables can be considered as a position and its conjugate momentum. Therefore, phase space is isomorphic to the sphere S_2 . The Hamiltonian is simply

$$\mathcal{H}(\mathbf{S}) = \mathbf{H} \cdot \mathbf{S}. \quad (A3.27)$$

More precisely, it is simple to verify that if one parametrizes \mathbf{S} as

$$\mathbf{S} = s(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),$$

then equation (A3.26) is generated by the action

$$\mathcal{A}(\theta, \varphi) = \int [s\dot{\varphi}(t) \cos \theta(t) - \mathcal{H}(\varphi(t), \theta(t))] dt. \quad (A3.28)$$

We note in the expression that $\cos \theta$ and φ play the role of conjugate variables. However, two remarks are in order: first we have integrated the symplectic 2-form but, for this purpose, we have been forced to introduce a parametrization of the sphere that is singular at $\theta = 0$ and π . When the trajectory contains the north or the south pole of the sphere, the integral is not defined. The 2-form cannot be integrated globally into a 1-form because the integral $\int d\varphi d\cos \theta$, which is the area on the sphere, is defined only mod (4π) .

In classical mechanics, these properties play no direct role, since only equations of motion are physical.

Note that the symplectic form has other useful representations (ϵ_{ijk} is the completely antisymmetric tensor, with $\epsilon_{123} = 1$),

$$d\cos \theta \wedge d\varphi = \frac{1}{2}s^{-3} \sum_{i,j,k} \epsilon_{ijk} S_i dS_j \wedge dS_k = 2i \sum_i dz_i \wedge d\bar{z}_i,$$

where z_i is a two-component complex vector of length 1, corresponding to the isomorphism between S_2 and the symmetric space CP_1 (see Section 39.5):

$$\mathbf{S} = s \sum_{i,j} \bar{z}_i \boldsymbol{\sigma}_{ij} z_j, \text{ with } \sum_i \bar{z}_i z_i = 1,$$

where $\boldsymbol{\sigma}$ is the set of Pauli matrices. In a special *gauge*, the vector \mathbf{z} can also be written as

$$\begin{aligned} z_1 &= e^{i\varphi/2} \cos(\theta/2), \\ z_2 &= e^{-i\varphi/2} \sin(\theta/2). \end{aligned}$$

A3.3.2 Quantization of spin degrees of freedom

We consider the path integral representation of the corresponding evolution operator in QM. The action itself now appears explicitly and the problem discussed in the preceding section becomes relevant. The path integral exists only if the integrand $e^{i\mathcal{A}/\hbar}$ is defined. Since the total area is defined only mod (4π) , the integrand must be invariant under such a change. This implies that $4\pi s$ must be a multiple of $2\pi\hbar$: the parameter s/\hbar is quantized and can take only half-integer values. This is a generic property in QM: the amplitude of *topological* contributions to the action, that is, contributions that are not globally defined, is *quantized*. The magnetic monopole of Section A3.3.3, or the Chern–Simons term in three-dimensional non-Abelian gauge theories, provide other examples of such a situation.

In the parametrization (A3.28), the action can be written as (we now set $\hbar = 1$)

$$\mathcal{A}(\theta, \varphi) = \int [\dot{\varphi}(t)(\gamma + \cos \theta(t)) - \mathcal{H}(\varphi(t), \theta(t))] dt, \quad (\text{A3.29})$$

an expression that differs from expression (A3.28) only by a total derivative: by choosing $\gamma = 1/2, 0$ for s half-integer, integer, respectively, one renders $e^{i\mathcal{A}}$ regular near $\theta = 0, \pi$.

To relate this action to the usual operator formulation of the angular momentum relations, we first note that, classically,

$$S_{\pm} = e^{\pm i\varphi} (s^2 - S_z^2)^{1/2}.$$

After quantization, S_z becomes p_φ , the conjugate momentum of the angular variables φ :

$$S_z \equiv p_\varphi = \frac{1}{i} \frac{d}{d\varphi}.$$

It can be verified that the quantum operators S_\pm can be written as

$$S_\pm = e^{\pm i\varphi/2} (s^2 - p_\varphi^2)^{1/2} e^{\pm i\varphi/2}.$$

Then, using

$$j(j+1) = \mathbf{S}^2 = S_z^2 + S_+ S_- S_z = s^2 - 1/4,$$

we find the relation between the angular momentum j and the parameter s ,

$$s = j + 1/2. \quad (A3.30)$$

In particular, since s is quantized, we recover a property of QM: quantization of spin.

If we denote by m the eigenvalues of S_z , we observe that, in the φ -configuration space, the corresponding eigenvectors have the form $e^{im\varphi}$, and the projector K on the basis is

$$K(\varphi'', \varphi') = \sum_{m=-j}^{m=j} e^{im(\varphi' - \varphi'')} = \frac{\sin[(j+1/2)(\varphi' - \varphi'')]}{\sin[(\varphi' - \varphi'')/2]}. \quad (A3.31)$$

We can compare this expression with the short-time path integral representation, which leads to

$$\begin{aligned} K(\varphi'', \varphi') &\propto \sum_n \int_{-s}^s dp_\varphi e^{ip_\varphi(\varphi' - \varphi'' + 2n\pi)}, \\ &\propto \sin[s(\varphi' - \varphi'')] \sum_n \frac{2(-1)^n}{(\varphi' - \varphi'' + 2n\pi)}. \end{aligned}$$

The sum over n must be regularized, but the factor $\sin s(\varphi' - \varphi'')$ is consistent with equation (A3.31) and the identification (A3.30).

A final word of caution: although the path integral quantization of spin variables is quite useful to study the classical limit (which is also the limit $s \rightarrow \infty$), the problem of operator ordering in products leads to the usual ambiguities in explicit calculations.

A3.3.3 The magnetic monopole

Electromagnetism provides another example of the situation encountered in Section A3.3.1. In a magnetic system, the only physical quantity which appears in the classical equations of motion is the magnetic field. The contribution of the magnetic term to the action can be generally written as the integral of a 2-form (involving the magnetic field), since this form is closed:

$$\mathcal{A}_{\text{mag.}} = e \int \sum_{i,j} F_{ij}(x) dx_i \wedge dx_j, \text{ with } F_{ij}(x) = \sum_k \epsilon_{ijk} B_k(x).$$

Because the 2-form F is closed, one can integrate it locally by introducing a vector potential, which is a 1-form (see equation (3.22)):

$$F_{ij}(x) = \partial_i A_j(x) - \partial_j A_i(x), \Rightarrow \mathcal{A}_{\text{mag.}} = e \int \mathbf{A}(\mathbf{x}) \cdot d\mathbf{x}.$$

However, if the 2-form F is not exact, the vector potential cannot be globally defined. This is precisely what happens when a magnetic field is generated by a magnetic monopole.

The formal duality symmetry between magnetic and electric fields in Maxwell's equations has led Dirac to speculate about the existence of yet undiscovered isolated magnetic charges, magnetic equivalents of electric charges. An isolated magnetic charge g creates a magnetic field \mathbf{B} of the form

$$\mathbf{B} = g \frac{\mathbf{x}}{|\mathbf{x}|^3} .$$

The field \mathbf{B} is also a singular solution to the free static Maxwell's equations. It has an infinite energy and hence, in this form, it is irrelevant to physics. However, in non-Abelian gauge theories with spontaneous symmetry breaking, finite energy solutions (solitons) can be found, which coincide at large distance with Dirac's magnetic monopoles.

The integral of the magnetic field over a closed surface containing the magnetic charge is $4\pi g$, as one immediately verifies by using polar coordinates $\{r, \theta, \varphi\}$ with the monopole at the origin:

$$\int \sum_{i,j} F_{ij}(x) dx_i \wedge dx_j = g \int \frac{r}{r^3} \times r^2 d\cos \theta d\varphi = 4\pi g .$$

If the vector potential could be globally defined, the integral would obviously vanish. More directly, if one tries to calculate the corresponding vector potential, one finds, in a family of gauges,

$$A_i(x) = g \sum_{j,k} \epsilon_{ijk} n_k x_j \frac{\mathbf{n} \cdot \mathbf{x}}{r(r^2 - (\mathbf{n} \cdot \mathbf{x})^2)} ,$$

where \mathbf{n} is a constant unit vector. We observe that the potential is singular along the line of direction \mathbf{n} passing through the origin. This line of singularities is not physical and can be displaced, but not removed.

Again, this property has no classical consequences. However, in QM, since the classical action can be defined only $(\text{mod } 4\pi eg)$, the weight factor e^{iA} is only defined if

$$4\pi eg = 0 \pmod{2\pi} \Rightarrow 2eg = \text{integer} ,$$

in which we recognize Dirac's quantization condition.

Note that when this condition is fulfilled, parallel transport (see Section 21.7) is globally defined in \mathbb{R}^3 .

4 Quantum statistical physics: Functional integration formalism

In this chapter, we introduce the representation by functional integrals of the density matrix at thermal equilibrium in non-relativistic quantum mechanics (QM), with many degrees of freedom, in the grand canonical formulation.

A class of quantum Hamiltonians can be expressed in terms of creation and annihilation operators, instead of the more usual position and momentum operators, a method adapted to the study of perturbed harmonic oscillators. In the holomorphic formalism, these operators act by multiplication and differentiation on a Hilbert space of analytic functions. Alternatively, they can also be represented by kernels, functions of complex variables z, \bar{z} that, in the classical limit correspond to a complex parametrization of phase space. The formalism is adapted to the description of many-body boson systems.

A path integral representation of the density matrix at thermal equilibrium, where paths belong to complex spaces, instead of the more usual position–momentum phase space, corresponds to this formalism.

A parallel formalism can then be set-up to describe systems with many fermion degrees of freedom, with Grassmann variables replacing complex variables.

Both formalisms can be generalized to quantum gases of Bose and Fermi particles in the grand canonical formulation. *Field integral* representations of the corresponding quantum partition functions are then derived.

4.1 One-dimensional QM: Holomorphic representation

We briefly recall the main ideas and properties of the holomorphic representation. To motivate the construction, we consider the quantum harmonic oscillator

$$H_0 = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{q}^2, \quad \omega > 0, \quad (4.1)$$

where \hat{q}, \hat{p} are the position and momentum operators, respectively, with the commutation relation $[\hat{q}, \hat{p}] = i\hbar$. We introduce the annihilation and creation operators a, a^\dagger defined by

$$\hat{p} - i\omega\hat{q} = -i\sqrt{2\hbar\omega}a, \quad \hat{p} + i\omega\hat{q} = i\sqrt{2\hbar\omega}a^\dagger \quad \Rightarrow [a, a^\dagger] = 1. \quad (4.2)$$

In terms of a, a^\dagger , the Hamiltonian takes the standard form $\hbar\omega(a^\dagger a + 1/2)$. In what follows, we omit the constant energy shift $\hbar\omega/2$ and consider

$$H_0 = \hbar\omega a^\dagger a. \quad (4.3)$$

We then introduce a complex variable z and represent a^\dagger and a by the operators z and $\partial/\partial z$, acting by multiplication and differentiation on a complex vector space of analytic functions of the form

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{n!} f_n z^n, \quad (4.4)$$

which have the same commutation relations

$$a^\dagger \mapsto z, \quad a \mapsto \partial/\partial z, \quad [\partial/\partial z, z] = 1. \quad (4.5)$$

The Hamiltonian H_0 is then represented by

$$H_0 = \hbar\omega z \frac{\partial}{\partial z}. \quad (4.6)$$

In this representation, the eigenfunctions of H_0 are the monomials z^n , since

$$\hbar\omega z \frac{\partial}{\partial z} z^n = n\hbar\omega z^n.$$

The action of the operator $U_0(t) = e^{-H_0 t/\hbar}$ (again we first keep the \hbar normalization of real-time evolution) on analytic functions is

$$U_0(t)f(z) = e^{-H_0 t/\hbar} f(z) = f(e^{-\omega t} z). \quad (4.7)$$

4.1.1 Hilbert space of analytic functions

The vector space of analytic functions can be endowed with a structure of Hilbert space [15]: the scalar product of two functions g and f is defined by

$$(g, f) = \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} \overline{g(z)} f(z), \quad (4.8)$$

where this type of complex integrals has been defined and discussed in Section 1.4.

With this scalar product, the operators z and $\partial/\partial z$ are Hermitian conjugate and the eigenfunctions of the harmonic oscillator form an orthogonal basis:

$$\int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} \bar{z}^n z^m = n! \delta_{mn}, \quad (4.9)$$

a result that follows, for instance, from Wick's theorem (1.30). In particular, the norm of a function of the form (4.4) is finite if

$$(f, f) = \sum_n \frac{1}{n!} |f_n|^2 < \infty.$$

Therefore, square integrable functions belong to a subset of entire analytic functions.

Using the orthogonality relations (4.9), one defines an analogue of Dirac's δ -function:

$$f(0) = \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} f(z). \quad (4.10)$$

4.1.2 Operator kernels

Since the functions $z^n / \sqrt{n!}$ form an orthonormal basis, the identity operator can be represented by the kernel

$$e^{z\bar{z}} = \sum_{n=0} \frac{\bar{z}^n}{\sqrt{n!}} \frac{z^n}{\sqrt{n!}}.$$

This property is expressed by the identity

$$\int \frac{d\bar{z}' dz'}{2i\pi} e^{\bar{z}' z} e^{-z'\bar{z}'} f(z') = f(z), \quad (4.11)$$

which is a direct consequence of the identity (4.10).

More generally, any operator function of a and a^\dagger can first be written in ‘normal’ order by commuting all creation operators on the left of all annihilation operators. It becomes a linear combination of operators of the form

$$a^{\dagger m} a^n \mapsto z^m (\partial/\partial z)^n,$$

in the representation (4.5). To a normal-ordered operator $O(z, \partial/\partial z)$, we then associate a kernel $\mathcal{O}(z, \bar{z})$, for which we also use the notation of matrix elements $\langle z | \mathcal{O} | \bar{z} \rangle$ (the bra–ket notation of QM) obtained by acting on the identity,

$$\mathcal{O}(z, \bar{z}) \equiv \langle z | \mathcal{O} | \bar{z} \rangle = O(z, \partial/\partial z) e^{z\bar{z}} = O(z, \bar{z}) e^{z\bar{z}}. \quad (4.12)$$

Acting with $O(z, \partial/\partial z)$ on both sides of equation (4.11), one infers

$$(Of)(z) = \int \frac{d\bar{z}' dz'}{2i\pi} \mathcal{O}(z, \bar{z}') e^{-z'\bar{z}'} f(z').$$

The kernel associated with the product of the two operators is then given by

$$\int \frac{d\bar{z}' dz'}{2i\pi} \langle z | \mathcal{O}_2 | \bar{z}' \rangle e^{-z'\bar{z}'} \langle z' | \mathcal{O}_1 | \bar{z} \rangle = \langle z | \mathcal{O}_2 \mathcal{O}_1 | \bar{z} \rangle. \quad (4.13)$$

From the form (4.8) of the scalar product, one also infers the expression of the trace of an operator,

$$\text{tr } \mathcal{O} = \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} \mathcal{O}(z, \bar{z}), \quad (4.14)$$

a form consistent with the cyclic property of the trace, as one can verify by taking the trace of equation (4.13).

With the definition (4.12), the matrix elements of the Hamiltonian (4.6) and the operator $U_0(t)$ are, respectively,

$$\langle z | H_0 | \bar{z} \rangle = \hbar\omega z\bar{z} e^{z\bar{z}}, \quad \langle z | U_0(t) | \bar{z} \rangle = e^{z\bar{z} e^{-\omega t}}. \quad (4.15)$$

The partition function $\mathcal{Z}_0(\beta)$ corresponding to H_0 is the trace of $U_0(\hbar\beta)$. Using equations (4.14) and (4.15), one finds the expected result,

$$\mathcal{Z}_0(\beta) = \text{tr } U_0(\hbar\beta) = \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} e^{z\bar{z} e^{-\omega\hbar\beta}} = \frac{1}{1 - e^{-\hbar\omega\beta}}. \quad (4.16)$$

Some properties of kernels.

(i) From the property that z and $\partial/\partial z$ are Hermitian conjugate, follows that Hermitian conjugation of operators is represented by the formal complex conjugation of kernels:

$$O \mapsto O^\dagger \Leftrightarrow \mathcal{O}(z, \bar{z}) \mapsto \overline{\mathcal{O}}(z, \bar{z}).$$

Clearly, with this definition H_0 and $U_0(t)$ are Hermitian.

(ii) To an operator O that has the matrix elements O_{mn} in the harmonic oscillator basis, is associated the kernel $\sum_{m,n} O_{mn} (z^m / \sqrt{m!})(\bar{z}^n / \sqrt{n!})$.

(iii) A somewhat similar representation, in the case of phase-space variables, is the mixed position–momentum representation $\langle q | \mathcal{O} | p \rangle$, which is obtained by a Fourier transformation on one argument,

$$\langle q | \mathcal{O} | p \rangle = \int dq' e^{ipq'/\hbar} \langle q | \mathcal{O} | q' \rangle.$$

4.2 Holomorphic path integral

We now derive a path integral representation of the matrix elements of the solution of equation (2.4)

$$\hbar \frac{\partial U}{\partial t}(t, t') = -H(t)U(t, t'), \quad U(t', t') = \mathbf{1}, \quad (4.17)$$

based on the holomorphic formalism, first in the example of the harmonic oscillator, and then for more general Hamiltonians [16, 17, 2] (see also the contribution of Faddeev in Ref. [12]). We expect the path integral to be related to the form (3.21) by a simple change of variables of the form (4.2) (quantum operators being replaced by classical variables), but the boundary conditions and boundary terms require a specific analysis.

4.2.1 The harmonic oscillator

We first expand the exact expression (4.15), for small Euclidean time ε , as

$$\langle z | U_0(\varepsilon) | \bar{z} \rangle = \exp(z\bar{z}(1 - \omega\varepsilon) + O(\varepsilon^2)). \quad (4.18)$$

To calculate U_0 for a finite time interval $t'' - t'$, we then write it as

$$U_0(t'' - t') = \lim_{n \rightarrow \infty} U_0^n (\varepsilon = (t'' - t')/n) \quad (4.19)$$

and express the products in terms of kernels using the rule (4.13),

$$\langle z'' | U_0(t'', t') | \bar{z}' \rangle = \lim_{n \rightarrow \infty} \int \prod_{k=1}^{n-1} \frac{d\bar{z}_k dz_k}{2i\pi} \exp[-\mathcal{S}_\varepsilon(z, \bar{z})], \quad (4.20)$$

with

$$\mathcal{S}_\varepsilon(z, \bar{z}) = \left[- \sum_{k=1}^{n-1} \bar{z}_k (z_{k+1} - z_k) - \bar{z}_0 z_1 + \omega \varepsilon \sum_{k=0}^{n-1} \bar{z}_k z_{k+1} \right], \quad (4.21)$$

and the boundary conditions $\bar{z}_0 = \bar{z}'$, $z_n = z''$.

In the formal continuum limit $n \rightarrow \infty$, $\varepsilon \rightarrow 0$, the expression (4.20) leads to the path integral representation

$$\langle z'' | U_0(t'', t') | \bar{z}' \rangle = \int \left[\frac{d\bar{z}(t) dz(t)}{2i\pi} \right] \exp[-\mathcal{S}_0(z, \bar{z})], \quad (4.22)$$

with $(\dot{z}(t) \equiv dz/dt)$

$$\mathcal{S}_0(z, \bar{z}) = -\bar{z}(t') z(t') + \int_{t'}^{t''} dt \bar{z}(t) [-\dot{z}(t) + \omega z(t)], \quad (4.23)$$

and the boundary conditions $z(t'') = z''$, $\bar{z}(t') = \bar{z}'$. The path $\{z(t), \bar{z}(t)\}$ is a trajectory in phase space in the complex parametrization.

The symmetry of the action between initial and final times, which is not explicit, can be verified by an integration by parts of the term $\bar{z}\dot{z}$, but the validity of time integration within a path integral requires a specific prescription (and the boundary term becomes $-\bar{z}(t'') z(t'')$).

The partition function. To calculate the trace of U_0 , we return to the discretized form, because boundary terms are involved. Using equation (4.14), we obtain

$$\mathcal{Z}_0(\beta) = \text{tr } U_0(\hbar\beta/2, -\hbar\beta/2) = \lim_{n \rightarrow \infty} \int \prod_{k=1}^n \frac{dz_k d\bar{z}_k}{2i\pi} \exp [-\mathcal{S}_\varepsilon(\bar{z}, z)], \quad (4.24)$$

where \mathcal{S}_ε now has different boundary terms,

$$\mathcal{S}_\varepsilon(\bar{z}, z) = \sum_{k=1}^n [-\bar{z}_{k-1}(z_k - z_{k-1}) + \hbar\omega\varepsilon\bar{z}_{k-1}z_k], \quad (4.25)$$

with the identification $\bar{z}_n = \bar{z}_0$ and $z_n = z_0$.

In the continuum limit, we obtain the path integral representation of the partition function

$$\mathcal{Z}_0(\beta) = \int \left[\frac{dz(t) d\bar{z}(t)}{2i\pi} \right] \exp [-\mathcal{S}_0(\bar{z}, z)], \quad (4.26)$$

with

$$\mathcal{S}_0(\bar{z}, z) = \int_{-\beta/2}^{\beta/2} dt \bar{z}(t) [-\dot{z}(t) + \hbar\omega z(t)], \quad (4.27)$$

and the *periodic boundary conditions* $z(-\beta/2) = z(\beta/2)$, $\bar{z}(-\beta/2) = \bar{z}(\beta/2)$.

4.2.2 Linear coupling to an external source: Generating functional

The expression (4.26) generalizes simply to a system linearly coupled to external sources $\bar{b}(t)$ and $b(t)$ with the Hamiltonian

$$H = \hbar\omega z \partial/\partial z - b(t) \partial/\partial z - \bar{b}(t) z. \quad (4.28)$$

At order ε , the solution $U_G(t + \varepsilon, t)$ of equation (2.4) is $1 - \varepsilon H(t)$ and, therefore,

$$\begin{aligned} \langle z | U_G(t + \varepsilon, t) | \bar{z} \rangle &= e^{\bar{z}z} [1 - \varepsilon(\omega z \bar{z} - b(t)\bar{z} - \bar{b}(t)z)] + O(\varepsilon^2) \\ &= \exp [\bar{z}z - \varepsilon(\omega z \bar{z} - b(t)\bar{z} - \bar{b}(t)z) + O(\varepsilon^2)]. \end{aligned}$$

In the continuum limit, the action \mathcal{S}_0 in the path integral (4.22) is replaced by the action

$$\mathcal{S}_G(z, \bar{z}) = -\bar{z}(t')z(t') + \int_{t'}^{t''} dt \{ \bar{z}(t) [-\dot{z}(t) + \omega z(t)] - \bar{z}(t)b(t) - \bar{b}(t)z(t) \}. \quad (4.29)$$

The trace of U_G takes the form (equation (4.14)),

$$\begin{aligned} \text{tr } U_G(\hbar\beta/2, -\hbar\beta/2) &= \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} U_G(\hbar\beta/2, -\hbar\beta/2; z, \bar{z}) \\ &= \int \left[\frac{d\bar{z}(t) dz(t)}{2i\pi} \right] \exp [-\mathcal{S}_G(z, \bar{z})], \end{aligned} \quad (4.30)$$

with

$$\mathcal{S}_G(z, \bar{z}) = \int_{-\beta/2}^{\beta/2} dt \{ \bar{z}(t) [-\dot{z}(t) + \hbar\omega z(t)] - \bar{z}(t)b(t) - \bar{b}(t)z(t) \},$$

and the periodic boundary conditions $z(-\beta/2) = z(\beta/2)$, $\bar{z}(-\beta/2) = \bar{z}(\beta/2)$.

We leave the explicit calculation of the Gaussian path integral as an exercise, since an analogous calculation is presented in the fermion case. The result is

$$\mathrm{tr} U_G(\hbar\beta/2, -\hbar\beta/2) = \mathcal{Z}_0(\beta) \exp \left[\int_{\beta/2}^{\beta/2} dt du \bar{b}(u) \Delta(t-u) b(t) \right], \quad (4.31)$$

where $\mathcal{Z}_0(\beta)$ is the partition function (4.16) of the harmonic oscillator and the function

$$\Delta(t) = \frac{1}{2} e^{-\hbar\omega t} [\mathrm{sgn}(t) + 1/\tanh(\hbar\omega\beta/2)], \quad (4.32)$$

($\mathrm{sgn}(t) = 1$ for $t > 0$, $\mathrm{sgn}(-t) = -\mathrm{sgn}(t)$, $d\mathrm{sgn}(t)/dt = 2\delta(t)$) is the solution of the differential equation

$$\dot{\Delta}(t) + \hbar\omega\Delta(t) = \delta(t), \quad (4.33)$$

in the interval $[-\beta/2, \beta/2]$ with periodic boundary conditions.

In particular, the Gaussian two-point function with weight e^{-S_0} (equation (4.27)), which we call propagator with reference to real propagation, and which is the basic element of perturbation theory, is given by

$$\begin{aligned} \langle \bar{z}(t_2) z(t_1) \rangle_0 &= \mathcal{Z}_0^{-1}(\beta) \frac{\delta^2}{\delta b(t_2) \delta \bar{b}(t_1)} \exp \left[\int_{\beta/2}^{\beta/2} dt du \bar{b}(t) \Delta(u-t) b(u) \right] \Big|_{b=\bar{b} \equiv 0} \\ &= \Delta(t_2 - t_1). \end{aligned} \quad (4.34)$$

The Fourier representation of the propagator is

$$\Delta(t) = \frac{1}{2\pi} \int \frac{d\kappa e^{ikt}}{i\kappa + \hbar\omega}. \quad (4.35)$$

Partition function. The partition function $\mathcal{Z}_0(\beta)$ (equation (4.26)) can be related to the propagator by the method of Section 2.6.2. Differentiating the path integral (4.26) with respect to ω , one finds

$$\frac{\partial \ln \mathcal{Z}_0(\beta)}{\partial \omega} = -\hbar \int_{-\beta/2}^{\beta/2} dt \langle z(t) \bar{z}(t) \rangle_0 = -\hbar\beta\Delta(0).$$

A difficulty then arises, because $\Delta(t)$ is not continuous at $t = 0$; this, again, is the $\mathrm{sgn}(0)$ problem (a problem that can only be solved by a regularization of the path integral). Integrating, one obtains (using the limit $\beta \rightarrow \infty$ to determine the integration constant)

$$\mathcal{Z}_0(\beta) = \frac{e^{-\hbar\beta\omega \mathrm{sgn}(0)/2}}{2 \sinh(\hbar\beta\omega/2)}.$$

The ambiguity due to $\mathrm{sgn}(0)$ has the nature of a constant shift of the Hamiltonian, a natural consequence of the order problem between the operators z and $\partial/\partial z$. The rule consistent with the normal order (4.3) is to set $\mathrm{sgn}(0) = -1$ and yields

$$\mathcal{Z}_0(\beta) = \frac{1}{1 - e^{-\hbar\beta\omega}}. \quad (4.36)$$

However, such a choice is inconvenient for perturbative calculations, because it breaks time reversal symmetry, and it is better to choose $\mathrm{sgn}(0) = 0$. This corresponds to a symmetric form $\frac{1}{2}(z\partial/\partial z + \partial/\partial zz)$ and yields the standard harmonic oscillator. The additional contribution can then be cancelled, if necessary, by shifting physical parameters.

4.2.3 General one-dimensional Hamiltonian

Quite generally, one expresses the quantum Hamiltonian $H(p, q)$ in terms of creation and annihilation operators and commutes all creation operators to the left (normal order). In the Hamiltonian $h(a^\dagger, a)$, one replaces operators by the corresponding classical variables, as explained in Section 4.1. One obtains the matrix elements $\langle z|U(t + \varepsilon, t)|\bar{z}\rangle$ at order ε and, following the method of Section 4.2.1, derives the path integral representation

$$\langle z''|U(t'', t')|\bar{z}'\rangle = \int \left[\frac{d\bar{z}(t)dz(t)}{2i\pi} \right] \exp[-S(z, \bar{z})], \quad \text{with} \quad (4.37)$$

$$S(z(t), \bar{z}(t)) = -\bar{z}(t')z(t') + \int_{t'}^{t''} dt [-\bar{z}(t)\dot{z}(t) + h(z(t), \bar{z}(t))/\hbar], \quad (4.38)$$

and the boundary conditions $z(t'') = z'', \bar{z}(t') = \bar{z}'$.

Such a path integral can be used to generate a perturbative expansion. However, the expansion is plagued by singularities, reflections of the quantization problem and the order between the quantum operators z and $\partial/\partial z$. In particular, the undefined quantity $\Delta(0)$ appears. As we have already explained, normal order corresponds to the choice of $\text{sgn}(0) = -1$, which is somewhat inconvenient, and it is preferable to choose $\text{sgn}(0) = 0$ and to modify $h(z, \bar{z})$ to suppress the unwanted additional contributions.

Real parametrization of phase space. From the classical point of view, momentum and position variables (p, q) and complex variables (\bar{z}, z) are two different parametrizations of phase space related by

$$p - i\omega q = -i\sqrt{2\hbar\omega}\bar{z}, \quad p + i\omega q = i\sqrt{2\hbar\omega}z. \quad (4.39)$$

Path integrals, because they involve classical Hamiltonians, extend somewhat this correspondence to QM. If, in the path integral over phase space derived in Chapter 3 (equation (3.21)), one changes variables $(p(t), q(t)) \mapsto (\bar{z}(t), z(t))$, where the variables are related at all times by the relations (4.39), one finds a form (4.37, 4.38), with

$$h(z, \bar{z}) = H(i(z - \bar{z})\sqrt{\hbar\omega/2}, (z + \bar{z})\sqrt{\hbar/2\omega}).$$

The differences come only from boundary terms and boundary conditions, and this justifies our starting from first principles again. But once the modifications are known, one can infer the holomorphic path integral directly from the phase-space integral.

Both formalisms may have problems generated by ordering operators in products. However, the transformation (4.39) is not innocuous, since even simple Hamiltonians of the form $p^2 + V(q)$ have quantization problems after the transformation.

Partition function. The partition function is given by

$$\mathcal{Z}(\beta) = \int \frac{d\bar{z} dz}{2i\pi} e^{-z\bar{z}} \langle z|U(\hbar\beta/2, -\hbar\beta/2)|\bar{z}\rangle = \int \left[\frac{d\bar{z}(t)dz(t)}{2i\pi} \right] \exp[-S(z, \bar{z})],$$

with

$$S(z, \bar{z}) = \int_{-\beta/2}^{\beta/2} dt [-\bar{z}(t)\dot{z}(t) + h(z(t), \bar{z}(t))],$$

and the periodic boundary conditions $z(-\beta/2) = z(\beta/2), \bar{z}(-\beta/2) = \bar{z}(\beta/2)$.

It can be calculated by expanding around the harmonic oscillator, $h(z, \bar{z}) = \omega z\bar{z} + h_I(z, \bar{z})$, and evaluating perturbative terms by using Wick's theorem (1.30) together with the two-point function (4.32).

Remark. If the Hamiltonian has the form discussed in Chapter 2,

$$H = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{q}^2 + V_1(\hat{q}),$$

perturbation theory requires only the Gaussian two-point function of $q(t)$, which is proportional to $z(t) + \bar{z}(t)$. We see from the action (4.29) that this sum can be generated by taking a real source $b(t)$ ($\bar{b}(t) = b(t)$) and acting by functional differentiation with respect to $b(t)$ on U_G . Thus, the expressions (4.31) and (4.32) can be symmetrized in time. After the rescaling, $b(t) \mapsto b(t)\sqrt{\hbar/2\omega}$, one finds

$$\text{tr } U_G(\hbar\beta/2, -\hbar\beta/2) = \frac{1}{2\sinh(\omega\beta/2)} \exp \left[\frac{1}{2} \int dt_1 dt_2 b(t_1) \Delta(t_1 - t_2) b(t_2) \right],$$

with

$$\Delta(t) = \hbar \frac{\cosh \hbar\omega(\beta/2 - |t|)}{2\omega \sinh(\hbar\omega\beta/2)},$$

a result consistent with equation (2.57).

4.3 Several degrees of freedom. Boson interpretation

In the holomorphic formalism, a Hamiltonian H_0 , sum of ν independent harmonic oscillators with frequencies $\omega_i > 0$, can be expressed as the differential operator

$$H_0 = \sum_{i=1}^{\nu} \hbar\omega_i z_i \frac{\partial}{\partial z_i},$$

acting on analytic functions of ν complex variables, which can be expanded in the form

$$\psi(\mathbf{z}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1, i_2, \dots, i_n=1}^{\nu} \psi_{i_1 i_2 \dots i_n} z_{i_1} z_{i_2} \dots z_{i_n}, \quad (4.40)$$

where the coefficient $\psi_{i_1 i_2 \dots i_n}$ is symmetric in its n indices i_1, i_2, \dots, i_n . The monomial $z_{i_1} z_{i_2} \dots z_{i_n}$ is an eigenstate of H_0 with the energy $\hbar(\omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_n})$.

Many-body boson interpretation. Due to the additive character of the spectrum of the harmonic oscillator, the energy eigenvalues also have the interpretation of the total energy of non-interacting particles. In this alternative interpretation, one-particle states are associated with the energies $\hbar\omega_i$. Moreover, $\psi_{i_1 i_2 \dots i_n}$ is then the component of $\psi(\mathbf{z})$ on an n -particle state and, since $\psi_{i_1 i_2 \dots i_n}$ is symmetric in its n indices, these particles obey the Bose–Einstein statistics: the holomorphic formalism makes a description of general boson states, which are linear combinations of states with an arbitrary number of particles, possible.

We choose the functions $\Psi(\mathbf{z})$ to be normalizable with respect to the scalar product

$$(\psi_1, \psi_2) = \int \left(\prod_{i=1}^{\nu} \frac{dz_i d\bar{z}_i}{2i\pi} e^{-z_i \bar{z}_i} \right) \overline{\psi_1(\mathbf{z})} \psi_2(\mathbf{z}), \quad (4.41)$$

which generalizes the scalar product (4.8). Then,

$$\|\Psi\|^2 = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1, i_2, \dots, i_n} |\psi_{i_1 i_2 \dots i_n}|^2 < \infty.$$

The normalizable functions (4.40) belong to a Hilbert space of entire functions of ν variables. When $\|\Psi\| = 1$, the n th term in the sum gives the probability for the quantum system to be in an n -particle state, the factor $1/n!$ cancelling the over-counting of states implied by the unrestricted summation over all indices $\{i_k\}$.

The kernel of the identity is now

$$\mathcal{I}(z, \bar{z}) = \exp(\sum_{i=1}^{\nu} \bar{z}_i z_i). \quad (4.42)$$

The kernels of the Hamiltonian and the number operator N follow:

$$\langle z | H_0 | \bar{z} \rangle = \mathcal{I}(z, \bar{z}) \sum_i \hbar \omega_i z_i \bar{z}_i, \quad \langle z | N | \bar{z} \rangle = \mathcal{I}(z, \bar{z}) \sum_i z_i \bar{z}_i.$$

The partition function for free bosons that can occupy ν different states of energy $\hbar \omega_i$ is then given by the path integral

$$\mathcal{Z}_0(\beta) = \int \left[\frac{dz(t)d\bar{z}(t)}{2i\pi} \right] \exp[-\mathcal{S}_0(\bar{z}, z)], \quad (4.43)$$

$$\mathcal{S}_0(\bar{z}, z) = \int_{-\beta/2}^{\beta/2} dt \sum_i \bar{z}_i(t) [-\dot{z}(t) + \hbar \omega_i z_i(t)]. \quad (4.44)$$

By adding to the action \mathcal{S}_0 , a polynomial of higher degree in the variables z_i and \bar{z}_i , it is possible to describe interactions between bosons.

This formalism is well-suited to the study of quantum statistical systems of bosons in the grand canonical formulation, as discussed in Section 4.4.2.

4.4 The Bose gas. Field integral representation

We now consider a non-relativistic gas of bosons, that is, identical quantum particles obeying the Bose–Einstein statistics of mass m , in continuum space. For simplicity, we assume that the particles have no internal degrees of freedom, like spin.

A formalism with a fixed number of particles is useful for models where the n -body problem can be solved exactly, or approximately, but more cumbersome in general.

The grand canonical formalism, where the number of particles can vary, and is fixed only on average, is better adapted in many situations where collective effects are important, like in the theory of phase transitions.

4.4.1 Matrix density at thermal equilibrium: Fixed number of particles

The density matrix at thermal equilibrium for a fixed number of bosons, can be calculated from a *path integral*, simple extension of the form derived in Chapter 2, but where the boson character is taken into account.

We consider an n -particle Hamiltonian in d dimensions of the general form

$$H_n = T_n + \mathcal{V}_n(\mathbf{x}_1, \dots, \mathbf{x}_n), \quad \mathbf{x}_i \in \mathbb{R}^d, \quad (4.45)$$

where T_n is the kinetic term,

$$T_n = -\frac{\hbar^2}{2m} \sum_{i=1}^n \nabla_i^2, \quad (4.46)$$

with $\nabla_i \equiv \nabla_{x_i} \equiv (\partial/\partial x_i^1, \partial/\partial x_i^2, \dots, \partial/\partial x_i^d)$, and where the potential $\mathcal{V}_n(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is a symmetric function of the n positions. The Hamiltonian acts on square integrable symmetric wave functions of n positions, $\psi_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$, which belong to the Hilbert space \mathfrak{H}_n of n -particle boson states.

The partition function is the trace of the statistical operator $e^{-\beta H}$, but, in the trace, the statistical properties of the quantum particles have to be taken into account. This implies that the trace in the partition function has to be restricted to the relevant subspace by the insertion of the corresponding projectors. The partition function can be expressed as a path integral of the form,

$$\mathcal{Z}(n, \beta) = \int [d^d x_i(t)] e^{-S(x)/\hbar}, \quad (4.47)$$

with

$$S(x)/\hbar = \int_0^\beta dt \left[\sum_i \frac{1}{2m\hbar^2} (\dot{x}_i(t))^2 + V_n(\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)) \right], \quad (4.48)$$

but averaged over the $n!$ different boundary conditions corresponding to all permutations P , $\mathbf{x}_i(\beta) = \mathbf{x}_{P(i)}(0)$.

4.4.2 Second quantization. Field integral representation

We now discuss the thermodynamics of a gas of quantum particles in the grand canonical formalism in which the number of particles is fixed only on average by tuning a parameter μ coupled to the number of particles and called *chemical potential*. Generalizing the formalism of Section 4.3, we use a representation of bosons often referred to as *second quantization*. We derive an expression for the partition function that has the form of a *field integral*, a generalization of the holomorphic path integral (4.44) of Section 4.3.

The n-particle potential. We consider here only potentials that are sums of one-body and two-body potentials (to simplify the notation, we now substitute $\mathbf{x} \mapsto x$):

$$\mathcal{V}_n(x_1, x_2, \dots, x_n) = \sum_{i=1}^n V_1(x_i) + \sum_{i < j \leq n} V_2(x_i, x_j), \quad (4.49)$$

($V_2(x, y) = V_2(y, x)$) but the generalization to other many-body potentials is simple.

4.4.3 Fock space

We consider the direct sum $\bigoplus \mathfrak{H}_n$, $n = 1, \dots, \infty$ of Hilbert spaces. Since the wave functions are symmetric, we introduce a complex field $\varphi(x)$, and a generating functional of n -particle wave functions (see Section 2.5.3)

$$\Psi(\varphi) = \sum_{n=0} \frac{1}{n!} \left(\int \prod_i d^d x_i \varphi(x_i) \right) \psi_n(x_1, \dots, x_n), \quad (4.50)$$

which generalizes expression (4.40). The scalar product of two functionals is given by the *field integral* (see Section 4.4.5)

$$(\Psi_1, \Psi_2) = \left\langle \overline{\Psi_1(\varphi)} \Psi_2(\varphi) \right\rangle \equiv \int [d\varphi d\bar{\varphi}] \overline{\Psi_1(\varphi)} \Psi_2(\varphi) \exp \left[- \int d^d x \bar{\varphi}(x) \varphi(x) \right], \quad (4.51)$$

which generalizes expression (4.41), and which is normalized by the condition

$$(1, 1) = \langle 1 \rangle \equiv \int [d\varphi d\bar{\varphi}] \exp \left[- \int d^d x \bar{\varphi}(x) \varphi(x) \right] = 1.$$

Since the measure is Gaussian, to calculate expectation values one only needs the two-point function. The kernel in the quadratic form is the identity. Its inverse is also the identity and, therefore,

$$\langle \bar{\varphi}(x)\varphi(y) \rangle = \delta^{(d)}(x - y).$$

Using then Wick's theorem, it is simple to verify that the norm of the functional Ψ is then given by

$$|\Psi|^2 = (\Psi, \Psi) = \sum_{n=0} \frac{1}{n!} \left(\int \prod_i d^d x_i \right) |\psi_n(x_1, \dots, x_n)|^2.$$

When $\|\Psi\| = 1$, the n th term in the sum is the probability that the boson system is in an n -particle state. The complex vector space of functionals with finite norm is called a Fock space, and this formalism is often called second quantization.

4.4.4 Hamiltonian in Fock space

We denote by \mathbf{H} the Hamiltonian in Fock space, whose restriction to n -particle states is H_n , \mathbf{T} the kinetic term whose restriction is T_n , and \mathbf{V}_1 , \mathbf{V}_2 the one- and two-body potentials, respectively. Thus,

$$\mathbf{H} = \mathbf{T} + \mathbf{V}_1 + \mathbf{V}_2. \quad (4.52)$$

We then represent the kinetic term \mathbf{T} and the potential terms \mathbf{V}_1 and \mathbf{V}_2 by operators acting on Ψ by functional differentiation.

We first calculate

$$\begin{aligned} \int d^d x \varphi(x) \nabla_x^2 \frac{\delta}{\delta \varphi(x)} \Psi(\varphi) &= \int d^d x \varphi(x) \nabla_x^2 \sum_n \frac{1}{(n-1)!} \\ &\times \int \left(\prod_{i<n} d^d x_i \varphi(x_i) \right) \psi_n(x_1, \dots, x_{n-1}, x). \end{aligned}$$

The variable x in the right-hand side can be renamed x_n . A symmetrization of the coefficient of $\prod_{i \leq n} \varphi(x_i)$ then yields a factor $1/n$ and the sum of all gradients squared. We have reconstructed, up to a factor, the kinetic term (4.46). We conclude that

$$\mathbf{T}\Psi(\varphi) = -\frac{\hbar^2}{2m} \int d^d x \varphi(x) \nabla_x^2 \frac{\delta}{\delta \varphi(x)} \Psi(\varphi). \quad (4.53)$$

For the one-body potential \mathbf{V}_1 , an analogous argument shows that

$$\mathbf{V}_1\Psi(\varphi) = \int d^d x \varphi(x) V_1(x) \frac{\delta}{\delta \varphi(x)} \Psi(\varphi).$$

To generate a two-body potential, it is necessary to differentiate twice with respect to φ at two different points. It is simple to verify the representation

$$\mathbf{V}_2\Psi(\varphi) = \frac{1}{2} \int d^d x d^d y \varphi(x) \varphi(y) V_2(x, y) \frac{\delta^2}{\delta \varphi(x) \delta \varphi(y)} \Psi(\varphi). \quad (4.54)$$

We now have a representation of the full Hamiltonian \mathbf{H} acting on generating functionals.

The representation of the operator number of particles \mathbf{N} is given by

$$\mathbf{N} = \int d^d x \varphi(x) \frac{\delta}{\delta \varphi(x)} \Rightarrow [\mathbf{N}, \mathbf{H}] = 0. \quad (4.55)$$

In what follows, we consider as full Hamiltonian the sum $\mathbf{H} - \mu\mathbf{N}$, where the chemical potential μ determines the average value of \mathbf{N} .

Because, in this construction, the role of coordinates is played by fields $\varphi(x)$, we are constructing a non-relativistic (in general non-local) *quantum field theory (QFT)*.

4.4.5 Kernels of operators and field integral representation

Within the holomorphic formalism, we now use functional methods to derive a representation of matrix elements of the statistical operator as *field integrals*, generalization of the path integrals discussed in Section 4.2.1: in the field integral, the summation over fields replaces the summation over paths. The form of the field integral actually follows rather directly from results already obtained in QM, provided one interprets the function $\varphi(x)$ as a set of complex variables depending on a continuous index x and $\Psi(\varphi)$ as an analytic ‘functional’ in the holomorphic formalism of Section 4.1. Note that the matrix elements of the statistical operator depend on two fields and represent an operator acting on the space of generating functionals.

We denote by $\bar{\varphi}(x)$ the field conjugate to $\varphi(x)$. The kernel of the identity operator $\mathcal{I}(\varphi, \bar{\varphi})$, generalization of expression (4.42) and consistent with the definition (4.51) of the scalar product, takes the form

$$\mathcal{I}(\varphi, \bar{\varphi}) = \exp \left[\int d^d x \varphi(x) \bar{\varphi}(x) \right]. \quad (4.56)$$

The kernel associated with operators defined in terms of the field $\varphi(x)$ and functional derivatives is then obtained by acting on $\mathcal{I}(\varphi, \bar{\varphi})$. Note that the Hamiltonian here is directly written in normal order. Its matrix elements are

$$\begin{aligned} \langle \varphi | \mathbf{H} | \bar{\varphi} \rangle &= \mathcal{I}(\varphi, \bar{\varphi}) \left\{ \int d^d x \varphi(x) \left[-\frac{\hbar^2}{2m} \nabla^2 + V_1(x) \right] \bar{\varphi}(x) \right. \\ &\quad \left. + \frac{1}{2} \int d^d x d^d y \varphi(x) \varphi(y) V_2(x, y) \bar{\varphi}(x) \bar{\varphi}(y) \right\}. \end{aligned} \quad (4.57)$$

The kernel associated with the particle number operator is

$$\langle \varphi | \mathbf{N} | \bar{\varphi} \rangle = \mathcal{I}(\varphi, \bar{\varphi}) \int d^d x \varphi(x) \bar{\varphi}(x). \quad (4.58)$$

Considering the space coordinate x as a continuous index, we can adapt the expressions of Section 4.2.1, in particular equation (4.38), to this more general situation, in much the same way as we have generalized the simple integrals of Chapter 1 to path integrals. We then obtain a field integral, because we no longer integrate over paths, but instead, over *fields* $\{\varphi(t, x), \bar{\varphi}(t, x)\}$. Here, one finds

$$\langle \varphi'' | \mathbf{U}(t'', t') | \bar{\varphi}' \rangle = \langle \varphi'' | e^{-(t''-t)(\mathbf{H}-\mu \mathbf{N})/\hbar} | \bar{\varphi}' \rangle = \int [d\bar{\varphi}(t, x) d\varphi(t, x)] e^{-S(\varphi, \bar{\varphi})}, \quad (4.59)$$

with the boundary conditions $\bar{\varphi}(t, x') \equiv \bar{\varphi}'(x)$, $\varphi(t, x'') \equiv \varphi''(x)$, and the Euclidean action

$$\begin{aligned} S(\varphi, \bar{\varphi}) &= -\bar{\varphi}(t, x') \varphi(t, x') + \int dt d^d x \bar{\varphi}(t, x) \left(-\frac{\partial}{\partial t} - \frac{\hbar}{2m} \nabla_x^2 + \frac{V_1(x) - \mu}{\hbar} \right) \varphi(t, x) \\ &\quad + \frac{1}{2\hbar} \int dt d^d x d^d y \bar{\varphi}(t, x) \varphi(t, x) V_2(x, y) \bar{\varphi}(t, y) \varphi(t, y). \end{aligned} \quad (4.60)$$

Setting $t = \hbar\beta$, one obtains the statistical operator and then the partition function,

$$\mathcal{Z}(\beta) = \int [d\bar{\varphi}(t, x)d\varphi(t, x)] \exp[-\mathcal{S}(\varphi, \bar{\varphi})], \quad (4.61)$$

with the periodic boundary conditions

$$\varphi(-\beta/2, x) = \bar{\varphi}(\beta/2, x), \quad \varphi(-\beta/2, x) = \varphi(\beta/2, x),$$

and the Euclidean action is

$$\begin{aligned} \mathcal{S}(\varphi, \bar{\varphi}) &= \int dt d^d x \bar{\varphi}(t, x) \left(-\frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla_x^2 + V_1(x) - \mu \right) \varphi(t, x) \\ &\quad + \frac{1}{2} \int dt d^d x d^d y \bar{\varphi}(t, x) \varphi(t, x) V_2(x, y) \bar{\varphi}(t, y) \varphi(t, y). \end{aligned} \quad (4.62)$$

We have derived a representation of the partition function in the form of a field integral for a *non-relativistic QFT* (see, for example, Ref. [19]). The generalization to a relativistic theory is then mainly a matter of implementing relativistic invariance.

U(1) symmetry. The particle number conservation has an interesting consequence for the field integral: it has a $U(1)$ symmetry, corresponding to the transformations

$$\varphi(t, x) \mapsto e^{i\kappa} \varphi(t, x), \quad \bar{\varphi}(t, x) \mapsto e^{-i\kappa} \bar{\varphi}(t, x),$$

with κ any real constant. We show in Section 15.10 that interesting physics is associated with the spontaneous breaking of the symmetry.

4.4.6 The Gaussian model

As an illustration, we now calculate the partition function of the Gaussian model, which corresponds to a gas of independent particles in a one-body external potential V_1 . We begin with the free gas. We want to show that well-known results are recovered. Although functional methods are not required for these simple examples, their study will help us to take into account the effect of additional interactions.

Free Bose gas. In the absence of a potential, the partition function can be calculated exactly. To take advantage of translation invariance, we expand the fields $\varphi, \bar{\varphi}$ in Fourier modes. In d -space dimensions,

$$\bar{\varphi}(t, \mathbf{x}) = \int d^d p e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \tilde{\varphi}^*(t, \mathbf{p}), \quad \varphi(t, \mathbf{x}) = \int d^d p e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \tilde{\varphi}(t, \mathbf{p}). \quad (4.63)$$

The Jacobian is trivial, and the action becomes

$$\mathcal{S}_0(\tilde{\varphi}, \tilde{\varphi}^*) = (2\pi\hbar)^d \int dt d^d p \tilde{\varphi}^*(t, \mathbf{p}) \left(-\frac{\partial}{\partial t} + \frac{\mathbf{p}^2}{2m} - \mu \right) \tilde{\varphi}(t, \mathbf{p}). \quad (4.64)$$

We first calculate the partition function in a hypercubic box of linear size L with periodic boundary conditions and then take the thermodynamic limit. The fields $\varphi(t, \mathbf{x}), \bar{\varphi}(t, \mathbf{x})$ are periodic functions of period L of all space coordinates. After Fourier transformation, the corresponding momenta belong to the lattice

$$\mathbf{p} = 2\pi\hbar\mathbf{n}/L, \quad \mathbf{n} \in \mathbb{Z}^d,$$

and the action reads

$$\mathcal{S}_0(\tilde{\varphi}^*, \tilde{\varphi}) = (2\pi\hbar)^{2d} \int dt \sum_{\mathbf{n}} \tilde{\varphi}^*(t, \mathbf{p}) \left(-\frac{\partial}{\partial t} + \frac{\mathbf{p}^2}{2m} - \mu \right) \tilde{\varphi}(t, \mathbf{p}). \quad (4.65)$$

The field integral factorizes and the result is the product over all values of \mathbf{n} of a partition function of the form (4.36). The free energy, $\mathcal{W}_0(\beta) = \beta^{-1} \ln \mathcal{Z}_0(\beta)$, is

$$\mathcal{W}_0(\beta) = -\frac{1}{\beta} \sum_{\mathbf{n} \in \mathbb{Z}^d} \ln \left(1 - e^{-\beta(\mathbf{p}^2/2m - \mu)} \right). \quad (4.66)$$

In the thermodynamic limit $L \rightarrow \infty$, the free energy per unit volume, which is the pressure Π , becomes ($d\mathbf{n} = d\mathbf{p}L/2\pi\hbar$)

$$\Pi = L^{-d} \mathcal{W}_0(\beta) = -\frac{1}{\beta} \int \frac{d^d p}{(2\pi\hbar)^d} \ln \left(1 - e^{-\beta(\mathbf{p}^2/2m - \mu)} \right). \quad (4.67)$$

Note that the Bose gas is stable only if the chemical potential is non-positive.

The derivative of $\ln \mathcal{Z}$ with respect to β , at $\beta\mu$ fixed, yields the average energy density

$$\langle H \rangle = -\frac{1}{L^d} \left. \frac{\partial \ln \mathcal{Z}}{\partial \beta} \right|_{\beta\mu \text{ fixed}} = \frac{1}{(2\pi\hbar)^d} \int \frac{d^d p \mathbf{p}^2/2m}{e^{\beta(\mathbf{p}^2/2m - \mu)} - 1}. \quad (4.68)$$

The derivative of $\ln \mathcal{Z}$ with respect to $\beta\mu$ (β fixed) yields the average particle number, and thus the gas density

$$\rho = L^{-d} \langle N \rangle = \frac{1}{L^d \beta} \int dt d^d x \langle \varphi(t, x) \bar{\varphi}(t, x) \rangle = \frac{1}{(2\pi\hbar)^d} \int \frac{d^d p}{e^{\beta(\mathbf{p}^2/2m - \mu)} - 1}. \quad (4.69)$$

This equation, called *equation of state*, is known to exhibit the phenomenon of Bose–Einstein condensation. At fixed temperature $T = 1/\beta$, when μ increases the density ρ increases. When the dimension d of space is larger than 2, since μ cannot become positive, ρ is bounded by the value ρ_c of the integral calculated for $\mu = 0$:

$$\rho \leq \rho_c = \frac{1}{(2\pi\hbar)^d} \int \frac{d^d p}{e^{\beta\mathbf{p}^2/2m} - 1} = \zeta(d/2) \left(\frac{mT}{2\pi\hbar^2} \right)^{d/2},$$

where $\zeta(s)$ is Riemann's ζ -function. Conversely, at fixed density, the equation of state has a solution up to a minimal temperature T_0 given by

$$T_0 = \frac{2\pi\hbar^2}{m} \left(\frac{\rho}{\zeta(d/2)} \right)^{2/d}.$$

To understand the physics at lower temperatures $T < T_0$, it is useful to return to a large but finite box, where the momentum modes are discrete. One then discovers that a macroscopic fraction of the free Bose gas condenses in the ground state, which corresponds to the zero-momentum mode. This is the phenomenon of *Bose–Einstein condensation*.

In two dimensions, because ρ_c diverges, there is no condensation.

General one-body potential. To calculate the Gaussian field integral, we replace the Fourier expansion (4.63) by an expansion on the eigenfunctions of the one-body Hamiltonian H_1 . The same arguments then lead to

$$\mathcal{W}(\beta) = -\text{tr} \ln \left(1 - e^{\beta\mu - \beta H_1} \right).$$

Similarly,

$$\langle N \rangle = \text{tr} [1 - e^{\beta\mu - \beta H_1}]^{-1}, \quad \langle H \rangle = \text{tr} H_1 [1 - e^{\beta\mu - \beta H_1}]^{-1}.$$

An analysis, based on the arguments of Section 2.8, shows that, in the semi-classical limit, the free energy becomes

$$\mathcal{W}(\beta) = -\frac{1}{\beta} \int \frac{d^d p d^d x}{(2\pi\hbar)^d} \ln \left(1 - e^{\beta\mu - \beta H_1(p, x)} \right),$$

where $H_1(p, x)$ is the classical Hamiltonian.

A simple example is provided by a harmonic well, of the kind relevant for the magnetic traps of some Bose–Einstein condensation experiments. For simplicity, we assume that the trap is spherical. Then,

$$H_1 = \frac{1}{2m} \mathbf{p}^2 + \frac{1}{2} m\omega^2 \mathbf{x}^2.$$

In the semi-classical limit, the average number of particles in the trap, for $\mu = 0$, is

$$\langle N \rangle = \frac{\zeta(d)}{(\hbar\beta\omega)^d},$$

an approximation valid for temperatures large enough, $T \gg \hbar\omega$. Condensation now occurs for any dimension $d > 1$ below the temperature

$$T_0 = 1/\beta = \hbar\omega(\langle N \rangle)^{1/d}.$$

At lower temperatures, the discrete nature of the quantum spectrum becomes relevant, the relation between particle numbers and chemical potential can be satisfied up to $\mu = d\hbar\omega/2$, and below T_0 particles accumulate in the ground state.

4.4.7 Pair potentials: The example of the $\delta(x)$ -function potential

When one is interested only in long-wavelength phenomena, one can often approximate a short-range pair-potential by a δ -function pseudo potential. The action then simplifies and becomes *local* (here, we assume $V_1 \equiv 0$):

$$\mathcal{S}(\varphi, \bar{\varphi}) = \int dt d^d x \left[\bar{\varphi}(t, x) \left(-\frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla^2 - \mu \right) \varphi(t, x) + \frac{g}{2} (\bar{\varphi}(t, x) \varphi(t, x))^2 \right]. \quad (4.70)$$

For simple potentials, the strength g of the interaction is proportional to $\hbar^2 a^{d-2}/m$, where a is the s-wave scattering length. a must be positive, that is, correspond to a repulsive interaction, for the boson system to be stable.

Note that μ is no longer restricted to be negative. For $\mu > 0$, the minimum of the $\varphi, \bar{\varphi}$ potential now corresponds to a non-vanishing value of $|\varphi|$. This suggests the possibility of a phase transition.

In one space dimension, the model is exactly integrable and solvable by Bethe's Ansatz, in the sense that all eigenstates of the Hamiltonian are linear combinations of a finite number of plane waves. If thermal fluctuations are neglected, the model leads to a classical field equation, called non-linear Schrödinger equation, which is also integrable.

In three dimensions, the model is especially interesting: in the presence of the interaction, the condensation temperature of the Bose gas becomes the transition temperature for a phase transition where the $U(1)$ symmetry is spontaneously broken. Therefore, it describes the physics of the helium superfluid transition (see Section 15.10).

However, note that a Hamiltonian with a δ -function potential is only defined in one dimension. In higher dimensions, it leads to divergences that have to be dealt with (regularized, see Chapter 8) by adding an artificial short-distance structure. In two dimensions, the model is renormalizable, as power counting shows, and not renormalizable in higher dimensions (see Chapter 8).

In all dimensions $d > 0$, a problem of quantization also appears. Since

$$[\delta/\delta\varphi(x), \varphi(x)] = \delta^{(d)}(0)',$$

it leads to divergences. A careful derivation shows that normal order eliminates all divergences in one dimension.

Perturbative expansion. As an exercise, we calculate the correction of order g to the free energy. We need the Gaussian two-point $\langle\varphi\bar{\varphi}\rangle$ correlation function. In partial Fourier representation, from equations (4.64), (4.32), and (4.34), we infer ($\vartheta(t) = (1 + \text{sgn}(t))/2$),

$$\tilde{\Delta}(t, p) = \langle \tilde{\varphi}^*(t, p) \tilde{\varphi}(0, p) \rangle = \frac{e^{-\omega(p)t}}{(2\pi\hbar)^d} \left(\vartheta(t) + \frac{1}{e^{\beta\omega(p)} - 1} \right), \text{ with } \omega(p) = \frac{p^2}{2m} - \mu.$$

We recognize the sum of the two contributions generated by quantum and thermal fluctuations. After Fourier transformation over time, one finds (equation (4.35))

$$\tilde{\Delta}(\omega, p) = \frac{1}{(2\pi\hbar)^{d+1}} \frac{1}{i\omega + p^2/2m - \mu}. \quad (4.71)$$

Using Wick's theorem, one obtains

$$\mathcal{W}(\beta) = \mathcal{W}_0(\beta) - \frac{g}{2\beta} \int dt d^d x (\langle \varphi(t, x) \bar{\varphi}(t, x) \rangle)^2 + O(g^2).$$

Translation invariance implies that the two-point function at coinciding points is independent of t, x and thus the integral over space and Euclidean time generates a factor $L^d \beta$. Then one faces the $\text{sgn}(0)$ problem. The convention dictated by normal order is

$$\Delta(0, \mathbf{p}) = \frac{1}{e^{\beta(p^2/2m-\mu)} - 1},$$

and the result is then simply the square of the leading order density:

$$\mathcal{W}(\beta) = \mathcal{W}_0(\beta) - g L^d \left[\frac{1}{(2\pi\hbar)^d} \int \frac{d^d p}{e^{\beta(p^2/2m-\mu)} - 1} \right]^2 + O(g^2).$$

High temperature. The periodic boundary conditions in the Euclidean time direction imply that the field can be expanded on a basis of periodic functions of period β ,

$$\varphi(t, x) = \sum_{\nu \in \mathbb{Z}} e^{2i\pi\nu t/\beta} \varphi_\nu(x), \quad \bar{\varphi}(t, x) = \sum_{\nu \in \mathbb{Z}} e^{2i\pi\nu t/\beta} \bar{\varphi}_\nu(x).$$

At high temperature, $\beta \rightarrow 0$, in the action the time derivative suppresses the contribution of all non-zero modes. In this limit, the field can be approximated by the zero mode. Note that high temperature is defined with respect to, for instance, the correlation length ξ that characterizes the decay of correlation functions in space directions.

In terms of the thermal wavelength λ_{th} , the condition, which can be written as

$$\lambda_{\text{th}} = \hbar \sqrt{\beta/m} = \sqrt{\hbar^2/mT} \ll \xi,$$

is satisfied at high temperature for finite correlation length, or at finite temperature when the correlation length diverges, that is, near a second-order phase transition.

The action (4.70) can then be approximated as

$$\mathcal{S}(\varphi, \bar{\varphi}) = \beta \int d^d x \left[\bar{\varphi}_0(x) \left(-\hbar^2 \nabla^2 / 2m - \mu \right) \varphi_0(x) + \frac{1}{2} g (\bar{\varphi}_0(x) \varphi_0(x))^2 \right], \quad (4.72)$$

an action that we study systematically in the context of Euclidean QFT, and which is relevant to the superfluid phase transition (see Section 15.10).

4.5 Fermion representation and complex Grassmann algebras

To construct a formalism for fermions parallel to the one describing bosons as in Section 4.3, we now introduce Grassmann algebras. Grassmann algebras together with the rules of differentiation and integration have been described in Sections 1.5–1.7. We assume that the fermions can occupy a finite number ν of states and their number is not fixed.

4.5.1 Analytic Grassmann functions, scalar product

We consider a Grassmann algebra \mathfrak{C} with two sets of generators $\{\theta_i, \bar{\theta}_i\}$, $i = 1, \dots, \nu$. We generalize the considerations of Section 4.3 to fermions. ‘Analytic’ Grassmann functions are elements of the subalgebra \mathfrak{A} of functions of θ_i variables only:

$$\frac{\partial \psi}{\partial \bar{\theta}_i} = 0, \quad \forall i.$$

Analytic Grassmann functions form a complex vector space of dimension 2^ν . This is in direct relation with Fermi–Dirac statistics for a system with ν states: each state can either be empty or occupied by one fermion. Analytic functions have an expansion of the form

$$\psi(\theta) = \sum_{n=0}^{\nu} \frac{1}{n!} \sum_{i_1, i_2, \dots, i_n} \psi_{i_1 i_2 \dots i_n} \theta_{i_1} \theta_{i_2} \dots \theta_{i_n}.$$

Since the product of Grassmann generators is antisymmetric, the complex coefficients $\psi_{i_1 i_2 \dots i_n}$, amplitude of the vector on an n -particle state, can be chosen antisymmetric in all indices and this enforces the Pauli principle for fermions.

Scalar product of Grassmann analytic functions. In the same way as for analytic complex functions (expression (4.41)), a scalar product between two analytic Grassmann functions ψ_1 and ψ_2 can be defined by (two different forms are possible, see *e.g.*, Ref. [18]),

$$(\psi_1, \psi_2) = \int \left(\prod_i d\theta_i d\bar{\theta}_i \right) \exp \left(\sum_i \bar{\theta}_i \theta_i \right) \psi_1^\dagger(\theta) \psi_2(\theta). \quad (4.73)$$

With this definition, the norm of a vector ψ is positive and given by

$$\|\psi\|^2 = (\psi, \psi) = \sum_{n=0}^{\nu} \frac{1}{n!} \sum_{i_1, i_2, \dots, i_n} |\psi_{i_1 i_2 \dots i_n}|^2.$$

If the function ψ has unit norm, the quantity $\frac{1}{n!} \sum_{i_1, i_2, \dots, i_n} |\psi_{i_1 i_2 \dots i_n}|^2$ is the probability to find the quantum system in an n -fermion state.

The ' δ '-function. In Grassmann algebras, the role of the Dirac δ -function is played by the function θ itself, since

$$\int d\theta \theta f(\theta) = f(0),$$

in which $f(0)$ means the constant part of the affine function $f(\theta)$. The $\delta(\theta)$ -function has the convenient integral representation, analogous to a Fourier transform,

$$f(0) = \int d\theta d\bar{\theta} e^{\bar{\theta}\theta} f(\theta). \quad (4.74)$$

4.5.2 Operator algebra and kernels

Operator algebra. We have already introduced in Section 1.6.1 the algebra of operators $\{\theta_i, \partial/\partial\theta_i\}$ acting by left multiplication and differentiation on the algebra \mathfrak{A} of Grassmann functions. They satisfy the commutation relations of the usual fermion annihilation and creation operators (see relations (1.48)) and are Hermitian conjugate with respect to the scalar product (4.73).

For example, a Hamiltonian H_0 for non-interacting fermions that can occupy ν states of energies $\hbar\omega_i$ is represented by

$$H_0 = \sum_{i=1}^{\nu} \hbar\omega_i \theta_i \frac{\partial}{\partial\theta_i}, \quad (4.75)$$

as one can verify by explicit calculation,

$$H_0 \sum_{i_1, i_2, \dots, i_n} \psi_{i_1 i_2 \dots i_n} \theta_{i_1} \theta_{i_2} \dots \theta_{i_n} = \sum_{i_1, i_2, \dots, i_n} \hbar(\omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_n}) \psi_{i_1 i_2 \dots i_n} \theta_{i_1} \theta_{i_2} \dots \theta_{i_n}.$$

The operator number of particles N , which commutes with the Hamiltonian when the number of fermions is conserved, is obtained from H_0 by replacing all $\hbar\omega_i$ by 1:

$$N = \sum_{i=1}^{\nu} \theta_i \frac{\partial}{\partial\theta_i}. \quad (4.76)$$

The identity as a kernel. A representation of the identity $\mathcal{I}(\theta, \bar{\theta})$ as a kernel (element of \mathfrak{C}) follows from the existence of an orthonormal basis:

$$\mathcal{I}(\theta, \bar{\theta}) = \prod_i (1 + \theta_i \bar{\theta}_i) = \exp\left(-\sum_i \bar{\theta}_i \theta_i\right). \quad (4.77)$$

A direct verification relies on the identity (4.74),

$$\begin{aligned} \int \prod_i d\theta'_i d\bar{\theta}'_i \mathcal{I}(\theta, \bar{\theta}') \exp\left(\sum_i \bar{\theta}'_i \theta'_i\right) f(\theta') &= \int \prod_i d\theta'_i d\bar{\theta}'_i \exp\left(\sum_i \bar{\theta}'_i (\theta'_i - \theta_i)\right) f(\theta') \\ &= f(\theta). \end{aligned} \quad (4.78)$$

Operator algebra and kernels. Using the anticommutation relations, we commute all multiplication operators to the left of all differentiation operators, writing all operators in *normal order*. We then use equation (4.78) to associate to each element of this operator algebra a kernel,

$$\theta_{i_1} \theta_{i_2} \dots \theta_{i_p} \frac{\partial}{\partial \theta_{j_1}} \frac{\partial}{\partial \theta_{j_2}} \dots \frac{\partial}{\partial \theta_{j_q}} \mathcal{I}(\theta, \bar{\theta}) = \theta_{i_1} \theta_{i_2} \dots \theta_{i_p} \bar{\theta}_{j_1} \bar{\theta}_{j_2} \dots \bar{\theta}_{j_q} \mathcal{I}(\theta, \bar{\theta}), \quad (4.79)$$

which belongs to the Grassmann algebra \mathfrak{C} .

Such an operator $O(\theta, \partial/\partial\theta)$ defined by its kernel $\mathcal{O}(\theta, \bar{\theta})$,

$$O(\theta, \partial/\partial\theta) \mapsto \mathcal{O}(\theta, \bar{\theta}) \equiv \langle \theta | \mathcal{O} | \bar{\theta} \rangle = O(\theta, \bar{\theta}) \mathcal{I}(\theta, \bar{\theta}), \quad (4.80)$$

then acts on a function as

$$(O f)(\theta) = \int \prod_i d\theta'_i d\bar{\theta}'_i \mathcal{O}(\theta, \bar{\theta}') \exp\left(\sum_i \bar{\theta}'_i \theta'_i\right) f(\theta'). \quad (4.81)$$

The kernel associated with the product $\mathcal{O}_2 \mathcal{O}_1$ is given by

$$\langle \theta | \mathcal{O}_2 \mathcal{O}_1 | \bar{\theta} \rangle = \int \prod_i d\theta'_i d\bar{\theta}'_i \langle \theta | \mathcal{O}_2 | \bar{\theta}' \rangle \exp\left(\sum_i \bar{\theta}'_i \theta'_i\right) \langle \theta' | \mathcal{O}_1 | \bar{\theta} \rangle. \quad (4.82)$$

Trace. The trace of an operator is

$$\text{tr } \mathcal{O} = \int \prod_i d\bar{\theta}_i d\theta_i \exp\left(-\sum_i \bar{\theta}_i \theta_i\right) \langle \theta | \mathcal{O} | \bar{\theta} \rangle. \quad (4.83)$$

Note the difference of sign in the exponentials between expressions (4.82) and (4.83). One can verify that the sign in (4.83) is consistent with the cyclic property of the trace.

Hamiltonian and number operator. The Hamiltonian (4.75) is then represented by

$$\langle \theta | H_0 | \bar{\theta} \rangle = \mathcal{I}(\theta, \bar{\theta}) \sum_i \hbar \omega_i \theta_i \bar{\theta}_i, \quad (4.84)$$

and the particle number operator by

$$\langle \theta | N | \bar{\theta} \rangle = \mathcal{I}(\theta, \bar{\theta}) \sum_i \theta_i \bar{\theta}_i. \quad (4.85)$$

Complex conjugation and Hermitian conjugation of operators. The Hermitian conjugation for operator kernels is the Hermitian conjugation in the algebra \mathfrak{C} as defined by (equations (1.67)). One can verify that H_0 and N are Hermitian.

4.5.3 An example: One state system

As a consequence of the Pauli principle, the vector space for particles obeying the Fermi–Dirac statistics (fermions) that can occupy only one state reduces to a two-dimensional complex vector space: a state can only be empty (the vacuum) or occupied once.

A superposition of the zero-particle and one-particle states is represented by the affine complex function

$$\psi(\theta) = \psi_0 + \psi_1 \theta, \quad (\psi_0, \psi_1) \in \mathbb{C}^2.$$

If the function ψ is normalized by the scalar product (4.73), $\|\psi\|^2 = (\psi, \psi) = 1$, $|\psi_0|^2$ and $|\psi_1|^2$ represent the respective probability for the system to be in the vacuum state or in the occupied state. The Hamiltonian

$$H_0 = \hbar\omega\theta \frac{\partial}{\partial\theta}, \quad \omega > 0, \quad (4.86)$$

has the eigenvalue 0 for the vacuum state and $\hbar\omega$ for the one-particle state. The corresponding kernel is

$$\langle \theta | H_0 | \bar{\theta} \rangle = \hbar\omega\theta \frac{\partial}{\partial\theta} e^{-\bar{\theta}\theta} = \hbar\omega\theta\bar{\theta} e^{-\bar{\theta}\theta} = -\hbar\omega\bar{\theta}\theta. \quad (4.87)$$

One can verify that the matrix elements of the operator $U_0(t) = e^{-tH_0/\hbar}$ are

$$\langle \theta | U_0(t) | \bar{\theta} \rangle = e^{-\bar{\theta}\theta e^{-\hbar\omega t}}. \quad (4.88)$$

With our definition, both kernels are invariant under complex conjugation. From the explicit expression (4.88) and the definition (4.83), one infers the partition function

$$\mathcal{Z}_0(\beta) = \text{tr } U_0(\hbar\beta) = \int d\bar{\theta} d\theta e^{-\bar{\theta}\theta} e^{-\bar{\theta}\theta e^{-\hbar\omega\beta}} = 1 + e^{-\hbar\omega\beta}. \quad (4.89)$$

Negative energy states. If, in the Hamiltonian (4.86), ω is negative, the occupied state has an energy lower than the vacuum. This means that the two states have been misidentified; they have to be interchanged. A simple transformation deals with the problem. In the kernel (4.87), one sets $\eta = \bar{\theta}$, $\bar{\eta} = \theta$.

4.6 Path integrals with fermions

The derivation of a path integral representation for the statistical operator of fermion systems closely follows the method of Sections 4.1 and 4.2.1, with the Grassmann variables replacing the complex variables.

We first construct a path integral representation for the statistical operator (4.88). Path integrals are certainly not required to deal with this simple example. However, the path integral representation is useful, because it can easily be generalized to an arbitrary number of possible fermion states.

We expand the exact expression (4.88) for a small time-interval ε as

$$\langle \theta | U_0(\varepsilon) | \bar{\theta} \rangle = \exp [-\bar{\theta}\theta(1 - \omega\varepsilon) + O(\varepsilon^2)]. \quad (4.90)$$

The product form (4.82) makes it possible to write the statistical operator at finite time as

$$\langle \theta'' | U_0(t'', t') | \bar{\theta}' \rangle = \lim_{n \rightarrow \infty} \int \prod_{k=1}^{n-1} d\theta_k d\bar{\theta}_k \exp [-\mathcal{S}_\varepsilon(\theta, \bar{\theta})], \quad (4.91)$$

with

$$\mathcal{S}_\varepsilon(\theta, \bar{\theta}) = \bar{\theta}_0\theta_1 + \sum_{k=1}^{n-1} \bar{\theta}_k(\theta_{k+1} - \theta_k) - \varepsilon \sum_{k=0}^{n-1} \omega\bar{\theta}_k\theta_{k+1}, \quad (4.92)$$

where $\varepsilon = (t'' - t')/n$ and $\bar{\theta}_0 \equiv \bar{\theta}'$, $\theta_n \equiv \theta''$.

The formal large n limit leads for the matrix elements to the path integral representation

$$\langle \theta'' | U_0(t'', t') | \bar{\theta}' \rangle = \int_{\bar{\theta}(t')=\bar{\theta}'}^{\theta(t'')=\theta''} [d\theta(t)d\bar{\theta}(t)] \exp [-S_0(\theta, \bar{\theta})], \quad (4.93)$$

with ($\dot{\theta} \equiv d\theta/dt$)

$$S_0(\theta, \bar{\theta}) = \int_{t'}^{t''} dt \bar{\theta}(t) [\dot{\theta}(t) - \hbar\omega\theta(t)] + \bar{\theta}(t')\theta(t'). \quad (4.94)$$

Partition function. As in the holomorphic case, to calculate the trace, using the definition (4.83), we first return to the discretized expression. We find

$$S_\varepsilon(\theta, \bar{\theta}) = \bar{\theta}_0(\theta_n + \theta_0) + \sum_{k=1}^n [\bar{\theta}_{k-1}(\theta_k - \theta_{k-1}) - \hbar\omega\varepsilon\bar{\theta}_{k-1}\theta_k], \quad (4.95)$$

with the integration measure

$$\prod_{k=1}^{n-1} d\theta_k d\bar{\theta}_k d\bar{\theta}_0 d\theta_n.$$

Note the difference in sign with respect to the commuting case. We have now to set $\bar{\theta}_0 = -\bar{\theta}_n$, and similarly $\theta_n = -\theta_0$, in such a way that

$$S_\varepsilon(\theta, \bar{\theta}) = \sum_{k=1}^n [\bar{\theta}_{k-1}(\theta_k - \bar{\theta}_{k-1}) - \hbar\omega\varepsilon\bar{\theta}_k\theta_{k-1}].$$

In the continuum limit, we thus obtain

$$\mathcal{Z}_0(\beta) = \text{tr } U(\hbar\beta/2, -\hbar\beta/2) = \int [d\bar{\theta}(t)d\theta(t)] \exp [-S_0(\theta, \bar{\theta})], \quad (4.96)$$

with

$$S_0(\theta, \bar{\theta}) = \int_{-\beta/2}^{\beta/2} dt \bar{\theta}(t) [\dot{\theta}(t) - \hbar\omega\theta(t)], \quad (4.97)$$

and the *anti-periodic boundary conditions* $\theta(-\beta/2) = -\theta(\beta/2)$, $\bar{\theta}(-\beta/2) = -\bar{\theta}(\beta/2)$.

General Gaussian integral. We now introduce a general Gaussian integral, which will be useful later for generating correlation functions and perturbative expansions. We must first enlarge the Grassmann algebra \mathfrak{C} generated by the elements $\bar{\theta}(t), \theta(t)$, by adding two new sets of generators $\bar{\eta}(t)$ and $\eta(t)$. In the path integral (4.96), we then replace the action $S_0(\theta, \bar{\theta})$ by

$$S_G(\theta, \bar{\theta}) = \int_{-\beta/2}^{\beta/2} dt \left\{ \bar{\theta}(t) [\dot{\theta}(t) - \hbar\omega\theta(t)] - \bar{\eta}(t)\theta(t) - \bar{\theta}(t)\eta(t) \right\}. \quad (4.98)$$

The path integral can be calculated by a simple extension of the method used in the case of the integral (1.71). One looks for the saddle point, solution of the equations obtained by varying $\bar{\theta}(t)$ and $\theta(t)$, respectively:

$$\dot{\theta}(t) - \hbar\omega\theta(t) - \eta(t) = 0, \quad (4.99)$$

$$\dot{\bar{\theta}}(t) + \hbar\omega\bar{\theta}(t) + \bar{\eta}(t) = 0. \quad (4.100)$$

The solution of equation (4.99) with anti-periodic boundary conditions can be written as

$$\theta(t) = - \int_{-\beta/2}^{\beta/2} \Delta(u-t)\eta(u)du, \quad (4.101)$$

with

$$\Delta(t) = \frac{1}{2} e^{-\hbar\omega t} [\operatorname{sgn}(t) + \tanh(\hbar\omega\beta/2)] \quad (4.102)$$

($\operatorname{sgn}(t)$ is the sign function, $\operatorname{sgn}(t) = 1$ for $t > 0$, $\operatorname{sgn}(-t) = -\operatorname{sgn}(t)$).

The function $\Delta(t)$ is the solution of the differential equation

$$\dot{\Delta}(t) + \hbar\omega\Delta(t) = \delta(t),$$

with anti-periodic boundary conditions: $\Delta(\beta/2) = -\Delta(-\beta/2)$.

The explicit solution of equation (4.100) is not needed. To calculate the action at the saddle point, one integrates by parts $\bar{\theta}\dot{\theta}$, uses equation (4.100), and obtains

$$\mathcal{S}_G = - \int_{-\beta/2}^{\beta/2} dt \bar{\eta}(t)\theta(t) = \int_{-\beta/2}^{\beta/2} dt du \bar{\eta}(t)\Delta(u-t)\eta(u). \quad (4.103)$$

After a translation of $\theta(t)$ and $\bar{\theta}(t)$ by the solutions of the saddle point equations, the remaining path integral is the partition function $\mathcal{Z}_0(\beta)$. Thus,

$$\operatorname{tr} U_G(\hbar\beta/2, -\hbar\beta/2) = (1 + e^{-\hbar\omega\beta}) \exp \left[- \int_{-\beta/2}^{\beta/2} dt du \bar{\eta}(t)\Delta(u-t)\eta(u) \right]. \quad (4.104)$$

As in the complex case, correlation functions are obtained by differentiating the path integral with the action (4.98) and expression (4.104) with respect to η and $\bar{\eta}$. For example,

$$\langle \bar{\theta}(t_2)\theta(t_1) \rangle = -\mathcal{Z}_0^{-1}(\beta) \left. \frac{\delta}{\delta\eta(t_2)} \frac{\delta}{\delta\bar{\eta}(t_1)} \operatorname{tr} U_G(\hbar\beta/2, -\hbar\beta/2) \right|_{\eta=\bar{\eta}=0} = \Delta(t_2 - t_1). \quad (4.105)$$

Remark. By contrast, the path integral with periodic boundary conditions appears in the calculation of $\operatorname{tr}(-1)^F e^{-\beta H}$, in which F is the fermion number, and is obtained by integrating expression (4.93) with $e^{\theta''\bar{\theta}'}$.

The Gaussian integration and the sgn(0) problem. As in the complex case of Section 4.2.1, the partition function $\mathcal{Z}_0(\beta)$ can be calculated, up to a numerical factor, by relating it to $\Delta(t)$. From equation (4.96), one infers

$$\frac{\partial \ln \mathcal{Z}_0(\beta)}{\partial\omega} = \hbar \int \langle \bar{\theta}(t)\theta(t) \rangle = \hbar\beta\Delta(0).$$

Again, the function $\Delta(t)$ is not continuous at $t = 0$, and thus the result is not defined. The choice $\operatorname{sgn}(0) = -1$ is consistent with the normal order. Integrating, one then obtains

$$\mathcal{Z}_0(\beta) = 1 + e^{-\hbar\beta\omega}. \quad (4.106)$$

Another choice corresponds to a constant shift of the Hamiltonian, natural consequence of the order problem between the operators θ and $\partial/\partial\theta$. For example, if we choose $\operatorname{sgn}(0) = 0$, one obtains

$$\mathcal{Z}_0(\beta) = 2 \cosh(\hbar\beta\omega/2),$$

which corresponds to the Hamiltonian

$$H_0 = \frac{1}{2}\hbar\omega[\theta, \partial/\partial\theta].$$

4.6.1 Generalization

The partition function corresponding to the Hamiltonian (4.75) has a path integral representation with an action that is a straightforward generalization of the form (4.97)

$$\mathcal{S}_0(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \int_{-\beta/2}^{\beta/2} dt \sum_i \bar{\theta}_i(t) \left[\dot{\theta}_i(t) - \hbar\omega_i \theta_i(t) \right], \quad (4.107)$$

and with the *anti-periodic boundary conditions*

$$\boldsymbol{\theta}(-\beta/2) = -\boldsymbol{\theta}(\beta/2), \quad \bar{\boldsymbol{\theta}}(-\beta/2) = -\bar{\boldsymbol{\theta}}(\beta/2).$$

Interactions. Interactions between particles can be introduced by considering Hamiltonians which are more general elements of the operator algebra defined in Section 4.5.2. If H has been normal-ordered, which means that, with the help of the commutation relations, all operators θ_i have been moved to the left of all operators $\partial/\partial\theta_i$ in the monomials contributing to H , then

$$H(\boldsymbol{\theta}, \partial/\partial\boldsymbol{\theta}) \mapsto H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})\mathcal{I}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}),$$

where \mathcal{I} has been defined by equation (4.77). Equation (4.90) generalizes in the form

$$\langle \boldsymbol{\theta} | U(t) | \bar{\boldsymbol{\theta}} \rangle = \exp \left[- \sum_i \theta_i \bar{\theta}_i - t H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) + O(t^2) \right]. \quad (4.108)$$

At finite time, a path integral representation follows:

$$\langle \boldsymbol{\theta}'' | U(t'', t') | \bar{\boldsymbol{\theta}}' \rangle = \int_{\bar{\boldsymbol{\theta}}(t')=\bar{\boldsymbol{\theta}}'}^{\boldsymbol{\theta}(t'')=\boldsymbol{\theta}''} [d\boldsymbol{\theta}(t)d\bar{\boldsymbol{\theta}}(t)] \exp [-\mathcal{S}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})], \quad (4.109)$$

with

$$\mathcal{S}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \sum_i \bar{\theta}_i(t') \theta_i(t') + \int_{t'}^{t''} dt \left\{ \sum_i \bar{\theta}_i(t) \dot{\theta}_i(t) + H[\boldsymbol{\theta}(t), \bar{\boldsymbol{\theta}}(t)] \right\}. \quad (4.110)$$

We have shown in Section 1.7 how to calculate Gaussian integrals and average of polynomials, with the help of Wick's theorem. Here, we can use the same methods to expand the path integral (4.110) in perturbation theory. However, problems due to operator ordering arise again. In particular, as in the case of the holomorphic path integral, perturbative calculations involve $\text{sgn}(0)$. The correct ansatz, consistent with our construction, is to again set $\text{sgn}(0) = -1$. The more convenient, left-right symmetric choice $\text{sgn}(0) = 0$ requires an appropriate modification of the function $H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})$.

4.7 The Fermi gas. Field integral representation

The considerations of Section 4.6 and, in particular, the construction of the path integral for fermions, make it possible to generalize the method of Section 4.4.5 to fermions.

We start from the Hamiltonian \mathbf{H} of equation (4.45) in the n -particle subspace. Since we now deal with identical fermions (for simplicity without internal degrees of freedom), the n -particle wave functions ψ_n are antisymmetric.

Fock space. The construction of a generating functional requires the introduction of fields $\varphi(x)$, generators of a Grassmann algebra:

$$\varphi(x)\varphi(x') + \varphi(x')\varphi(x) = 0.$$

We introduce a functional that generates antisymmetric fermion wave functions:

$$\Psi(\varphi) = \sum_{n=0} \frac{1}{n!} \left(\int \prod_i d^d x_i \varphi(x_i) \right) \psi_n(x_1, \dots, x_n). \quad (4.111)$$

The scalar product of two generating functionals is defined in terms of a *Grassmann field integral*, which generalizes expression (4.73):

$$(\Psi_1, \Psi_2) = \int [d\varphi(x)d\bar{\varphi}(x)] \Psi_1^\dagger(\varphi)\Psi_2(\varphi) \exp \left[\int d^d x \bar{\varphi}(x)\varphi(x) \right], \quad (4.112)$$

normalized by the condition $(1, 1) = 1$. Using Wick's theorem, one then finds

$$\|\Psi\|^2 = (\Psi, \Psi) = \sum_{n=0} \frac{1}{n!} \int d^d x_1 \cdots d^d x_n |\psi_n(x_1, \dots, x_n)|^2 < \infty.$$

When $\|\Psi\| = 1$, the n th term in the sum is the probability that the fermion system is in an n -particle state.

The Hamiltonian. The kernel of the identity that generalizes expression (4.77) and is consistent with the scalar product (4.112) is

$$\mathcal{I}(\varphi, \bar{\varphi}) = \left(\exp - \int d^d x \bar{\varphi}(x)\varphi(x) \right).$$

For the Hamiltonian, compared to the holomorphic case, the kinetic term does not change. The potential terms also remain the same with some specified order of the fields. With the conventions of Section 4.6, one finds for \mathbf{H} ,

$$\begin{aligned} \langle \varphi | \mathbf{H} | \bar{\varphi} \rangle &= \mathcal{I}(\varphi, \bar{\varphi}) \left\{ \int d^d x \varphi(x) \left[-\frac{\hbar^2}{2m} \nabla^2 + V_1(x) \right] \bar{\varphi}(x) \right. \\ &\quad \left. + \frac{1}{2} \int d^d x d^d y \varphi(x)\varphi(y)V_2(x, y)\bar{\varphi}(y)\bar{\varphi}(x) \right\}. \end{aligned} \quad (4.113)$$

The field integral. The derivation of the field integral representation follows the same steps as in the boson case; one just has to be careful about the order of factors and the signs that may appear. The partition function is then given by the field integral

$$\mathcal{Z}(\beta) = \int [d\varphi(t, x)d\bar{\varphi}(t, x)] e^{-S(\varphi, \bar{\varphi})}, \quad (4.114)$$

where the fields $\bar{\varphi}(t, x), \varphi(t, x)$ satisfy *anti-periodic boundary conditions* in the Euclidean time direction and the Euclidean action is

$$\begin{aligned} S(\varphi, \bar{\varphi}) &= \int dt d^d x \bar{\varphi}(t, x) \left(\frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 - V_1(x) + \mu \right) \varphi(t, x) \\ &\quad + \frac{1}{2} \int dt d^d x d^d y \bar{\varphi}(t, x)\varphi(t, x)V_2(x, y)\bar{\varphi}(t, y)\varphi(t, y). \end{aligned} \quad (4.115)$$

4.7.1 Simple examples

The free Fermi gas. Let us briefly examine the free gas, the generalization to one-body potentials being straightforward. As in the case of the Bose gas, one expands the fields in Fourier components. Using the result (4.89), which gives the partition function for one degree of freedom, one obtains the free energy density and the mean energy density

$$L^{-d}\mathcal{W}(\beta) = \int \frac{d^dp}{(2\pi\hbar)^d} \ln \left(1 + e^{-\beta(p^2/2m - \mu)} \right), \quad \langle H \rangle = \int \frac{d^dp}{(2\pi\hbar)^d} \frac{p^2/2m}{e^{\beta(p^2/2m - \mu)} + 1}. \quad (4.116)$$

The sign of the chemical potential is now arbitrary; the Pauli principle prevents the collapse of the fermion system. The density takes the form

$$\rho = \frac{1}{(2\pi\hbar)^d} \int \frac{d^dp}{e^{\beta(p^2/2m - \mu)} + 1}. \quad (4.117)$$

The $\delta(x)$ -function potential. When the two-body potential is short-range and one is interested in long-wavelength phenomena, one can again approximate the potential by a δ -function. The action then becomes local (we again assume $V_1 \equiv 0$). However, in the case of fermions without internal degrees of freedom, the interaction term vanishes, because the square of a Grassmann variable vanishes, and the fermions thus remain free. Four-fermion interactions of this simple form can be non-trivial only if the fermions have an internal quantum number. A simple example is provided by an N -component fermion φ^α and an action with $U(N)$ symmetry:

$$\begin{aligned} \mathcal{S}(\varphi, \bar{\varphi}) = & \int dt d^dx \left[\sum_\alpha \bar{\varphi}^\alpha(t, x) \left(\frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla_x^2 + \mu \right) \varphi^\alpha(t, x) \right. \\ & \left. + \frac{g}{2} \left(\sum_\alpha \bar{\varphi}^\alpha(t, x) \varphi^\alpha(t, x) \right)^2 \right]. \end{aligned} \quad (4.118)$$

In the case $N = 2$, the quantum number can be associated with a non-relativistic spin (as the spin of an electron). Note that the sign of g is now arbitrary, both attractive and repulsive forces are admissible.

In dimension $d = 1$, the model with the action (4.118) is integrable. It is the non-relativistic limit of the massive Thirring model, which is also integrable (see Section 30.6).

In higher dimensions, in the case of attractive forces, the model can describe phase transitions with fermion pair $\langle \varphi \bar{\varphi} \rangle$ condensation. This situation can, for instance, be studied in the large N limit (see Section 20.5).

4.7.2 Non-relativistic Fermi gas at low temperatures, in one dimension

The equation of state (4.117) shows that, at low temperature (β large) and $\mu > 0$, only the states with momenta k below the Fermi surface, which satisfy

$$|k| \leq k_F = \sqrt{2m\mu}, \quad (4.119)$$

are occupied. Moreover, in the presence of weak interactions, only excitations with momenta close to the Fermi surface $|k| = k_F$ are important.

In one dimension, the Fermi surface reduces to the two points $k = \pm k_F$. We thus set

$$k = \pm k_F + q, \quad |q| \ll k_F.$$

The Fourier transform of the Gaussian two-point function or propagator reduces to

$$\tilde{\Delta}(\omega, k) = \frac{1}{-i\omega + k^2/2m - \mu} \sim \sum_{\epsilon=\pm 1} \frac{1}{-i\omega + \epsilon q k_F/m}.$$

This corresponds to approximating, locally near the Fermi points, the dispersion curve by straight lines. We recognize the propagators of a massless *relativistic fermion* (Chapter 12), the two values $\epsilon = \pm 1$ corresponding to right and left movers.

The quantity $c = m/\hbar k_F$ plays the role of the speed of light.

More directly, the initial fermion fields $\varphi(t, x)$, $\bar{\varphi}(t, x)$ can be parametrized in terms of four fields $\eta_i(t, x)$, $\bar{\eta}_i(t, x)$, $i = 1, 2$ which have only small-momentum components:

$$\begin{aligned} \varphi(t, x) &= e^{-ik_F x/\hbar} \eta_1(t, x) + e^{ik_F x/\hbar} \eta_2(t, x), \\ \bar{\varphi}(t, x) &= -i e^{-ik_F x/\hbar} \bar{\eta}_1(t, x) + i e^{ik_F x/\hbar} \bar{\eta}_2(t, x). \end{aligned}$$

After substitution, all terms in the action with explicit exponential factors $e^{\pm ik_F x/\hbar}$, which correspond to large momentum exchanges, can be neglected at leading order. The free fermion action then becomes

$$\mathcal{S}_0(\bar{\eta}, \eta) = \int dt dx \bar{\eta}(t, x) [\sigma_2 \partial_t + c \sigma_1 \partial_x] \eta(t, x),$$

where the two-component fermion fields $\{\bar{\eta}, \eta\}$ can be identified with a massless Dirac fermion field, and σ_μ , $\mu = 1, 2, 3$, are the Pauli matrices (see Section A12). Note that, in these covariant relativistic conventions, $\eta^\dagger = \bar{\eta} \sigma_2$.

We now consider again N -component fermion fields φ_α and the example of the action (4.118) with a two-body δ -function potential. After the same substitution (again all terms proportional to non-trivial powers of $e^{\pm ik_F x/\hbar}$ being neglected), the action becomes

$$\begin{aligned} \mathcal{S}(\bar{\eta}, \eta) &= \int dt dx \left[\sum_\alpha \bar{\eta}^\alpha(t, x) (\sigma_2 \partial_t + c \sigma_1 \partial_x) \eta^\alpha(t, x) + \frac{g}{2} \left(\sum_\alpha \bar{\eta}^\alpha(t, x) \sigma_2 \eta^\alpha(t, x) \right)^2 \right. \\ &\quad \left. + \frac{g}{4} \left(\sum_\alpha \bar{\eta}^\alpha(t, x) \eta^\alpha(t, x) \right)^2 - \frac{g}{4} \left(\sum_\alpha \bar{\eta}^\alpha(t, x) \sigma_3 \eta^\alpha(t, x) \right)^2 \right]. \end{aligned}$$

The two-component fermion fields $\{\bar{\eta}^\alpha, \eta^\alpha\}$ represent N massless Dirac fermions. The action has a chiral symmetry corresponding to the transformations

$$\eta^\alpha \mapsto e^{i\sigma_3 \theta} \eta^\alpha, \quad \bar{\eta}^\alpha \mapsto \bar{\eta}^\alpha e^{i\sigma_3 \theta},$$

which prevents the addition of a fermion mass term (see Section A12).

Since the non-relativistic spin appears as an external quantum number, a spin 1/2 fermion yields a doublet of massless Dirac fermions. A two-fermion model ($N = 2$), which is considered in Sections 30.7 and 31.6, thus describes self-interacting one-dimensional non-relativistic fermions at low temperatures. All results derived in this way have an interpretation in terms of the Kosterlitz–Thouless phase transition, or in terms of systems like the Luttinger liquid (the equivalent of the Fermi liquid of higher dimensions), or one-dimensional conductors [20].

In higher space dimensions, the situation is no longer that simple and does not lead to massless relativistic fermions any more.

5 Quantum evolution: From particles to non-relativistic fields

In Chapter 4, we have discussed quantum statistical operators and thus, formally, evolution in imaginary or Euclidean time. By contrast, to calculate scattering S -matrix elements, quantities relevant to particle physics, it is necessary to consider quantum evolution operators in real time.

To a large extent, the functional integral representations of the evolution operator can be inferred from the corresponding expressions of Chapter 4 by returning from imaginary to real time. However, the representation of the S -matrix requires an additional discussion.

To illustrate the power of the formalism, we will show how to recover the perturbative expansion of the scattering amplitude, some semi-classical approximations, and the eikonal approximation.

When the asymptotic states at large time are eigenstates of the harmonic oscillator, instead of free particles, the holomorphic formalism becomes useful. A simple generalization of the path integral of Section 4.1 leads to the corresponding path integral representation of the S -matrix. In the case of the Bose gas, the evolution operator is then given by a holomorphic field integral.

Using the parallel formalism of Section 4.7, an analogous representation for the evolution operator of a system of non-relativistic fermions can then be derived.

5.1 Time evolution and scattering matrix in quantum mechanics (QM)

Quantum evolution is associated with an operator acting linearly on the Hilbert space of states. For an isolated system, conservation of probabilities implies that the evolution operator must be unitary. Moreover, one assumes that the evolution of an isolated system is Markovian (without memory effects).

5.1.1 Evolution operator and S -matrix

The unitary evolution operator $U(t'', t')$ from time t' to time t'' satisfies the relation

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1). \quad (5.1)$$

Assuming that the evolution operator is a smooth function of times, we expand $U(t+\varepsilon, t)$ for an infinitesimal time interval ε as

$$U(t + \varepsilon, t) = \mathbf{1} - i\varepsilon H(t)/\hbar + O(\varepsilon^2),$$

where $H(t)$ is the quantum Hamiltonian. Equation (5.1) then implies

$$i\hbar \frac{\partial U}{\partial t}(t, t') = H(t)U(t, t'), \quad \text{with } U(t', t') = \mathbf{1}. \quad (5.2)$$

When the operator $H(t)$ is time independent, $U(t'', t') = U(t'' - t') = e^{-i(t'' - t')H/\hbar}$. The evolution operators $U(t)$ then form a representation of the symmetry group of time translations. The generator of the Lie algebra $\partial/\partial t$ is represented by the operator $-iH/\hbar$.

The S-matrix. The scattering S -matrix is obtained by comparing the quantum evolution to the free evolution in the absence of interactions. More precisely, the S -matrix can be defined as the limit of the evolution operator in the interaction representation

$$S = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} e^{iH_0 t''/\hbar} U(t'', t') e^{-iH_0 t'/\hbar}, \quad (5.3)$$

where the Hamiltonian H_0 is the free Hamiltonian corresponding to H : its eigenstates are the free or asymptotic states at large time of quantum evolution.

5.1.2 One-particle system

For a quantum particle in a potential, in d space dimensions, examples of free and interacting Hamiltonians are

$$H_0 = \mathbf{p}^2/2m, \quad H = \mathbf{p}^2/2m + V(\mathbf{q}, t). \quad (5.4)$$

The large time limit and, therefore, the S -matrix exist only if the potential decreases fast enough at large distance, or large time in such a way that for $|t| \rightarrow \infty$ the evolution converges fast enough towards the free evolution corresponding to H_0 .

The free evolution operator $U_0 = e^{-i(t''-t')H_0/\hbar}$, corresponding to the Hamiltonian H_0 in equation (5.4), in d space dimensions, is given by

$$\begin{aligned} \langle \mathbf{q}'' | U_0(t'', t') | \mathbf{q}' \rangle &= \frac{1}{(2\pi)^d} \int d^d p \exp \frac{i}{\hbar} [\mathbf{p} \cdot (\mathbf{q}'' - \mathbf{q}') - \mathbf{p}^2(t'' - t')/2m] \\ &= \left(\frac{m}{2i\pi\hbar(t'' - t')} \right)^{d/2} \exp \left[\frac{i}{\hbar} \frac{m(\mathbf{q}'' - \mathbf{q}')^2}{2(t'' - t')} \right]. \end{aligned} \quad (5.5)$$

In the momentum basis, the relation (5.3) between the S -matrix and the evolution operator takes the form

$$\langle \mathbf{p}'' | S | \mathbf{p}' \rangle = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} e^{iE''t''/\hbar} \langle \mathbf{p}'' | U(t'', t') | \mathbf{p}' \rangle e^{-iE't'/\hbar}, \quad (5.6)$$

where

$$E' = E(\mathbf{p}'), \quad E'' = E(\mathbf{p}''), \quad E(\mathbf{p}) = \mathbf{p}^2/2m.$$

The limits $t' \rightarrow -\infty$, $t'' \rightarrow +\infty$ have to be understood mathematically in the sense of distributions (one should use test functions in the form of wave packets).

In the momentum basis, the S -matrix is, in general, parametrized in terms of the scattering matrix \mathcal{T} :

$$S = 1 - i\mathcal{T}, \Rightarrow \langle \mathbf{p}'' | S | \mathbf{p}' \rangle = (2\pi\hbar)^d \delta^{(d)}(\mathbf{p}'' - \mathbf{p}') - i \langle \mathbf{p}'' | \mathcal{T} | \mathbf{p}' \rangle. \quad (5.7)$$

When the potential is time-independent, energy is conserved, and one can express \mathcal{T} in the form

$$\langle \mathbf{p}'' | \mathcal{T} | \mathbf{p}' \rangle = -2\pi\delta(E'' - E')T(\mathbf{p}'', \mathbf{p}'). \quad (5.8)$$

5.1.3 Path integrals

To calculate matrix elements of the evolution operator, we start from expression (2.19) and proceed by analytic continuation replacing t by $t e^{i\varphi}$ in all expressions and rotating in the complex t plane from $\varphi = 0$ to $\varphi = \pi/2$ in the positive direction (but leaving the argument of $H(t)$ unchanged).

For example, a Hamiltonian of the form $H = p^2/2m + V(q, t)$ leads to the representation

$$\langle q'' | U(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [dq(t)] \exp [i\mathcal{A}(q)/\hbar], \quad (5.9)$$

where $\mathcal{A}(q)$, which replaces the Euclidean action, is now the usual classical action, integral of the Lagrangian,

$$\mathcal{A}(q) = \int_{t'}^{t''} dt \mathcal{L}(q(t), \dot{q}(t); t), \quad \text{with here } \mathcal{L}(q, \dot{q}; t) = \frac{1}{2}m\dot{q}^2 - V(q, t). \quad (5.10)$$

Expression (5.9) establishes a beautiful relation between classical and QM. In a quantum evolution, all paths contribute to the path integral and they are weighted with the complex weight $e^{i\mathcal{A}/\hbar}$ involving the classical action.

The classical variational principle. Paths close to the extrema of the action, that is, paths that are solutions of the classical equations of motion, give the largest contributions to the path integral. In particular, if for classical paths the value of the classical action is large compared to \hbar , paths close to the classical paths completely dominate the path integral. Therefore, *the path integral description of QM explains why fundamental classical equations satisfy a variational principle*.

Phase-space formulation. In the real-time formulation, the path integral over phase space (3.21) is replaced by

$$\langle q'' | U(t'', t') | q' \rangle = \int [dp(t)dq(t)] \exp [i\mathcal{A}(p, q)/\hbar]. \quad (5.11)$$

The quantity $\mathcal{A}(p, q)$ is still the classical action, but in the Hamiltonian formalism. In terms of the classical Hamiltonian $\mathcal{H}(p(t), q(t); t)$, it reads

$$\mathcal{A}(p, q) = \int_{t'}^{t''} [p(t)\dot{q}(t) - \mathcal{H}(p(t), q(t); t)] dt. \quad (5.12)$$

Even in this more general situation, the quantum evolution is obtained by summing over all paths weighted with the complex weight $e^{i\mathcal{A}/\hbar}$, where \mathcal{A} is the classical action in the phase space formalism.

5.2 Path integral and S -matrix: Perturbation theory

We now show how the path integral giving the evolution operator can be calculated in the form of an expansion in powers of the potential. The path integral formalism actually organizes the perturbative expansion in a way similar to the operator formalism recalled in Section A5. From the expansion of the evolution operator, we then derive the expansion of the elements of the S -matrix.

We consider the time-independent Hamiltonian

$$H = p^2/2m + V(x). \quad (5.13)$$

The classical actions corresponding to the free Hamiltonian H_0 and to H , respectively, are

$$\mathcal{A}_0(x) = \int_{t'}^{t''} \frac{1}{2} m \dot{x}^2(t) dt, \quad \mathcal{A}(x) = \int_{t'}^{t''} [\frac{1}{2} m \dot{x}^2(t) - V(x(t))] dt. \quad (5.14)$$

We expand the path integral (5.9) in powers of the potential,

$$\langle x'' | U(t'', t') | x' \rangle = \int_{x(t')=x'}^{x(t'')=x''} [dx(t)] \exp[i\mathcal{A}(x)] = \sum_{\ell} \langle x'' | U^{(\ell)}(t'', t') | x' \rangle$$

(setting, for convenience, $\hbar = 1$), with

$$\langle x'' | U^{(\ell)}(t'', t') | x' \rangle = \frac{(-i)^\ell}{\ell!} \int_{x(t')=x'}^{x(t'')=x''} [dx(t)] e^{i\mathcal{A}_0(x)} \left[\int_{t'}^{t''} V(x(t)) dt \right]^\ell. \quad (5.15)$$

Since the potential must vanish at large distances, it makes sense to assume that it has a Fourier representation of the form,

$$V(x) = (2\pi)^{-d} \int d^d k e^{ikx} \tilde{V}(k). \quad (5.16)$$

Substituting the representation (5.16) into the path integral, one finds,

$$\begin{aligned} \langle x'' | U^{(\ell)}(t'', t') | x' \rangle &= \frac{(-i)^\ell}{\ell!} \int_{t'}^{t''} \prod_j d\tau_j \int \prod_{j=1}^{\ell} \tilde{V}(k_j) \frac{d^d k_j}{(2\pi)^d} \\ &\times \int_{x(t')=x'}^{x(t'')=x''} [dx(t)] \exp i \left[\int_{t'}^{t''} \frac{1}{2} m \dot{x}^2(t) dt + \sum_j k_j x(\tau_j) \right]. \end{aligned} \quad (5.17)$$

Each term in the perturbative expansion involves only the calculation of a Gaussian path integral. The integrand in expression (5.17) is symmetric in the times $\tau_1, \tau_2, \dots, \tau_\ell$. Therefore, one can order them as $t'' \geq \tau_\ell \geq \tau_{\ell-1} \dots \geq \tau_1 \geq t'$ and suppress the factor $1/\ell!$.

Since the path integral is Gaussian, it is obtained, up to a normalization, by replacing the path $x(t)$ by the solution of the classical equation of motion:

$$-m\ddot{x} + \sum_j k_j \delta(t - \tau_j) = 0 \Rightarrow \dot{x}(\tau_{j+}) - \dot{x}(\tau_{j-}) = k_j/m.$$

This gives a simple interpretation to the terms in the perturbative expansion: the leading path contributing to the ℓ th order is a succession of free motions, where, at times τ_1, \dots, τ_ℓ , the momentum changes by the amounts k_1, \dots, k_ℓ . We then integrate the corresponding phase factor over all times and over all momenta weighted with $\tilde{V}(k)$.

The term of order zero in V yields a contribution $(2\pi)^d \delta(p'' - p')$ to the S -matrix. To calculate the general term explicitly, we introduce δ -functions in the form

$$\exp[ik_j x(\tau_j)] = \int d^d x_j \delta(x_j - x(\tau_j)) \exp(i k_j x_j).$$

Then, in each interval $\tau_j \leq t \leq \tau_{j+1}$, one recognizes the matrix elements of the free evolution operator, which has the representation (5.5).

Moreover, we calculate the Fourier transform with respect to x' and x'' , calling p' and p'' , respectively, the corresponding momenta. We then find

$$\begin{aligned} \langle p'' | U^{(\ell)}(t'', t') | p' \rangle &= (-i)^\ell \int \prod_{j=1}^{\ell} d\tau_j \tilde{V}(k_j) \frac{d^d k_j}{(2\pi)^d} d^d x_j \prod_{j=2}^{\ell} \frac{d^d p_j}{(2\pi)^d} \\ &\times \exp \left[\sum_{j=1}^{\ell+1} -ip_j^2(\tau_j - \tau_{j-1})/2m + ip_j(x_j - x_{j-1}) + ik_j x_j \right], \end{aligned}$$

with the conventions

$$\tau_0 = t', \quad \tau_{\ell+1} = t'', \quad p_{\ell+1} = p'', \quad p_1 = p', \quad x_0 = x_{\ell+1} = 0.$$

The integration over the variables x_j yields δ -functions, which determine the variables k_j : $k_j = p_{j+1} - p_j$. After factorization of the free evolution operator on both sides, the limits $t'' \rightarrow +\infty$, $t' \rightarrow -\infty$ can be taken. The corresponding S -matrix elements follow:

$$\begin{aligned} \langle p'' | S^{(\ell)} | p' \rangle &= (-i)^\ell \int \prod_{j=1}^{\ell} d\tau_j \prod_{j=2}^{\ell} \frac{d^d p_j}{(2\pi)^d} e^{ip''^2 \tau_\ell / 2m} \tilde{V}(p'' - p_\ell) \\ &\times e^{ip_\ell^2 (\tau_{\ell-1} - \tau_\ell) / 2m} \tilde{V}(p_\ell - p_{\ell-1}) \cdots e^{ip_2^2 (\tau_1 - \tau_2) / 2m} \tilde{V}(p_2 - p') e^{-ip'^2 \tau_1 / 2m}. \quad (5.18) \end{aligned}$$

We still have to integrate over the times τ_j . We set

$$\tau_{j+1} = \tau_j + u_j, \quad u_j \geq 0.$$

The remaining integral over τ_1 yields a δ -function of energy conservation: $2\pi\delta(E'' - E')$, which reflects time-translation invariance. The integrals over the variables u_j on the positive axis yield mathematical distributions:

$$\int_0^{+\infty} du_j e^{i(E'' - E(p_j))u_j} = \frac{i}{E'' - E(p_j) + i\epsilon}, \quad \text{with } E(p) \equiv p^2/2m \text{ and } \epsilon \rightarrow 0_+.$$

The $i\epsilon$ term in the denominator identifies the distribution as a boundary value of an analytic function and indicates how to avoid the pole at a $p_j^2 = p''^2$. The final result is then

$$\langle p'' | S^{(\ell)} | p' \rangle = -2i\pi\delta(E'' - E') \int \tilde{V}(p'' - p_\ell) \prod_j \frac{d^d p_j}{(2\pi)^d} \frac{\tilde{V}(p_j - p_{j-1})}{E'' + i\epsilon - E(p_j)}.$$

In this form, the perturbation series is a geometric series whose sum satisfies an integral equation, called the Lippmann–Schwinger equation. In terms of the operator $T(E)$, where E is generically complex, the solution of

$$T(E) = V - VG_0(E)T(E), \quad \text{with } G_0(E) = (H_0 - E)^{-1},$$

the quantity $T(p'', p')$ that appears in equation (5.8) is given by

$$T(p'', p') = \langle p'' | T(E + i\epsilon) | p' \rangle \text{ for } E = p'^2/2m = p''^2/2m.$$

5.3 Path integral and S -matrix: Semi-classical expansions

The representation (5.9) of the evolution operator leads to a path integral representation for elements of the scattering S -matrix, which is particularly well-suited to the study of the semi-classical limit, and which we describe here.

5.3.1 Path integral and S -matrix

We calculate the elements of the S -matrix between two wave packets:

$$\langle \psi_2 | S | \psi_1 \rangle = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} \int dq' dq'' \langle \psi_2 | e^{iH_0 t''/\hbar} | \mathbf{q}'' \rangle \langle \mathbf{q}'' | U(t'', t') | \mathbf{q}' \rangle \langle \mathbf{q}' | e^{-iH_0 t'/\hbar} | \psi_1 \rangle. \quad (5.19)$$

Introducing the two wave functions $\tilde{\psi}_1(\mathbf{p})$ and $\tilde{\psi}_2(\mathbf{p})$ in the momentum basis associated with the vectors $|\psi_1\rangle$ and $|\psi_2\rangle$, we define

$$\psi_1(\mathbf{q}, t) = \langle q | e^{-iH_0 t/\hbar} | \psi_1 \rangle = \int \frac{d^d p}{(2\pi)^d} \tilde{\psi}_1(\mathbf{p}) \exp \left[i \left(\mathbf{p} \cdot \mathbf{q} - t \frac{\mathbf{p}^2}{2m} \right) / \hbar \right] \quad (5.20)$$

and a similar expression for ψ_2 .

When t becomes large, the phase in expression (5.20) varies rapidly, and the integral is then dominated by the stationary points of the phase:

$$\frac{\partial}{\partial \mathbf{p}} \left(\mathbf{p} \cdot \mathbf{q} - t \frac{\mathbf{p}^2}{2m} \right) = 0 \implies \mathbf{q} = t \frac{\mathbf{p}}{m}. \quad (5.21)$$

The integral (5.20) is thus equivalent to

$$\psi_1(\mathbf{q}, t) \underset{|t| \rightarrow \infty}{\sim} \tilde{\psi}_1(\mathbf{p}) \frac{1}{(2\pi)^{d/2}} \left(\frac{m\hbar}{|t|} \right)^{d/2} \exp \left(\frac{i\pi}{4} \operatorname{sgn} t + it \frac{\mathbf{p}^2}{2m\hbar} \right), \quad (5.22)$$

with

$$\mathbf{p} = m\mathbf{q}/t.$$

We then change variables in integral (5.19), setting

$$\mathbf{q}' = t' \mathbf{p}' / m, \quad \mathbf{q}'' = t'' \mathbf{p}'' / m, \quad (5.23)$$

and obtain

$$\begin{aligned} \langle \psi_2 | S | \psi_1 \rangle &\propto \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} \int dp' dp'' \tilde{\psi}_2^*(\mathbf{p}'') \tilde{\psi}_1(\mathbf{p}') \exp \left[\frac{i}{\hbar} \left(t'' \frac{\mathbf{p}''^2}{2m} - t' \frac{\mathbf{p}'^2}{2m} \right) \right] \\ &\times \langle t'' \mathbf{p}'' / m | U(t'', t') | t' \mathbf{p}' / m \rangle. \end{aligned} \quad (5.24)$$

We introduce the path integral representation (5.9) of the evolution operator in the equation and find

$$\langle t'' \mathbf{p}'' / m | U(t'', t') | t' \mathbf{p}' / m \rangle = \int_{q(t')=t' \mathbf{p}' / m}^{q(t'')=t'' \mathbf{p}'' / m} [dq(t)] \exp(i\mathcal{A}(q)/\hbar).$$

We conclude that the S -matrix is obtained by calculating the path integral with classical scattering boundary conditions, that is, summing over paths solutions at large times of the free classical equation of motion [21]. In particular, if we know how to solve the classical equations of motion with such boundary conditions, we can calculate the evolution operator and thus the S -matrix for \hbar small. This leads to semi-classical approximations of the S -matrix. A calculation in this spirit is presented in the next section.

5.3.2 One dimension: Semi-classical limit

To illustrate the previous considerations, we consider the one-dimensional Hamiltonian

$$H = p^2/2m + V(x),$$

and perform the explicit calculation of the S -matrix for $\hbar \rightarrow 0$, using the path integral representation and the stationary phase method. We assume a generic analytic potential, decreasing fast enough at large distances. In one dimension, in the semi-classical limit, due to unitarity the scattering generates a pure phase shift.

Forward scattering. We assume that the energy is larger than the maximal value of the potential so that classical forward scattering is possible. Integrating once the classical equation of motion, we obtain

$$\frac{1}{2}m\dot{x}^2(\tau) + V(x(\tau)) = \frac{1}{2}\kappa^2/m, \quad \kappa > 0,$$

with the boundary conditions ($\tau'' > \tau'$)

$$x(\tau') = x' < x(\tau'') = x''.$$

We set $X = x'' - x'$, $T = \tau'' - \tau'$. We know from the general analysis of Section 5.3.1 that relevant trajectories correspond to classical scattering, with $mX/T = k$ finite when $\tau' \rightarrow -\infty$, $\tau'' \rightarrow \infty$. In these limits, the boundary condition

$$T = m \int_{x'}^{x''} \frac{dx}{\sqrt{\kappa^2 - 2mV(x)}},$$

leads to

$$\kappa = k + \frac{m}{T} \int_{-\infty}^{+\infty} dx \left(\frac{k}{\sqrt{k^2 - 2mV(x)}} - 1 \right) + O(T^{-2}).$$

The value of the action for the trajectory is then

$$\begin{aligned} \mathcal{A} &= \frac{\kappa^2 T}{2m} - 2 \int_{t'}^{t''} V(x(\tau)) d\tau = \frac{\kappa^2 T}{2m} - 2m \int_{x'}^{x''} \frac{V(x) dx}{\sqrt{\kappa^2 - 2mV(x)}} \\ &= \frac{\kappa^2 T}{2m} + \int_{-\infty}^{+\infty} dx \left(\sqrt{k^2 - 2mV(x)} - k \right) + O(T^{-1}). \end{aligned}$$

We calculate the Fourier transform,

$$S_+(k) = \int dx' dx'' e^{ik''x'' - ik'x'} e^{i\mathcal{A}/\hbar}.$$

Since the result depends only on $x'' - x'$, the Fourier transform contains a factor of momentum conservation $\delta(k'' - k')$. The remaining integral over X is calculated by the steepest descent method. At leading order, for T large, we need considering only the terms of order T for the saddle point. We find

$$X = k'T/m \Rightarrow k' = k.$$

As shown in Section 5.1.2, the terms proportional to T cancel with the free motion factors and, finally, we obtain the semi-classical Wentzel–Kramers–Brillouin (WKB) phase shift

$$\delta_+(k) = \frac{1}{i} \ln S_+(k) \sim \frac{1}{\hbar} \int_{-\infty}^{+\infty} dx \left(\sqrt{k^2 - 2mV(x)} - k \right).$$

Backward scattering. By contrast, we now assume that the energy is lower than the maximum value of the potential, a situation of classical reflection. We consider scattering from the left, that is, x', x'' large and negative. After an analogous calculation, we find the classical action:

$$\mathcal{A} = \frac{m}{2T} (x' + x'' - 2x_0)^2 + 2 \int_{-\infty}^{x_0} dx \left(\sqrt{k^2 - 2mV(x)} - k \right) + O(T^{-1}),$$

where x_0 is the reflection point, a solution of $k^2 = 2mV(x_0)$, and $k = m(2x_0 - x' - x'')/T$.

Since the result depends only on the combination $x' + x''$, after Fourier transformation we find the expected factor $\delta(k'' + k')$. The remaining integral over $X = x' + x''$, calculated by the steepest descent method, yields

$$2x_0 - X = k'T/m, \quad \text{with } k' = -k.$$

We find the corresponding WKB phase shift

$$\delta_-(k) = \frac{1}{i} \ln S_-(k) = \frac{2}{\hbar} \int_{-\infty}^{x_0} dx \left(\sqrt{k^2 - 2V(x)} - k \right) + \frac{2kx_0}{\hbar}.$$

Finally, note that, in the case of an analytic potential, the results in the forbidden region can be obtained by a proper analytic continuation in the potential and yield scattering amplitudes decreasing as $\exp(-\text{const.}/\hbar)$, a form typical of barrier penetration effects.

5.3.3 Eikonal approximation and path integral

From the path integral representation of the S -matrix, it is possible to derive a well-known approximation for the scattering amplitude, valid in the high-energy, low-momentum transfer regime: the eikonal approximation [22].

In the absence of a potential, the evolution operator is given by a Gaussian path integral, which can be calculated, up to a normalization, by replacing the path by the solution of the classical equation of motion. The solution that satisfies the boundary conditions implied by the representation (5.9) and corresponds to the free Hamiltonian $H_0 = \mathbf{p}^2/2m$, with $\mathbf{p} \in \mathbb{R}^d$, is

$$\mathbf{q}(t) = \mathbf{q}' + (\mathbf{q}'' - \mathbf{q}') \frac{t - t'}{t'' - t'}. \quad (5.25)$$

Translating the integration variables $\mathbf{q}(t)$ by the classical solution (5.25), we still have to calculate a normalization integral, which can be obtained by comparing with the exact result ($\hbar = 1$) (5.5).

The idea of the eikonal approximation is that, in the momentum regime

$$\mathbf{p}' = \mathbf{p} - \mathbf{k}/2, \quad \mathbf{p}'' = \mathbf{p} + \mathbf{k}/2, \quad \mathbf{p}^2 \rightarrow \infty, \quad \mathbf{p}^2 \gg \mathbf{k}^2,$$

the kinetic term $\mathbf{p}^2/2m$ dominates the action (a situation already encountered in Section 2.8.1). Therefore, the leading contributions to the path integral come from paths close to the straight lines of the free motion (5.25).

We thus calculate the action for the trajectories (5.25) and, in the integral over fluctuations around the classical trajectory, we keep only the kinetic term, which leads to a simple normalization factor. The calculation of the evolution operator for the Hamiltonian $H = \mathbf{p}^2/2m + V(\mathbf{q})$ is then straightforward. Setting $\mathbf{q}'' - \mathbf{q}' = \mathbf{s}$, $(\mathbf{q}'' + \mathbf{q}')/2 = \mathbf{x}$, one finds

$$\langle \mathbf{p} + \mathbf{k}/2 | U(t'', t') | \mathbf{p} - \mathbf{k}/2 \rangle \propto \int d^d s d^d x \exp [-i(\mathbf{p} \cdot \mathbf{s} + \mathbf{k} \cdot \mathbf{x}) + i\mathcal{A}(\mathbf{s}, \mathbf{x})], \quad (5.26)$$

in which the classical action is

$$\mathcal{A}(\mathbf{s}, \mathbf{x}) = \frac{im}{2} \frac{\mathbf{s}^2}{t'' - t'} - i \int_{t'}^{t''} dt V \left(\mathbf{x} - \frac{\mathbf{s}}{2} + \frac{t - t'}{t'' - t'} \mathbf{s} \right). \quad (5.27)$$

The normalization in equation (5.26) is determined by comparing with result (5.5) for the free motion.

Taking the large time limit, and neglecting the contribution of the potential, one finds that the integral over \mathbf{s} is dominated by the saddle point

$$\mathbf{s} = (t'' - t') \mathbf{p} / m. \quad (5.28)$$

After the substitution (5.28) and the shift $t - (t' + t'')/2 \mapsto t$, and thus $|t| \leq (t'' - t')/2$, the argument in the potential becomes $\mathbf{x} + t\mathbf{p}/m$. We assume that the potential decreases fast enough for the integral in (5.27) to have a large $t'' - t' \rightarrow \infty$ limit. The contribution of the potential to the action then becomes

$$-i \int_{-\infty}^{+\infty} dt V(\mathbf{x} + t\mathbf{p}/m).$$

After a shift of the integration variable t : $t \mapsto t - m\mathbf{x} \cdot \mathbf{p}/\mathbf{p}^2$, the argument of the potential V becomes $\mathbf{b} + t\mathbf{p}/m$ where \mathbf{b} is the component of \mathbf{x} orthogonal to \mathbf{p} :

$$\mathbf{b} = \mathbf{x} - \mathbf{p}(\mathbf{x} \cdot \mathbf{p}/\mathbf{p}^2). \quad (5.29)$$

The integral over the component of \mathbf{x} along \mathbf{p} can then be performed and implies $\mathbf{p} \cdot \mathbf{k} = 0$, which expresses energy conservation.

Finally, one obtains

$$\begin{aligned} \langle \mathbf{p} + \mathbf{k}/2 | U(t'', t') | \mathbf{p} - \mathbf{k}/2 \rangle &\simeq \delta(\mathbf{p} \cdot \mathbf{k}) \mathcal{N}(\mathbf{p}) \\ &\times \int d^{d-1} b e^{-i\mathbf{k} \cdot \mathbf{b}} \exp \left[-i \int dt V \left(\frac{\mathbf{p}t}{m} + \mathbf{b} \right) \right], \end{aligned} \quad (5.30)$$

with

$$\mathcal{N}(\mathbf{p}) \propto \exp \left[i(t'' - t') \frac{\mathbf{p}^2}{2m} \right]. \quad (5.31)$$

The insertion of the result (5.30) into equation (5.6) yields the elements of the scattering S -matrix and, thus, the transition operator \mathcal{T} . The scattering amplitude $T(\mathbf{p} + \mathbf{k}/2, \mathbf{p} - \mathbf{k}/2)$, defined by equations (5.7, 5.8), in the eikonal approximation is

$$\begin{aligned} T(\mathbf{p} + \mathbf{k}/2, \mathbf{p} - \mathbf{k}/2) &\\ &\simeq \frac{i|\mathbf{p}|}{m} \int \frac{d^{d-1} b}{(2\pi)^d} e^{-i\mathbf{k} \cdot \mathbf{b}} \left\{ \exp \left[-i \int_{-\infty}^{+\infty} dt V \left(\frac{\mathbf{p}t}{m} + \mathbf{b} \right) \right] - 1 \right\}. \end{aligned} \quad (5.32)$$

Eikonal approximation and Coulomb potential. We consider the Coulomb-like potential

$$V(\mathbf{q}) = \alpha / |\mathbf{q}|. \quad (5.33)$$

In this case, the integral over the potential in expression (5.32) diverges because the potential decreases too slowly at large distance. Integrating over a finite time interval, one finds

$$\int_{(t'-t'')/2}^{(t''-t')/2} dt V\left(\frac{\mathbf{p}t}{m} + \mathbf{x}\right) \underset{t''-t' \rightarrow \infty}{\sim} \frac{2\alpha m}{p} \ln((t'' - t')p/m b). \quad (5.34)$$

The appearance of this infinite phase has the following interpretation: since the Coulomb potential decreases too slowly at large distance, the classical trajectory converges too slowly towards the free motion which, in our definition of the S -matrix, has been taken as the reference motion. As a consequence, in the case of the Coulomb potential, only cross-sections are well-defined, not amplitudes.

Factorizing the infinite phase, one can complete the calculation of the scattering amplitude. Integrating over the vector \mathbf{b} , one obtains

$$\begin{aligned} T(\mathbf{p} + \mathbf{k}/2, \mathbf{p} - \mathbf{k}/2) &\simeq \frac{i\pi^{(d-1)/2}}{(2\pi)^d} \frac{p}{m} \exp\left[-i\frac{2\alpha m}{p} \ln((t'' - t')p/m b)\right] \\ &\times \frac{\Gamma[\frac{1}{2}(d-1) - \theta]}{\Gamma(\theta)} \left(\frac{\mathbf{k}^2}{4}\right)^{[\theta+(1-d)/2]}, \end{aligned} \quad (5.35)$$

with

$$\theta = -i\alpha m/p. \quad (5.36)$$

In three dimensions, the expression (5.35) is identical to the exact result. It also contains, for $\alpha < 0$, the correct Coulomb bound state energies E_n , which are given by the poles of the scattering amplitude

$$\theta = \frac{d-1}{2} + n \quad \Rightarrow E_n = \frac{\mathbf{p}^2}{2m} = -\frac{2\alpha^2 m}{(d-1+2n)^2}. \quad (5.37)$$

The eikonal approximation has a relativistic generalization in quantum electrodynamics. It again yields quite interesting expressions for the energy of bound states. It is obtained by an approximate summation of ladder and crossed ladder Feynman diagrams [23].

5.4 S -matrix and holomorphic formalism

The holomorphic formalism has been discussed in Section 4.1, and we adapt the expressions to real-time evolution here. The holomorphic formalism in real time is useful when the asymptotic states are eigenstates of the harmonic oscillator, a situation that one encounters naturally in quantum many-body theory (see Section 4.3) and relativistic QFT, as we start discussing in Section 6.1.1.

5.4.1 Path integrals

The path integral representation of the evolution operator is formally obtained by the continuation $t \mapsto it$. Then,

$$U(z'', \bar{z}'; t'', t') = \int \left[\frac{d\bar{z}(t)dz(t)}{2i\pi} \right] \exp [i\mathcal{A}(z, \bar{z})],$$

$$\mathcal{A}(z, \bar{z}) = -i\bar{z}(t')z(t') - \int_{t'}^{t''} dt [i\bar{z}(t)\dot{z}(t) + h(z(t), \bar{z}(t))],$$

with the boundary conditions $z(t'') = z'', \bar{z}(t') = \bar{z}'$.

From the evolution operator one can derive the corresponding S -matrix. If the S -matrix is defined by expression (5.3), where H_0 is the Hamiltonian of the harmonic oscillator (4.1),

$$H_0 = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{q}^2, \quad \omega > 0,$$

then,

$$S(z, \bar{z}) = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} \int \frac{dz'' d\bar{z}''}{2i\pi} \frac{dz' d\bar{z}'}{2i\pi} e^{-z''\bar{z}''} e^{-z'\bar{z}'} e^{i\omega t''/2} \exp \left(z\bar{z}'' e^{i\omega t''} \right)$$

$$\times U(z'', \bar{z}'; t'', t') e^{-i\omega t'/2} \exp \left(z'\bar{z} e^{-i\omega t'} \right).$$

Using equation (4.7), or integrating directly, one obtains

$$S(z, \bar{z}) = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} e^{i\omega t''/2} U(z e^{i\omega t''}, \bar{z} e^{-i\omega t'}; t'', t') e^{-i\omega t'/2}. \quad (5.38)$$

The coefficients of the expansion of $S(z, \bar{z})$ in powers of z and \bar{z} yield the matrix elements S_{mn} of the transition between the corresponding eigenstates of the harmonic oscillator,

$$S(z, \bar{z}) = \sum_{m,n} S_{mn} \frac{z^m}{\sqrt{m!}} \frac{\bar{z}^n}{\sqrt{n!}}.$$

As in the position representation (Section 5.1.2), the paths that give the leading contributions to the path integral representing the S -matrix are asymptotic, for large times, to the solutions of the classical equation of motion. For the harmonic oscillator H_0 , this implies

$$z(t'') \underset{t'' \rightarrow +\infty}{\sim} z e^{i\omega t''}, \quad \bar{z}(t') \underset{t' \rightarrow -\infty}{\sim} \bar{z} e^{-i\omega t'}.$$

5.4.2 Time-dependent force

In the case of a finite number of degrees of freedom, a simple application is the evaluation of transition rates between eigenstates of the harmonic oscillator induced by a time-dependent perturbation that vanishes for large positive and negative times. As an example, we apply the result (5.38) to the Hamiltonian (4.28), where we assume that the linearly coupled perturbation $b(t)\bar{z} + \bar{b}(t)z$ (thus coupled linearly to position and momentum) vanishes at $t \rightarrow \pm\infty$. After a short calculation, one finds

$$S(z, \bar{z}) = \exp \left[z\bar{z} + i \int_{-\infty}^{+\infty} dt (z e^{i\omega t} \bar{b}(t) + \bar{z} e^{-i\omega t} b(t)) \right. \\ \left. - \int_{-\infty}^{+\infty} dt_1 dt_2 \bar{b}(t_1) \theta(t_2 - t_1) e^{-i\omega(t_2-t_1)} b(t_2) \right]. \quad (5.39)$$

Using the formalism of Section 4.1, one verifies the unitarity of the S -matrix. Moreover, it is convenient to express the result in terms of the Fourier components of $b(t)$. Setting

$$b(t) = \int_{-\infty}^{+\infty} d\nu e^{i\nu t} \tilde{b}(\nu),$$

one obtains

$$S(z, \bar{z}) = \exp \left[z\bar{z} + 4\pi i \operatorname{Re}(\bar{z}\tilde{b}(\omega)) - 2\pi \int_{-\infty}^{+\infty} d\nu \frac{i}{\nu - \omega + i\varepsilon} |b(\nu)|^2 \right], \quad (5.40)$$

where $i\varepsilon$ means the limit of the expression when ε real positive goes to 0.

Coupling to position only. If the function $b(t)$ is real, the perturbation is coupled to the position $q(t)$ only. From equation (4.39), one infers that the coefficient of $b(t)$ is $\sqrt{2\omega}q(t)$. Then the expression (5.39) can be symmetrized in time and becomes

$$\begin{aligned} S(z, \bar{z}) = & \exp \left[z\bar{z} + i \int_{-\infty}^{+\infty} dt (z e^{i\omega t} + \bar{z} e^{-i\omega t}) b(t) \right. \\ & \left. - \frac{1}{2} \int_{-\infty}^{+\infty} dt_1 dt_2 b(t_1) e^{-i\omega|t_2-t_1|} b(t_2) \right]. \end{aligned} \quad (5.41)$$

In terms of Fourier components, one finds

$$S(z, \bar{z}) = \exp \left[z\bar{z} + 2\pi i (z\tilde{b}(-\omega) + \bar{z}\tilde{b}(\omega)) - \pi \int_{-\infty}^{+\infty} d\nu \frac{2i\omega}{\nu^2 - \omega^2 + i\varepsilon} |b(\nu)|^2 \right]. \quad (5.42)$$

It is interesting to compare this expression with a direct calculation of the path integral over real paths,

$$\mathcal{Z}(b) = \int [dq] \exp \left[i \int_{-\infty}^{+\infty} dt \left(\frac{1}{2} \dot{q}^2(t) - \frac{1}{2} \omega^2 q^2(t) + \sqrt{2\omega} b(t) q(t) \right) \right].$$

As such the path integral is undefined, because the classical equation of motion has non-trivial solutions. We thus define the real-time path integral as the analytic continuation of the Euclidean path integral. We perform a rotation in the time complex plane $t \mapsto t e^{i\theta}$, where θ varies between 0 (the Euclidean theory) and $\pi/2$, (the real-time theory). In the Fourier variable ν , the corresponding rotation is $\nu \mapsto \nu e^{-i\theta}$. Following the rotation, we find that this amounts to giving to ω^2 an infinitesimal negative imaginary part, $\omega^2 \mapsto \omega^2 - i\varepsilon$ with $\varepsilon \rightarrow 0_+$. The $i\varepsilon$ term then ensures the convergence of the Gaussian path integral and one finds

$$\mathcal{Z}(b) = \mathcal{Z}(0) \exp \left[-\frac{1}{2} \int_{-\infty}^{+\infty} dt_1 dt_2 b(t_1) e^{-i\omega|t_2-t_1|} b(t_2) \right].$$

We recognize the contribution quadratic in b in expression (5.41). Therefore, with this prescription, the S -matrix is given in terms of the modified path integral

$$\mathcal{Z}(b) = \int [dq] \exp i \int_{-\infty}^{+\infty} dt \left(\frac{1}{2} \dot{q}^2(t) - \frac{1}{2} \omega^2 q^2(t) + \sqrt{2\omega} b(t) (q(t) + q_0(t)) \right), \quad (5.43)$$

with

$$q_0(t) = (2\omega)^{-1/2} (z e^{i\omega t} + \bar{z} e^{-i\omega t}).$$

We note that the function $q_0(t)$ is the most general solution of the equation of motion of the classical harmonic oscillator. Then,

$$S(z, \bar{z}) = e^{z\bar{z}} \mathcal{Z}(b)/\mathcal{Z}(0).$$

One can also shift $q_0(t) + q(t) \mapsto q(t)$. Taking into account the equation of motion, one finds that the action in expression (5.43) can be written as

$$\mathcal{A}(q) = \int_{-\infty}^{+\infty} dt \left(\frac{1}{2} \dot{q}^2(t) - \frac{1}{2} \omega^2 q^2(t) - \frac{1}{2} \dot{q}_0^2(t) + \frac{1}{2} \omega^2 q_0^2(t) + \sqrt{2\omega} b(t) q(t) \right),$$

where the harmonic action of q_0 does not vanish because q_0 does not vanish asymptotically. It ensures the convergence of the time integral, because the function $q(t)$ now satisfies scattering boundary conditions, in agreement with general arguments.

5.5 The Bose gas: Evolution operator

The preceding formalism extends to the Bose gas discussed in Section 4.4.2. The evolution operator, in the formalism of the second quantization, and in the presence of a chemical potential μ coupled to the particle number \mathbf{N} , is given by a field integral continuation to real-time of the expression (4.59). One finds an expression of the form,

$$\langle \varphi'' | \mathbf{U}(t'', t') | \bar{\varphi}' \rangle = \langle \varphi'' | e^{-i(t''-t)(\mathbf{H}-\mu\mathbf{N})/\hbar} | \bar{\varphi}' \rangle = \int [d\bar{\varphi}(t, x) d\varphi(t, x)] e^{i\mathcal{A}(\varphi, \bar{\varphi})/\hbar}, \quad (5.44)$$

where the complex fields $\{\varphi(t, x), \bar{\varphi}(t, x)\}$ satisfy the boundary conditions

$$\bar{\varphi}(t, x') \equiv \bar{\varphi}'(x), \quad \varphi(t, x'') \equiv \varphi''(x).$$

In the example of an external potential V_1 and a pair potential V_2 , the action $\mathcal{A}(\varphi, \bar{\varphi})$, continuation of expression (4.60), becomes

$$\begin{aligned} \mathcal{A}(\varphi, \bar{\varphi}) = & -i\hbar \bar{\varphi}(t, x') \varphi(t, x') + \int dt d^d x \bar{\varphi}(t, x) \left(-i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla_x^2 - V_1(x) + \mu \right) \varphi(t, x) \\ & - \frac{1}{2} \int dt d^d x d^d y \bar{\varphi}(t, x) \varphi(t, x) V_2(x, y) \bar{\varphi}(t, y) \varphi(t, y). \end{aligned} \quad (5.45)$$

For $V_1(x) \equiv 0$ and for a pseudo-potential $V_2 = G\delta^{(d)}(x-y)$, the action becomes local:

$$\begin{aligned} \mathcal{A}(\varphi, \bar{\varphi}) = & -i\hbar \bar{\varphi}(t, x') \varphi(t, x') + \int dt d^d x \left[\bar{\varphi}(t, x) \left(-i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla_x^2 + \mu \right) \varphi(t, x) \right. \\ & \left. - \frac{1}{2} G (\bar{\varphi}(t, x) \varphi(t, x))^2 \right]. \end{aligned} \quad (5.46)$$

Note that below, but near the transition temperature of the Bose gas discussed in 4.4.7 (see also Section 15.10), the field φ is almost classical for small coupling. The field integral (5.44) in the limit $\hbar \rightarrow 0$ can then be evaluated by the stationary phase approximation, that is, by replacing the fields $\varphi, \bar{\varphi}$ by a solution of the classical field equations $\delta\mathcal{A}/\delta\varphi = \delta\mathcal{A}/\delta\bar{\varphi} = 0$. The evolution of the Bose gas is thus approximately described by a classical field equation

$$i\hbar \frac{\partial}{\partial t} \varphi(t, x) = \left(-\frac{\hbar^2}{2m} \nabla_x^2 + \mu - G\rho(t, x) \right) \varphi(t, x),$$

where $\rho(t, x)$ is the local condensate density:

$$\rho(t, x) = \bar{\varphi}(t, x) \varphi(t, x),$$

and $\bar{\varphi}$ and φ are complex conjugates. The equation has the form of a non-linear Schrödinger equation and is called the *Gross–Pitaevskii equation*.

5.6 Fermi gas: Evolution operator

With the conventions of Section 4.6, for a finite number of fermion states and a normal-ordered Hamiltonian of the form $h(\boldsymbol{\theta}, \partial/\partial\boldsymbol{\theta})$, the evolution operator is given by the path integral

$$\langle \boldsymbol{\theta}'' | U(t'', t') | \bar{\boldsymbol{\theta}}' \rangle = \int \prod_{\alpha} [d\theta_{\alpha}(t) d\bar{\theta}_{\alpha}(t)] \exp i\mathcal{A}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}), \quad (5.47)$$

with

$$\mathcal{A}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \int_{t'}^{t''} dt \left\{ i\bar{\boldsymbol{\theta}}(t) \cdot \dot{\boldsymbol{\theta}}(t) - h[\boldsymbol{\theta}(t), \bar{\boldsymbol{\theta}}(t)] \right\} + i\bar{\boldsymbol{\theta}}(t') \cdot \boldsymbol{\theta}(t'), \quad (5.48)$$

real-time continuation of the representation (4.110). We recall that Hermiticity now is equivalent to $\mathcal{A} = \mathcal{A}^{\dagger}$.

Note that, in the case of a free one-state Hamiltonian

$$h(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \omega \boldsymbol{\theta} \bar{\boldsymbol{\theta}} = -\omega \bar{\boldsymbol{\theta}} \boldsymbol{\theta},$$

if ω is negative, the empty state and the one-particle state have to be interchanged (Section 4.5.3). This is achieved by the transformation $\eta = \bar{\boldsymbol{\theta}}$, $\bar{\eta} = \boldsymbol{\theta}$. Then, after an integration by parts,

$$\mathcal{A}(\eta, \bar{\eta}) = \int_{t'}^{t''} dt [i\bar{\eta}(t)\dot{\eta}(t) + |\omega|\bar{\eta}(t)\eta(t)] - i\bar{\eta}(t'') \cdot \eta(t''),$$

and the action now corresponds to a one-particle state with positive energy.

The Fermi gas. The generalization of the formalism to fields makes it possible to describe the time evolution of the Fermi gas in the framework of second quantization, as introduced in Section 4.7. The evolution operator for the non-relativistic Fermi gas in the presence of a chemical potential μ coupled to the particle number \mathbf{N} , is given by a field integral continuation to real-time of expression (4.114). Here, one finds

$$\begin{aligned} \langle \varphi'' | \mathbf{U}(t'', t') | \bar{\varphi}' \rangle &= \langle \varphi'' | e^{-i(t''-t)(\mathbf{H}-\mu\mathbf{N})/\hbar} | \bar{\varphi}' \rangle \\ &= \int [d\bar{\varphi}(t, x) d\varphi(t, x)] \exp[i\mathcal{A}(\varphi, \bar{\varphi})/\hbar], \end{aligned} \quad (5.49)$$

where the fields $\{\varphi(t, x), \bar{\varphi}(t, x)\}$ are generators of a Grassmann algebra, and satisfy the boundary conditions

$$\bar{\varphi}(t, x') \equiv \bar{\varphi}'(x), \quad \varphi(t, x'') \equiv \varphi''(x).$$

In the example of an external potential V_1 and a pair potential V_2 , the action $\mathcal{A}(\varphi, \bar{\varphi})$ is the continuation of expression (4.115):

$$\begin{aligned} \mathcal{A}(\varphi, \bar{\varphi}) &= -i\hbar \bar{\varphi}(t, x') \varphi(t, x') + \int dt d^d x \bar{\varphi}(t, x) \left(-i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla_x^2 - V_1(x) + \mu \right) \varphi(t, x) \\ &\quad - \frac{1}{2} \int dt d^d x d^d y \bar{\varphi}(t, x) \varphi(t, x) V_2(x, y) \bar{\varphi}(t, y) \varphi(t, y). \end{aligned} \quad (5.50)$$

For example, in the absence of an external potential V_1 and for a pseudo-potential $V_2 = G\delta_{\alpha\beta}\delta^{(d)}(x-y)$, the action for an N -component Fermi field $\varphi^{\alpha}, \bar{\varphi}^{\alpha}$ becomes *local* and reads

$$\begin{aligned} \mathcal{A}(\varphi, \bar{\varphi}) &= -i\hbar \sum_{\alpha} \bar{\varphi}^{\alpha}(t, x') \varphi^{\alpha}(t, x') \\ &\quad + \int dt d^d x \left[\sum_{\alpha} \bar{\varphi}^{\alpha}(t, x) \left(-i\hbar \frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla_x^2 + \mu \right) \varphi^{\alpha}(t, x) - \frac{1}{2} G \left(\sum_{\alpha} \bar{\varphi}^{\alpha}(t, x) \varphi^{\alpha}(t, x) \right)^2 \right]. \end{aligned}$$

A5 Perturbation theory in the operator formalism

For completeness, and to illustrate the differences with the path integral formulation, we recall the basis of perturbation theory in the operator formalism of QM. To calculate the S -matrix, for example, we need an expression for the operator (see equation (5.3)):

$$\Omega(t) = e^{iH_0 t} e^{-iHt}, \quad (A5.1)$$

in which H_0 is the unperturbed Hamiltonian and

$$V = H - H_0$$

the perturbation.

The operator $\Omega(t)$ satisfies the equation

$$\dot{\Omega}(t) = -iV_I(t)\Omega(t), \quad (A5.2)$$

with the boundary condition

$$\Omega(0) = \mathbf{1},$$

and the operator $V_I(t)$ is the perturbation in the interaction representation [24],

$$V_I(t) = e^{iH_0 t} V e^{-iH_0 t}. \quad (A5.3)$$

One verifies that the solution of equation (A5.2) can be formally written as

$$\Omega(t) = \sum_{n=0}^{\infty} (-i)^n \int dt_1 dt_2 \cdots dt_n V_I(t_n) V_I(t_{n-1}) \cdots V_I(t_2) V_I(t_1), \quad (A5.4)$$

the domain of integration in the right-hand side being

$$0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq t.$$

If we now replace the product of perturbation terms that appears in the left-hand side by the time-ordered product, as defined in Section A2.3, the product becomes symmetric in the time arguments. Therefore, we can symmetrize the domain of integration provided we divide by a symmetry factor $n!$:

$$\Omega(t) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{\substack{0 \leq t_i \leq t \\ 1 \leq i \leq n}} dt_1 dt_2 \cdots dt_n T[V_I(t_1) V_I(t_2) \cdots V_I(t_n)]. \quad (A5.5)$$

This expression can be formally rewritten as

$$\Omega(t) = T \left[\exp \left(-i \int_0^t dt' V_I(t') \right) \right]. \quad (A5.6)$$

In particular, we can apply these results to a Hamiltonian H perturbed by a term linear in q . We can also write a path integral representation for the corresponding partition function. Comparing the expansion of the path integral in powers of the perturbation with expression (A5.5), we recover the relation between correlation functions and T-products established in Section 2.5 (see also Chapter 7 in Ref. [25]).

6 The neutral relativistic scalar field

With this chapter, we begin a discussion of relativistic quantum field theory (QFT). As the examples of quantum electrodynamics and general relativity show, a field theory provides a natural solution to the puzzle of the action at a distance, as generated by the Coulomb and Newton potentials, a concept which is, moreover, especially involved in a relativistic framework. Moreover, the known existence of the electromagnetic field, which pre-dates quantum mechanics (QM), made very early the idea of individual relativistic quantum particles unattractive, since it would not have unified fields and particles.

Using a field integral formalism, we construct a *local, relativistic-invariant* (that is, invariant under the pseudo-orthogonal group $O(1, d - 1)$, which preserves the Minkovski metric $(+1, -1, \dots, -1)$) for a self-interacting neutral scalar boson field, an extension of the non-relativistic QFT of the Bose gas in Section 4.4. Note that *locality* is a property that plays a essential role in most of this work. We postpone the study of relativistic fermions to Chapter 12, because we need to first introduce the spin group.

Using the holomorphic formalism, in a form that extends the construction of Section 4.4 to relativistic real-time evolution, we relate the field and particles. We obtain a representation of the S -matrix in terms of field integrals. After continuation to imaginary (or Euclidean) time, we apply the field-integral formalism to the partition function and Euclidean correlation functions. Conversely, we relate S -matrix elements to the continuation to real time of various kinds of correlation functions.

In Section 6.4, we argue that the massive ϕ^4 QFT has the quantum Bose gas with a pair potential, in the grand canonical formulation, described in Section 4.4.7, as a non-relativistic limit. We then discuss the field renormalization.

In this chapter, we set the speed c of light to $c = 1$, and $\hbar = 1$.

6.1 The relativistic scalar field

We consider a neutral (real), self-coupled, scalar boson field $\phi(t, x)$, which depends on time t and a $(d - 1)$ -dimensional space coordinate $x \equiv \{x_1, \dots, x_{d-1}\}$.

We construct a QFT by quantizing the relativistic classical Lagrangian density,

$$\mathcal{L}(\phi) = \frac{1}{2} (\partial_t \phi(t, x))^2 - \frac{1}{2} (\nabla_x \phi(t, x))^2 - V(\phi(t, x)), \quad x \in \mathbb{R}^{d-1}, \quad (6.1)$$

($\partial_t \equiv \partial/\partial t$, $\nabla_x \equiv (\partial/\partial x_1, \partial/\partial x_2, \dots, \partial/\partial x_{d-1})$), with the use of field integrals. Here, $V(\phi)$ is a function expandable in powers of ϕ , a polynomial in the simplest examples.

The Lagrangian density (6.1) is *local*: it only depends on the field $\phi(t, x)$ and its partial derivatives at the same space-time point. It is invariant under d -dimensional space-time translations, since space and time do not appear explicitly in the Lagrangian, relativistic invariant, that is, invariant under the group $O(1, d - 1)$, which, acting linearly on t and x , leaves the quadratic form $t^2 - x^2$ invariant.

To the Lagrangian (6.1) corresponds the classical action

$$\mathcal{A}(\phi) = \int dt d^{d-1}x \left\{ \frac{1}{2} \left[(\partial_t \phi(t, x))^2 - (\nabla_x \phi(t, x))^2 \right] - V(\phi(t, x)) \right\}. \quad (6.2)$$

An important example, which plays a fundamental role in the theory of critical phenomena (Section 15.3), and the Standard Model (SM) of particle physics (Chapter 23), is provided by the so-called ϕ^4 QFT, where $V(\phi)$ is a quartic, even potential:

$$V(\phi) = \frac{1}{2}r\phi^2 + \frac{1}{4!}g\phi^4. \quad (6.3)$$

For $g = 0$ (free-field theory), the physical mass of the particle associated with the field ϕ is \sqrt{r} . However, for $g > 0$, the parameter r that appears in the Lagrangian is expected to be negative: the existence of a continuous phase transition, necessary to generate a mass small enough, necessitates a double-well type potential, as many chapters will illustrate.

The Hamiltonian density, obtained by the Legendre transformation (see Section 1.8),

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_t\phi)}, \quad \mathcal{H}(\pi, \phi) = \pi(x)\partial_t\phi(t, x) - \mathcal{L},$$

where $\pi(x)$ is the conjugate momentum, is

$$\mathcal{H}(\pi, \phi) = \frac{1}{2}\pi^2(x) + \frac{1}{2}[\nabla_x\phi(x)]^2 + V(\phi(x)). \quad (6.4)$$

It is bounded from below if $V(\phi)$ is bounded from below. The total Hamiltonian is

$$\mathbf{H} = \int d^{d-1}x \mathcal{H}[\pi(x), \phi(x)]. \quad (6.5)$$

The action in the Hamiltonian (and phase space) formalism is (see equation (3.1)),

$$\mathcal{A}(\pi, \phi) = \int dt d^{d-1}x \left\{ \pi(t, x)\partial_t\phi(t, x) - \frac{1}{2}\pi^2(t, x) - \frac{1}{2}[\nabla_x\phi(t, x)]^2 - V(\phi(t, x)) \right\}. \quad (6.6)$$

The quantum Hamiltonian $\hat{\mathbf{H}}$ is obtained by replacing the classical fields $\{\pi, \phi\}$ by quantum operators $\{\hat{\pi}, \hat{\phi}\}$, which satisfy the commutation relations

$$[\hat{\pi}(x), \hat{\phi}(x')] = -i\delta^{d-1}(x - x'). \quad (6.7)$$

The properties of such a quantum theory can then be studied using the standard methods of operator QM. By contrast, we introduce a formalism of field integrals, which generalize the path integrals studied in Chapter 2, and are closely related to the field integrals introduced in Section 4.4.5. The positions q_i of particle mechanics are replaced here by the field $\phi(x)$. The transition from particle mechanics to field theory can be understood in much the same way as the transition from the discretized action (2.17) to the continuum time limit (2.18).

Evolution operator. Extending the arguments of Section 5.1.3 to field theory, we immediately obtain a *field integral* representation of the matrix elements of the evolution operator $U(t_2, t_1) = e^{-i(t_2-t_1)\hat{\mathbf{H}}}$:

$$\langle \phi_2 | U(t_2, t_1) | \phi_1 \rangle = \int [d\phi(t, x)] \exp[i\mathcal{A}(\phi)], \quad (6.8)$$

with the boundary conditions $\phi(t_1, x) = \phi_1(x)$, $\phi(t_2, x) = \phi_2(x)$, and where $\mathcal{A}(\phi)$ is the classical action (6.2), integrated between times t_1 and t_2 .

Note that, in Schrödinger's formulation of QM, wave functions become functionals of classical fields like $\phi_1(x)$ or $\phi_2(x)$ in expression (6.8), which correspond to an infinite number of usual variables.

6.1.1 Free-field theory and particle-field relation

First, we want to clarify the physical interpretation of relativistic fields and their relation with quantum particles. As in the example of the non-relativistic field theory of Section 4.4.2, the holomorphic formalism provides the proper tool to answer the question.

To explain the role of the holomorphic formalism in relativistic theory, it is useful to first briefly return to the construction of a free-field theory for a particle of mass m , which corresponds to the choice $V(\phi) = m^2\phi^2/2$. The phase-space action (6.6) becomes

$$\mathcal{A}_0(\pi, \phi) = \int dt d^{d-1}x \left\{ \pi(t, x) \partial_t \phi(t, x) - \frac{1}{2}\pi^2(t, x) - \frac{1}{2}[\nabla_x \phi(t, x)]^2 - \frac{1}{2}m^2\phi^2(t, x) \right\}. \quad (6.9)$$

Because the action is quadratic in the field, the field ϕ can be considered as a collection of harmonic oscillators. It is thus natural to introduce holomorphic formalism (see Section 4.2).

The $O(1, d - 1)$ invariant measure on the mass hyperboloid. We denote by p_0 the energy, p the $(d - 1)$ -component momentum, or by $\mathbf{p} \equiv (p_0, p)$ the energy-momentum, of a particle. The mass hyperboloid (or the mass shell) is defined by the $O(1, d - 1)$ invariant condition,

$$\mathbf{p}^2 \equiv p_0^2 - p^2 = m^2, \quad p_0 > 0, \quad (6.10)$$

where m is a mass. On the mass shell,

$$p_0 = \omega(p) \equiv \sqrt{p^2 + m^2}. \quad (6.11)$$

We introduce the $O(1, d - 1)$ invariant measure on the mass shell,

$$\delta_+(\mathbf{p}^2 - m^2) = \delta(\mathbf{p}^2 - m^2)\theta(p_0), \quad (6.12)$$

where $\theta(s)$ is the step function: $\theta(s) = 0$ for $s < 0$, $\theta(s) = 1$ for $s > 0$. Then, for any continuous function f of the energy-momentum \mathbf{p} ,

$$\int dp_0 d^{d-1}p \delta_+(\mathbf{p}^2 - m^2) f(\mathbf{p}) = \int \frac{d^{d-1}p}{2\omega(p)} f(\omega(p), p). \quad (6.13)$$

Covariant Fourier transformation and holomorphic representation. We introduce two fields $\varphi(t, p)$ and $\bar{\varphi}(t, p)$, generalizations of the complex functions $\bar{z}(t), z(t)$ of Section 4.2. We set,

$$\begin{aligned} \phi(t, x) &= \int \frac{d^{d-1}p}{2\omega(p)} [e^{ipx} \varphi(t, p) + e^{-ipx} \bar{\varphi}(t, p)], \\ \pi(t, x) &= i \int \frac{d^{d-1}p}{2\omega(p)} \omega(p) [e^{ipx} \varphi(t, p) - e^{-ipx} \bar{\varphi}(t, p)]. \end{aligned} \quad (6.14)$$

In the action (6.9), the different harmonic oscillators decouple in the momentum basis. The sign conventions ensure that $\bar{\varphi}$ and φ are complex conjugated when both π and ϕ are real.

In terms of $\varphi, \bar{\varphi}$, the free action (6.6) becomes

$$\mathcal{A}_0(\varphi, \bar{\varphi}) = -(2\pi)^{d-1} \int_{t'}^{t''} dt \frac{d^{d-1}p}{2\omega(p)} [i\bar{\varphi}(t, p) \partial_t \varphi(t, p) + \omega(p) \bar{\varphi}(t, p) \varphi(t, p)]. \quad (6.15)$$

This expression leads to a simple interpretation (see Sections 4.3, and 4.4) in the holomorphic representation: one-particle states are relativistic particles of mass m , momentum p and energy $p_0 = \omega(p)$. However, since the transformations (6.14) are not local in space, in the holomorphic formalism the *essential initial locality* of the action is no longer apparent.

6.1.2 Field integral and Fock's space

A representation of the matrix elements of the free evolution operator U_0 as a *field integral* in the holomorphic formalism follows:

$$U_0(t'', t'; \varphi'', \bar{\varphi}') = \int [\omega^{-1}(p) d\varphi(t, p) d\bar{\varphi}(t, p)] \exp [i\mathcal{A}_0(\varphi, \bar{\varphi})],$$

where $\mathcal{A}_0(\varphi, \bar{\varphi})$ is the action (6.15).

Fock's space. In the holomorphic formalism, the differences with the non-relativistic example discussed in Section 4.4.2 are mainly of a kinematic nature. However, for completeness, we briefly review the construction of the corresponding Fock's space. We work in the momentum representation, where the Hamiltonian is diagonal.

Let $\psi(p)$ be the wave function associated with a one-particle state of a particle of mass m . The scalar product of two states $|\Psi_1\rangle$, $|\Psi_2\rangle$ takes the relativistic covariant form

$$\langle \Psi_1 | \Psi_2 \rangle = (2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \psi_1^*(p) \psi_2(p). \quad (6.16)$$

We now introduce a complex field $\varphi(p)$ and the generating functional $\Psi(\varphi)$ of general n -particle wave functions for bosons,

$$\Psi(\varphi) = \sum_{n=0}^{\infty} \frac{(2\pi)^{n(d-1)}}{n!} \int \psi(p_1, p_2, \dots, p_n) \prod_{i=1}^n \frac{d^{d-1}p_i}{2\omega(p_i)} \varphi(p_i), \quad (6.17)$$

where the wave function $\psi(p_1, p_2, \dots, p_n)$ is totally symmetric in the momenta p_i . The direct sum of all n -particle spaces is called Fock space.

The scalar product of two vectors Ψ_1 and Ψ_2 takes the form of the field integral,

$$\langle \Psi_2 | \Psi_1 \rangle = \int \left[\frac{d\varphi(p) d\bar{\varphi}(p)}{\omega(p)} \right] \exp \left[-(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \bar{\varphi}(p) \varphi(p) \right] \overline{\Psi_2(\varphi)} \Psi_1(\varphi). \quad (6.18)$$

6.1.3 Hamiltonian and particle number operators

The free action (6.15) shows that the *one-particle Hamiltonian* has an energy spectrum of the form $\omega(p)$. Acting on $\Psi(\varphi)$, the free Hamiltonian is thus represented by the operator

$$\mathbf{H}_0 = \int d^{d-1}p \varphi(p) \omega(p) \frac{\delta}{\delta \varphi(p)} + E_0,$$

where E_0 is the zero-particle state (also called *vacuum*) and ground state energy.

The kernel representing the identity, which corresponds to the scalar product (6.18), is

$$\mathcal{I}(\varphi, \bar{\varphi}) = \exp \left[(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \bar{\varphi}(p) \varphi(p) \right]. \quad (6.19)$$

The kernel associated with the Hamiltonian follows:

$$\mathbf{H}_0 \mapsto \left[\frac{1}{2} (2\pi)^{d-1} \int d^{d-1}p \bar{\varphi}(p) \varphi(p) \right] \mathcal{I}(\varphi, \bar{\varphi}).$$

The free Hamiltonian commutes with the particle number operator

$$\mathbf{N} = \int d^{d-1}p \varphi(p) \frac{\delta}{\delta \varphi(p)} \mapsto (2\pi)^{d-1} \int \frac{d^{p-1}p}{2\omega(p)} \varphi(p) \bar{\varphi}(p) \mathcal{I}(\varphi, \bar{\varphi}) \Rightarrow [\mathbf{N}, \mathbf{H}_0] = 0,$$

a property that, in general, no longer holds in the presence of local interactions.

The vacuum energy. The vacuum state is the ground state $|0\rangle$ (in quantum bra–ket notation) of the Hamiltonian H_0 . The ground state or vacuum energy E_0 ,

$$\mathbf{H}_0 |0\rangle = E_0 |0\rangle,$$

in the usual quantization of the harmonic oscillator is formally given by the undefined expression

$$E_0 \propto \frac{1}{2} \int d^{d-1}p \omega(p).$$

To define more precisely E_0 , it is possible to quantize in a large box of linear size L , and to modify the theory at short distance or at large momenta in such a way that the Fourier modes are cut at some momentum scale Λ (a space lattice would provide such a cut-off). The Fourier variables p are then quantized:

$$p = p_n \equiv 2\pi n/L, \quad n \in \mathbb{Z}^{d-1},$$

and the vacuum energy becomes

$$E_0 = \frac{1}{2} \sum_n \omega(p_n).$$

For L large, sums can be replaced by integrals and $dn = L^{d-1}/(2\pi)^{d-1}dp$. The space volume factorizes, showing, as expected, that the energy is an extensive quantity:

$$E_0/L^{d-1} = \frac{1}{2} \int^\Lambda \frac{d^{d-1}p}{(2\pi)^{d-1}} \omega(p), \quad (6.20)$$

but the energy density is cut-off dependent. The large momentum divergence of the vacuum energy is not relevant here, because, in a non-gravitational theory, the Hamiltonian can always be shifted by a constant in such a way that the vacuum has zero energy (but this is no longer the case if the field theory is coupled to the gravitational field). However, even if the vacuum energy itself is not a physical observable, a variation (imposed for example by a change in boundary conditions) of the vacuum energy may be one (see Section A21.1).

6.1.4 Free-field two-point function

The ϕ -field two-point function, expressed as the expectation value of a time-ordered product of two fields (see, for example, Section 2.5 and Section A2.3), is given by

$$\langle 0 | T[\tilde{\phi}(t, p)\tilde{\phi}(0, p')] | 0 \rangle = \langle 0 | \tilde{\phi}(p) e^{-i\mathbf{H}_0|t|} \tilde{\phi}(p') | 0 \rangle = \frac{\delta^{d-1}(p + p')}{(2\pi)^{d-1}} \frac{e^{-i\omega(p)|t|}}{2\omega(p)}. \quad (6.21)$$

After Fourier transformation over time, one finds

$$\frac{1}{2\pi} \int e^{ip_0 t} dt \langle 0 | T[\tilde{\phi}(t, p)\tilde{\phi}(0, p')] | 0 \rangle = \frac{1}{(2\pi)^d} \delta^{d-1}(p + p') \frac{i}{p_0^2 - \omega^2(p) + i\varepsilon}, \quad (6.22)$$

with $\varepsilon \rightarrow 0_+$. The real-time two-point function is a distribution in the mathematical sense, boundary value of an analytic function:

$$\frac{i}{p_0^2 - \omega^2(p) + i\varepsilon} \equiv \frac{i}{\mathbf{p}^2 - m^2 + i\varepsilon} \equiv 2\pi\delta(\mathbf{p}^2 - m^2) + iPP \frac{1}{\mathbf{p}^2 - m^2},$$

where PP means principal part.

6.2 Quantum evolution and the S -matrix

Having determined the relation between fields and particles, we can now define the scattering S -matrix, and discuss some of its properties. The holomorphic formalism is especially useful in this respect. Various expressions for the scattering S -matrix follow.

6.2.1 The S -matrix: Scattering by an external source

We first calculate explicitly the scattering by an external source. The result will be the basis of a perturbative expression for the S -matrix in a general interacting theory.

To the action (6.2) where $V(\phi) = \frac{1}{2}m^2\phi^2$ (the free action),

$$\mathcal{A}_0(\phi) = \int dt d^{d-1}x \left\{ \frac{1}{2} \left[(\partial_t \phi(t, x))^2 - (\nabla_x \phi(t, x))^2 \right] - \frac{1}{2}m^2(\phi(t, x))^2 \right\}, \quad (6.23)$$

we add a source term which corresponds to the linear coupling of the field ϕ to an external classical source $J(t, x)$. The resulting action \mathcal{A}_G then takes the form

$$\mathcal{A}_G(\phi) = \mathcal{A}_0(\phi) - \int dt d^{d-1}x J(t, x)\phi(t, x).$$

In terms of the fields $\varphi, \bar{\varphi}$ (equation (6.14)), and the Fourier components of the source,

$$J(t, x) = \int d^{d-1}p e^{ipx} \tilde{J}(t, p),$$

the action becomes

$$\mathcal{A}_G(\varphi, \bar{\varphi}) = \mathcal{A}_0(\varphi, \bar{\varphi}) - (2\pi)^{d-1} \int dt \frac{d^{d-1}p}{2\omega(p)} \left[\tilde{J}(t, -p)\varphi(t, p) + \tilde{J}(t, p)\bar{\varphi}(t, p) \right],$$

where $\mathcal{A}_0(\varphi, \bar{\varphi})$ is the action (6.15).

A simple adaptation of expression (5.39) yields the holomorphic S -matrix

$$S_G(J, \varphi, \bar{\varphi}) = \exp \left[(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} K(p) \right],$$

with

$$\begin{aligned} K(p) &= \varphi(p)\bar{\varphi}(p) - i \int dt \left[\varphi(p) e^{i\omega(p)t} \tilde{J}(t, -p) + \bar{\varphi}(p) e^{-i\omega(p)t} \tilde{J}(t, p) \right] \\ &\quad - \frac{1}{2} \int dt_1 dt_2 \tilde{J}(t_1, -p) e^{-i|t_2-t_1|\omega(p)} \tilde{J}(t_2, p), \end{aligned}$$

where, in the last term, we have symmetrized in $p \mapsto -p$ and then $t_1 \leftrightarrow t_2$.

As explained in Section 5.4.1, the expansion of this functional in powers of φ and $\bar{\varphi}$ then yields the coefficients of the scattering matrix.

A more convenient expression is obtained by introducing the time Fourier components of $\tilde{J}(t, p)$,

$$\tilde{J}(t, p) = \int dp_0 e^{-ip_0 t} \tilde{J}(p_0, p).$$

Then,

$$\begin{aligned} \ln S_G(J, \varphi, \bar{\varphi}) &= (2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \varphi(p) \bar{\varphi}(p) - i(2\pi)^d \int dp_0 d^{d-1}p \tilde{J}(p_0, p) \\ &\quad \times [\delta_+(p_0^2 - p^2 - m^2)\varphi(-p) + \delta_-(p_0^2 - p^2 - m^2)\bar{\varphi}(p)] \\ &\quad - \frac{1}{2}(2\pi)^d \int dp_0 d^{d-1}p \tilde{J}(-p_0, -p) \tilde{\Delta}(p_0, p) \tilde{J}(p_0, p), \end{aligned} \quad (6.24)$$

where the notation (6.12) and

$$\delta_-(p_0^2 - p^2 - m^2) = \theta(-p_0)\delta(p_0^2 - p^2 - m^2),$$

have been used, and

$$\tilde{\Delta}(p_0, p) = \frac{i}{p_0^2 - p^2 - m^2 + i\varepsilon} \equiv \frac{i}{\mathbf{p}^2 - m^2 + i\varepsilon}. \quad (6.25)$$

In the coefficient of the term quadratic in J , we recognize the free two-point function (6.22).

6.2.2 General interacting theory

An interaction term $V_I(\phi)$ can then be added to the free action, where ϕ has to be expressed in terms of $\varphi, \bar{\varphi}$,

$$\mathcal{A}(\phi) = \int dt d^{d-1}x \left[\frac{1}{2} (\partial_t \phi(t, x))^2 - \frac{1}{2} (\nabla_x \phi(t, x))^2 - \frac{1}{2} m^2 \phi^2(t, x) \right] - V_I(\phi). \quad (6.26)$$

Using one of the functional expressions for the perturbative expansion that we derive in Section 7.2.2, we obtain the form of the S -matrix for the interacting theory:

$$S(\varphi, \bar{\varphi}) = \exp \left[-iV_I \left(\frac{1}{i} \frac{\delta}{\delta J} \right) \right] S_G(J, \varphi, \bar{\varphi}) \Big|_{J=0}. \quad (6.27)$$

The S -matrix thus has a Feynman diagram expansion with internal propagators $\tilde{\Delta}$, as given by expression (6.25) (see also Section 7.2.2). We note that we have indeed obtained a propagator, which otherwise would be singular on the *mass shell* $\mathbf{p}^2 = p_0^2 - p^2 = m^2$, with the well-known $i\varepsilon$ prescription.

Unitarity. With our conventions, the unitarity of the S -matrix takes the functional form

$$\begin{aligned} &\int [d\bar{\varphi}'(p) d\varphi'(p)] S^*(\varphi', \bar{\varphi}) S(\varphi', \bar{\varphi}) \exp \left[-(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \varphi'(p) \bar{\varphi}'(p) \right] \\ &= \exp \left[(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \varphi(p) \bar{\varphi}(p) \right]. \end{aligned} \quad (6.28)$$

Eigenstates. We have constructed our basis of states from the eigenstates of the unperturbed Hamiltonian. More generally, we can take another harmonic oscillator basis corresponding to a different mass, at the price of adding quadratic terms in the field to the interaction. Actually, and this will become clearer when we discuss the structure of the ground state in field theory, if we take an arbitrary basis, in general, all eigenstates of the interacting Hamiltonian will be orthogonal to all vectors of the basis (a property specific to systems with an infinite number of degrees of freedom, see Chapter 14).

Moreover, the Hamiltonian of a massive theory has a unique, translation-invariant, lowest-energy excited state (in the case of several fields this can be generalized to all super-selection sectors). The physical mass m (or inverse correlation length in the statistical language) is defined as the energy of this state. This defines the zero-momentum one-particle state. A more general one-particle state is obtained by boosting the zero-momentum state, that is, performing a $O(1, d - 1)$ transformation, and creating a one-particle state of momentum p and energy $\omega(p)$. Additional eigenstates have energies at least equal to $2m$.

Therefore, we have to take the true ground state of the complete Hamiltonian as vacuum state, and as asymptotic free states, free particles with the true physical mass. These conditions implicitly define a reference free theory with action \mathcal{A}_0 and ensure that it describes the asymptotic states at large times.

The vacuum state and the physical mass can be calculated in perturbation theory. To calculate scattering amplitudes, one has to perform order by order *field and mass renormalizations*, which involves, in particular, taking the physical mass as a parameter of the perturbative expansion by inverting the relation between the physical mass, as defined by the pole of the two-point function, and the coefficient of ϕ^2 , as it appears in the action.

Finally, the field integral has to be normalized by the condition $S(0, 0) = 1$, which means that we divide by a factor related to difference in energies between the true and the unperturbed ground state.

6.3 *S*-matrix and field asymptotic conditions

Since the action is only local when written in terms of the initial real field $\phi(x)$, it is useful to derive an expression of the *S*-matrix in the ϕ formalism. We know how to calculate the matrix elements of the evolution operator by integrating over the field ϕ . We now compare this expression of the evolution operator with the explicit form (6.27) of the *S*-matrix as derived from the holomorphic representation.

6.3.1 The Gaussian integral in an external source and the *S*-matrix

We first consider the Gaussian theory in an external source,

$$\mathcal{Z}_G(J) = \int [d\phi] \exp \left[i\mathcal{A}_0(\phi) + i \int dt d^{d-1}x J(t, x)\phi(t, x) \right], \quad (6.29)$$

where is the action (6.23). The action can be expressed in terms of the Fourier components of the field ϕ and the source J ,

$$\phi(t, x) = \int dp_0 d^{d-1}p e^{ipx - ip_0 t} \tilde{\phi}(p_0, p), \quad J(t, x) = \int dp_0 d^{d-1}p e^{ipx - ip_0 t} \tilde{J}(p_0, p).$$

The total action becomes

$$\mathcal{A}(\phi, J) = (2\pi)^d \int d^d \mathbf{p} \left[\frac{1}{2} \tilde{\phi}(-\mathbf{p}) (\mathbf{p}^2 - m^2) \tilde{\phi}(\mathbf{p}) + \tilde{J}(-\mathbf{p}) \tilde{\phi}(\mathbf{p}) \right]. \quad (6.30)$$

Unlike the Euclidean field integral (equation (6.44)), the field integral for the evolution operator has convergence problems because classical field equations have non-trivial solutions. This problem has already been discussed in Section 5.4.2, and we use the same strategy here.

We define the field integral as the analytical continuation in the time of the Euclidean field integral. We perform a rotation in the time complex plane, $t \mapsto t e^{i\theta}$, where θ varies between 0 (the Euclidean theory) and $\pi/2$, (the Minkowski theory). In the energy variable p_0 , the corresponding rotation is $p_0 \mapsto p_0 e^{-i\theta}$. As we have explained, this amounts to adding an infinitesimal negative imaginary part to m^2 which ensures the convergence of the field integral (6.29) for large fields. The generating functional $\mathcal{Z}_G(J)$ can then be calculated, and one finds

$$\ln \mathcal{Z}_G(J) = -\frac{1}{2}(2\pi)^d \int d^d \mathbf{p} \tilde{J}(\mathbf{p}) \tilde{\Delta}(\mathbf{p}) \tilde{J}(-\mathbf{p}), \quad (6.31)$$

where $\tilde{\Delta}(\mathbf{p})$ is the free propagator,

$$\tilde{\Delta}(\mathbf{p}) = \frac{i}{\mathbf{p}^2 - m^2 + i\varepsilon}. \quad (6.32)$$

We note that the propagator obtained by this prescription is identical to the internal propagator (6.25) that appears in the Feynman graph expansion of the S -matrix. The analytic continuation leads to an $i\varepsilon$ rule for real-time Feynman diagrams.

S-matrix in an external source. Comparing the expressions (6.31) and (6.24), we note that the quadratic term is reproduced, but not the term linear in the source. This term corresponds to an addition to the field $\tilde{\phi}$:

$$\tilde{\phi}(p_0, p) \mapsto \tilde{\phi}(p_0, p) + \tilde{\phi}_0(p_0, p),$$

with

$$\tilde{\phi}_0(p_0, p) = \delta(p_0^2 - p^2 - m^2) [\varphi(p)\theta(-p_0) + \bar{\varphi}(-p)\theta(p_0)]. \quad (6.33)$$

The additional term $\tilde{\phi}_0(p_0, p)$ is a general solution of the free classical field equation

$$(p_0^2 - p^2 - m^2)\tilde{\phi}_0(p_0, p) = 0,$$

in a specific parametrization. It is non-vanishing only on the mass-hyperboloid $\mathbf{p}^2 = m^2$. The parametrization (6.33) of the solution reflects the property that the mass hyperboloid has two disconnected components, depending on the sign of the energy p_0 .

As we have noted in Section 5.4.2, we can shift the field ϕ , taking $\phi + \phi_0$ as the integration variable. The shifted field ϕ then satisfies scattering boundary conditions and the S -matrix can thus be derived from the field integral

$$\mathcal{Z}_G(J) = \int [d\phi] \exp \left[i\mathcal{A}_0(\phi) - i\mathcal{A}_0(\phi_0) + i \int dt d^{d-1}x J(t, x)\phi(t, x) \right]. \quad (6.34)$$

The interpretation is the following: the field ϕ in the functional integral (6.29) satisfies general free-field boundary conditions $\phi \rightarrow \phi_0$, and its two-point function or propagator is then given by equation (6.32).

General interaction. The general functional representation (6.27) can then be rewritten in a different way. Introducing the form (6.34) of Z_G in equation (6.27), and applying functional derivatives, one finds

$$S(\varphi, \bar{\varphi}) = \mathcal{I}(\varphi, \bar{\varphi}) \int [d\phi] e^{i[\mathcal{A}(\phi) - \mathcal{A}_0(\phi_0)]}, \quad (6.35)$$

where \mathcal{I} is the identity kernel (6.19), and the field ϕ satisfies fixed free-field boundary conditions [26, 27].

The expression (6.35), up to the factor \mathcal{I} , is a field integral in the presence of a background field ϕ_0 . This field integral differs from the vacuum amplitude only in the boundary conditions, which are free-field boundary conditions. This result is consistent with the analysis of Section 5.3.1. We have shown that, in QM, S -matrix elements can be calculated from the path integral representation of the evolution operator, by integrating over paths which satisfy prescribed classical scattering boundary conditions, that is, which correspond to asymptotic free classical motion. In particular, the starting point of the semi-classical expansion is a classical scattering trajectory. The arguments can be generalized to quantum field theory with massive particles (to ensure proper cluster properties and thus the existence of an S -matrix).

Solitons. Considerations based on asymptotic field boundary conditions, or the more direct considerations of Section 6.1.1, lead to the same perturbative S -matrix. However, as suggested by the discussion given at the beginning of Section 5.3.1, the preceding considerations generalize to the scattering of *solitons*, that is, states obtained by expanding the field integral around finite energy static solutions of the complete classical field equations

$$\frac{\delta \mathcal{A}(\phi)}{\delta \phi(x)} = 0. \quad (6.36)$$

In this case, the S -matrix of soliton scattering is obtained by expanding the field integral around classical-scattering solutions of soliton type of the complete field equations [21].

6.3.2 S -matrix elements from correlation functions

This section somewhat anticipates some concepts and results that are presented in Chapters 7, 8, and 9, but otherwise contains material which has, more naturally, its place in this chapter.

We consider, for a general interaction, the expression

$$\mathcal{Z}(\mathcal{J}) = \exp \left[-iV_1 \left(\frac{1}{i} \frac{\delta}{\delta J} \right) \right] \exp \left[-\frac{1}{2} (\mathcal{J} + J) \Delta (\mathcal{J} + J) \right] \Big|_{J=0},$$

where \mathcal{J} is an additional external source and, in the Fourier representation,

$$\begin{aligned} (\mathcal{J} + J) \Delta (\mathcal{J} + J) &\equiv \int d^d p \tilde{J}(-\mathbf{p}) \frac{i(2\pi)^d}{\mathbf{p}^2 - m^2 + i\varepsilon} \tilde{J}(\mathbf{p}) + 2 \int d^d p \tilde{J}(-\mathbf{p}) \frac{i(2\pi)^d}{\mathbf{p}^2 - m^2 + i\varepsilon} \tilde{J}(\mathbf{p}) \\ &+ \int d^d p \tilde{J}(-\mathbf{p}) \frac{i(2\pi)^d}{\mathbf{p}^2 - m^2 + i\varepsilon} \tilde{J}(\mathbf{p}). \end{aligned} \quad (6.37)$$

Comparing the expression with expression (6.24), we note that the main formal difference between expressions (6.31) and (6.27) comes from the term linear in J :

$$\frac{i}{\mathbf{p}^2 - m^2 + i\varepsilon} \tilde{J}(\mathbf{p}) \mapsto -i\delta(\mathbf{p}^2 - m^2) [\varphi(\mathbf{p})\theta(-p_0) + \bar{\varphi}(-\mathbf{p})\theta(p_0)]. \quad (6.38)$$

Therefore, we conclude that S -matrix elements can be obtained from real-time correlation functions by first multiplying them by the product of external inverse propagators, and then by restricting the external momenta to the *mass shell* $\mathbf{p}^2 = m^2$. This does not imply that the result vanishes. Indeed, correlation functions have poles on the mass shell. The final answer is proportional to the mass-shell so-called *amputated correlation functions* (defined by equation (7.71)).

Disconnected contributions. The relation between correlation functions and S -matrix elements shows that the matrix elements, as defined here, have disconnected contributions. By contrast, the new functional

$$\mathcal{T}(\varphi, \bar{\varphi}) = i \ln S(\varphi, \bar{\varphi}), \quad (6.39)$$

is the generating functional of connected scattering amplitudes (see Section 7.3).

Crossing symmetry. We see that only a linear combination of φ and $\bar{\varphi}$ appears, with an identical coefficient, up to the sign of the energy. The sign of p_0 specifies the incoming and outgoing particles. This has deep implications, specific to relativistic QFT: in $d > 2$ dimensions, the same analytic functions lead to scattering amplitudes of different physical processes, a property known as *crossing symmetry*.

6.3.3 The ϕ^3 example: Tree approximation

We illustrate the analysis by calculating a four-point scattering amplitude in the simple, non-physical, ϕ^3 field theory, in the tree approximation. We consider the real-time action,

$$\mathcal{A}(\phi) = \int dt d^{d-1}x \left[\frac{1}{2} (\partial_t \phi(t, x))^2 - \frac{1}{2} (\nabla_x \phi(t, x))^2 - \frac{1}{2} m^2 \phi^2(t, x) - \frac{1}{3!} g \phi^3(t, x) \right].$$

We introduce the Fourier components of the field ϕ ($\mathbf{p} \equiv (p_0, \mathbf{p})$),

$$\phi(t, x) = \int dp_0 d^{d-1}p e^{-ip_0 t + ip \cdot x} \tilde{\phi}(\mathbf{p}). \quad (6.40)$$

In terms of Fourier components, the classical field equation takes the form

$$(\mathbf{p}^2 - m^2) \tilde{\phi}(\mathbf{p}) - \frac{1}{2} g \int d^d \mathbf{q} \tilde{\phi}(\mathbf{q}) \tilde{\phi}(\mathbf{p} - \mathbf{q}) = 0. \quad (6.41)$$

The equation can be solved as a series in the coupling constant g starting from the solution (6.33) of the free equation,

$$\tilde{\phi}_0(\mathbf{p}) = \delta(\mathbf{p}^2 - m^2) [\varphi(\mathbf{p}) \theta(-p_0) + \bar{\varphi}(-\mathbf{p}) \theta(p_0)].$$

The classical solution at order g is

$$\tilde{\phi}(\mathbf{p}) = \tilde{\phi}_0(\mathbf{p}) + \frac{1}{2} \frac{g}{\mathbf{p}^2 - m^2} \int d^d \mathbf{q} \tilde{\phi}_0(\mathbf{q}) \tilde{\phi}_0(\mathbf{p} - \mathbf{q}) + O(g^2).$$

By looking for perturbative solutions of the field equation, we have explicitly excluded scattering states corresponding to bound states or solitons (see the remark at the end of Section 6.3.1).

Using equation (6.41), we can rewrite \mathcal{T} in the tree approximation (see also Section 7.9.1) as

$$\mathcal{T}(\phi_0) = i \ln S(\phi_0) = -\frac{1}{12} g (2\pi)^d \int d^d \mathbf{p}_1 d^d \mathbf{p}_2 d^d \mathbf{p}_3 \delta^{(d)}(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \tilde{\phi}(\mathbf{p}_1) \tilde{\phi}(\mathbf{p}_2) \tilde{\phi}(\mathbf{p}_3).$$

We then replace ϕ by its expansion in powers of g . The term of order g , which would describe one ϕ particle decaying into two, vanishes by energy conservation. The next term of order g^2 has the form

$$-\frac{1}{8}g^2(2\pi)^d \int \delta^{(d)}(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_4) \frac{1}{(\mathbf{p}_1 + \mathbf{p}_2)^2 - m^2} \prod_{i=1}^4 d^d \mathbf{p}_i \tilde{\phi}_0(\mathbf{p}_i).$$

The connected four-particle scattering amplitude is then obtained by differentiating with respect to $\tilde{\phi}_0(\mathbf{p})$. The result is the product of a factor that contains the energy-momentum conservation,

$$(2\pi)^d \delta^{(d)}(\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_4),$$

and an amplitude,

$$-\frac{g^2}{(\mathbf{p}_1 + \mathbf{p}_2)^2 - m^2} - \frac{g^2}{(\mathbf{p}_1 + \mathbf{p}_3)^2 - m^2} - \frac{g^2}{(\mathbf{p}_1 + \mathbf{p}_4)^2 - m^2},$$

where $\mathbf{p}_i^2 = m^2$. The term we have calculated also contains, in principle, the decay of one particle into three, but again this process vanishes by energy conservation.

Higher orders in g yield 5, 6, and so on, particle scattering amplitudes.

6.4 The non-relativistic limit: The ϕ^4 QFT

Although relativistic QFT appears as a natural formal extension of non-relativistic QM, it is not quite obvious how non-relativistic QM is recovered in the non-relativistic limit.

Therefore, in the example of the *massive* relativistic ϕ^4 QFT, we briefly explain how non-relativistic QM emerges [28], in the low-momentum, compared to the mass, and large time, compared to the inverse mass, limits. We show that the limit takes the form of a quantum many-body model, of the kind discussed in Section 4.4.2. Since a careful discussion of the non-relativistic limit of QFT would be somewhat involved, we consider here only one example and give intuitive, non-rigorous arguments, which illustrate the main point.

The ϕ^4 QFT. We consider the *massive* scalar field theory, in $(d-1)$ space dimensions, with a ϕ^4 type interaction, corresponding to the action (6.2) with the potential (6.3),

$$\begin{aligned} \mathcal{A}(\phi) = & \int dt d^{d-1}x \left[\frac{1}{2} (\partial_t \phi(t, x))^2 - \frac{1}{2} (\nabla_x \phi(t, x))^2 - \frac{1}{2} m^2 \phi^2(t, x) - \frac{1}{2} (r - m^2) \phi^2(t, x) \right. \\ & \left. - \frac{1}{4!} g \phi^4(t, x) \right], \quad g > 0. \end{aligned} \quad (6.42)$$

The parameter $r(m, g)$ is adjusted in order to generate a particle of physical mass m .

The matrix elements of the evolution operator for the field ϕ , are given by equation (6.8). At least for a coupling g weak enough, the field integral is dominated by fields satisfying the free-field equation

$$(\partial_t^2 - \nabla_x^2 + m^2) \phi(t, x) = 0.$$

In the non-relativistic limit, the space variation is small compared to the time variation. If space variations are completely neglected, the solutions to the classical field equation simply reduce to $\phi(t, x) \propto e^{\pm i m t}$.

It is thus natural to introduce the holomorphic representation of fields, taking as the unperturbed harmonic oscillator action

$$\mathcal{A}_0(\phi) = \int dt d^{d-1}x \left[\frac{1}{2} (\partial_t \phi(t, x))^2 - \frac{1}{2} m^2 \phi^2(t, x) \right].$$

Denoting by $\varphi(t, x), \bar{\varphi}(t, x)$ the complex fields, in terms of which the field $\phi(t, x)$ reads

$$\phi(t, x) = (2m)^{-1/2} (\varphi(t, x) + \bar{\varphi}(t, x)),$$

we find the action

$$\begin{aligned} \mathcal{A}(\varphi, \bar{\varphi}) = & - \int dt d^{d-1}x \left\{ \bar{\varphi}(t, x) (i\partial_t + m) \varphi(t, x) + \frac{1}{4m} [\nabla_x (\varphi(t, x) + \bar{\varphi}(t, x))]^2 \right. \\ & \left. + \frac{1}{4m} (r - m^2) (\varphi(t, x) + \bar{\varphi}(t, x))^2 + \frac{g}{96m^2} (\varphi(t, x) + \bar{\varphi}(t, x))^4 \right\}. \end{aligned}$$

To separate fast and low time frequencies, we define the new fields

$$\varphi(t, x) \mapsto e^{imt} \varphi(t, x), \quad \bar{\varphi}(t, x) \mapsto e^{-imt} \bar{\varphi}(t, x),$$

where the new fields $\bar{\varphi}, \varphi$ have slow time variations compared to the factors e^{imt} . After the transformation, the monomials of the form $\bar{\varphi}^r \varphi^s$ are multiplied by a factor $e^{im(s-r)t}$. For $r \neq s$, the corresponding time integrals give small contributions due to the rapid time oscillations (note the similarity with the discussion of Section 4.7.2). Hence, at leading order, the only surviving terms are those that have an equal number of $\bar{\varphi}$ and φ factors. The limiting non-relativistic action is then

$$\begin{aligned} \mathcal{A}(\bar{\varphi}, \varphi) = & - \int dt d^{d-1}x \left[i\bar{\varphi}(t, x) \partial_t \varphi(t, x) + \frac{1}{2m} \nabla_x \bar{\varphi}(t, x) \nabla_x \varphi(t, x) \right. \\ & \left. + \frac{1}{2m} (r - m^2) \bar{\varphi}(t, x) \varphi(t, x) + \frac{g}{16m^2} (\bar{\varphi}(t, x) \varphi(t, x))^2 \right]. \end{aligned} \quad (6.43)$$

We recognize a real-time action written in terms of complex fields of the form (4.70). As expected, the Hamiltonian in the non-relativistic limit commutes with the particle number. This property, in general, is shared in relativistic QFT only by free Hamiltonians. By contrast, in the non-relativistic limit of a *massive theory*, all momenta are small compared to masses and, therefore, the number of particles is necessarily conserved.

Up to an infinite energy shift (the vacuum energy), the n -particle Hamiltonian H_n has the form

$$H_n = -\frac{1}{2m} \sum_{i=1}^n \nabla_{x_i}^2 + \frac{g}{8m^2} \sum_{i < j \leq n} \delta^{(d-1)}(x_i - x_j).$$

This is an n -particle Hamiltonian with a two-body contact repulsive potential. The additional $\varphi \bar{\varphi}$ term in equation (6.43) contributes only to the chemical potential.

The important conclusion of this analysis is that the low-energy, non-relativistic limit of relativistic QFT is many-body QM. The interaction being repulsive, we do not expect bound states even in the relativistic theory.

6.5 Quantum statistical physics

Since this Hamiltonian is quadratic in the momentum $\pi(x)$, it is a continuum generalization of Hamiltonians of the form $\sum_i p_i^2/2m + V(q)$. Following the method explained in Section 2.2, we thus immediately obtain a field integral representation of the matrix elements of the statistical operator $U(t_2, t_1) = e^{-(t_2 - t_1)\mathbf{H}}$:

$$\langle \phi_2 | U(t_2, t_1) | \phi_1 \rangle = \int [d\phi(t, x)] \exp [-\mathcal{S}(\phi)], \quad (6.44)$$

with the boundary conditions $\phi(t_1, x) = \phi_1(x)$, $\phi(t_2, x) = \phi_2(x)$, and where $\mathcal{S}(\phi)$ is the classical Euclidean action, derived by analytic continuation to Euclidean (imaginary) time (*i.e.*, $t \mapsto -it$) from the action in real time (6.2) (see Section 2.2), and integrated between times t_1 and t_2 :

$$\mathcal{S}(\phi) = \int_{t_1}^{t_2} dt \int d^{d-1}x \left\{ \frac{1}{2} \left[(\partial_t \phi(t, x))^2 + (\nabla_x \phi(t, x))^2 \right] + V(\phi(t, x)) \right\}. \quad (6.45)$$

One advantage of the Euclidean-time formulation, obvious in expression (6.45), is that time and space now play an equivalent role (except for possible boundary conditions). In particular, the initial non-compact $SO(1, d-1)$ pseudo-orthogonal symmetry, is replaced by the compact $SO(d)$ orthogonal symmetry.

Therefore, in what follows, we set $t \equiv x_0$, and $\mathbf{x} \equiv (x_0, \dots, x_{d-1})$.

6.5.1 Partition function and correlation functions

The quantum partition function. The quantum partition function $\mathcal{Z}(\beta)$, where β is the inverse temperature, is obtained by

$$\mathcal{Z}(\beta) = \text{tr } e^{-\beta \mathbf{H}} = \int [d\phi(x)] \exp [-\mathcal{S}(\phi)], \quad (6.46)$$

where the fields satisfy the periodic boundary conditions, in the Euclidean time direction, $\phi(\beta, x_1, \dots, x_{d-1}) = \phi(0, x_1, \dots, x_{d-1})$, and

$$\mathcal{S}(\phi) = \int_{x_0=0}^{x_0=\beta} d^d x \left[\frac{1}{2} (\nabla_x \phi(\mathbf{x}))^2 + V(\phi(\mathbf{x})) \right], \quad (6.47)$$

where $\nabla_x \equiv (\partial/\partial x_0, \partial/\partial x_1, \dots, \partial/\partial x_{d-1})$. The expression can be compared to the non-relativistic form (4.61).

Correlation functions. The integrand of the field integral (6.46) defines a positive measure for a classical statistical field theory (see Section 2.5). To the measure correspond correlation functions defined by ($x^j \equiv \{x_0^j, x_1^j, \dots, x_{d-1}^j\}$)

$$\langle \phi(\mathbf{x}_1) \cdots \phi(\mathbf{x}_n) \rangle = \mathcal{Z}^{-1}(\beta) \int [d\phi] \phi(\mathbf{x}_1) \cdots \phi(\mathbf{x}_n) e^{-\mathcal{S}(\phi)}, \quad (6.48)$$

in which $\mathcal{Z}(\beta)$ is the quantum partition function (6.46). Equal imaginary-time correlation functions are also the static correlation functions of the finite temperature QFT (see Chapter 33). To obtain the time-dependent quantum correlation functions, one has to proceed by analytic continuation towards real time in the argument of ϕ but not in β .

In the infinite β limit, these finite temperature correlation functions become the ground state (usually called *vacuum* in QFT) expectation values of time-ordered products of quantum field operators (see Section A2.3). After analytic continuation towards real time, they yield the Green's functions from which elements of the scattering S -matrix can be calculated (Section 6.3.2).

Zero temperature. Since finite temperature QFT will be discussed extensively in Chapter 33, we work in the remaining part of the chapter directly at zero temperature in infinite \mathbb{R}^{d-1} space and, therefore, in infinite \mathbb{R}^d Euclidean space: correlation functions then correspond to quantum vacuum expectation values. The Euclidean action is fully $O(d)$ -invariant, and one no longer needs to distinguish space and time; \mathbf{x} belongs to \mathbb{R}^d .

6.5.2 The problem of infinities or ultraviolet divergences

We have already seen (Sections 3.2.2, 3.3) that path integrals reduced to their formal definition in continuous time are sometimes undefined. In such cases it is possible, for example, to complete the formal definition with a limiting procedure based on time-discretization.

In Chapter 8, we show that in a local QFT, the problem is even more severe. With an action of the form (6.45), the theory expanded in powers of the interaction V (the perturbative expansion), is affected by short distance singularities or, equivalently, in a Fourier representation, by large momentum (ultraviolet) divergences.

These divergences show that QFT reduced to its straightforward definition is an incomplete theory that has, eventually, to be embedded into some non-local theory but whose non-locality is restricted to very short space-times. QFT is only an *effective theory* to describe large-scale physics. Since the embedding theory is unknown or too complicated, to define a relativistic local QFT, one modifies the theory at short distance or large momentum, a process called *regularization*.

An example of regularization is provided by a lattice regularization and this establishes a relation between QFT and the theory of continuous phase transitions in lattice models (see Chapter 15). Since such modifications are to a large extent arbitrary, it is then necessary to show that the physical properties of the theory at large distance or small momentum are independent of the modifications. This is the purpose of the *renormalization theory* and the interpretation of this short distance insensitivity is based on *renormalization group (RG)* arguments.

6.5.3 Connected and vertex functions

Connected correlation functions. By comparing the general form of the S -matrix with correlation functions, we have seen in Section 6.3.2 that connected S -matrix elements are related to the analytic continuation to real time ($t \equiv p_0 \mapsto it \equiv ip_0$) of Euclidean correlation functions in the mass-shell limit. More precisely, the relation between $\mathcal{W}(J)$, the generating functional of connected correlation functions (equation (7.19)), and the generating functional of connected S -matrix elements is

$$\mathcal{T}(\phi_0) = i\mathcal{W}(\Delta^{-1}J),$$

where J is inferred from equation (6.38), ϕ_0 is given by equation (6.33) and Δ is the propagator with a physical pole. *Connected correlation functions* $W^{(n)}$ can then be expressed in terms of *amputated functions* (inverting relation (7.71)).

In the Fourier representation, the relation is

$$\tilde{W}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) = \left[\prod_{i=1}^n \tilde{W}^{(2)}(\mathbf{p}_i) \right] \tilde{W}_{\text{amp.}}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n). \quad (6.49)$$

The Euclidean two-point function $\tilde{W}^{(2)}(\mathbf{p})$ has a pole located at $\mathbf{p}^2 = -m^2$, where m is the physical mass.

Near the pole

$$\tilde{W}^{(2)}(\mathbf{p}) \underset{\mathbf{p}^2 \rightarrow -m^2}{\sim} \frac{Z}{\mathbf{p}^2 + m^2}, \quad (6.50)$$

where Z is the field renormalization constant (Section 6.6).

We conclude that the coefficient $\mathcal{T}_r^{(n)}$ of $\mathcal{T}(\phi_0/\sqrt{Z})$ in the expansion in powers of ϕ_0 is given by

$$\mathcal{T}_r^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) = iZ^{n/2} \left. \tilde{W}_{\text{amp.}}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) \right|_{\mathbf{p}_i^2 = -m^2}. \quad (6.51)$$

The factor Z in the equation corresponds to a finite renormalization of the field such that the residue of the two-point function (equation (6.50)) on the physical pole $\mathbf{p}^2 = -m^2$ of the renormalized field is 1. It ensures that the matrix elements $\mathcal{T}_r^{(n)}$ satisfy the *unitarity* relations with the proper normalization.

The generalization to several particles is straightforward.

S-matrix and generating functional of vertex functions. We now also exhibit a direct relation with the analytic continuation of the generating functional of *vertex* (or one-particle-irreducible) functions $\Gamma(\chi)$, defined in Section 7.7. We start from the Euclidean form

$$e^{\mathcal{W}(J)} = \int [d\phi] e^{-\mathcal{A}(\phi) + (J\phi)}, \quad (6.52)$$

with the notation

$$(J\phi) \equiv \int d^d x J(\mathbf{x})\phi(\mathbf{x}).$$

We introduce $\Gamma(\chi)$, the Legendre transform of $\mathcal{W}(J)$, defined by

$$\mathcal{W}(J) + \Gamma(\chi) = \int d^d x J(\mathbf{x})\chi(\mathbf{x}), \quad \text{with} \quad \chi(\mathbf{x}) = \frac{\delta \mathcal{W}}{\delta J(\mathbf{x})}.$$

Then,

$$e^{-\Gamma(\chi) + (J\chi)} = \int [d\phi] e^{-\mathcal{A}(\phi) + (J\phi)}. \quad (6.53)$$

Using

$$J(\mathbf{x}) = \frac{\delta \Gamma}{\delta \chi(\mathbf{x})},$$

we can rewrite equation (6.53) as

$$e^{-\Gamma(\chi)} = \int [d\phi] \exp \left[-\mathcal{A}(\phi) + \int d^d x (\phi(\mathbf{x}) - \chi(\mathbf{x})) \frac{\delta \Gamma}{\delta \chi(\mathbf{x})} \right], \quad (6.54)$$

or, equivalently, after translating $\phi(\mathbf{x})$,

$$e^{-\Gamma(\chi)} = \int [d\phi] \exp \left[-\mathcal{A}(\phi + \chi) + \int d^d x \phi(\mathbf{x}) \frac{\delta \Gamma}{\delta \chi(\mathbf{x})} \right]. \quad (6.55)$$

In the limit of a vanishing source J ,

$$\frac{\delta \Gamma}{\delta \chi(\mathbf{x})} = 0. \quad (6.56)$$

The equation has propagating-type solutions only after continuation to real time. We then observe that $\Gamma(\chi)$ coincides with $i\mathcal{T}(\chi)$, when the solution $\chi(x)$ of this equation is expanded around ϕ_0 defined in equation (6.33).

6.5.4 Change of field variables

In most of the previous discussion, we have derived S -matrix elements from field correlation functions. We now show that S -matrix elements are, to some extent, invariant under local field transformations. Field correlation functions thus contain more information than the scattering matrix. This raises fundamental questions if one takes the viewpoint that only the scattering data are physical. On the other hand, such a property is important in theories like gauge theories, where all gauges are equivalent, or models defined on Riemannian manifolds, where the fields $\phi_i(x)$ correspond only to a particular choice of coordinates on the manifold.

We have seen that S -matrix elements can be calculated from connected correlation functions by taking the residues of the poles of the external propagators (see Section 6.3.2). For simplicity, we consider the example of only one species of field ϕ , which we define in such a way that it has a vanishing expectation value. Then, the connected elements of the scattering matrix $\mathcal{S}^{(n)}$ can be expressed in terms of the amputated correlation functions (equations (6.51) and (7.71)) as

$$\tilde{\mathcal{S}}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) = Z^{n/2} \tilde{W}_{\text{amp.}}^{(n)}(\mathbf{p}_1, \dots, \mathbf{p}_n) |_{\mathbf{p}^2 = -m^2}, \quad (6.57)$$

where Z is the field renormalization constant.

6.5.5 S -matrix and field representation

We now compare the S -matrix obtained from the ϕ field correlation functions to the S -matrix derived from the correlation functions of a different field $\phi'(x)$ related to $\phi(x)$ by

$$\phi'(x) = C_1 \phi(x) + \sum_2^\infty \frac{C_k}{k!} \phi^k(x), \quad C_1 \neq 0. \quad (6.58)$$

We assume, when necessary, that the theory has been regularized in such a way that the new correlation functions exist (see Chapter 8) and we neglect the divergent Jacobian, which does not contribute to the S -matrix.

Using relation (6.58), we can express the ϕ' correlation functions in terms of the ϕ correlation functions. The expansion of the ϕ' propagator shows immediately that the ϕ and ϕ' propagators have poles at the same location (see Fig. 6.1).

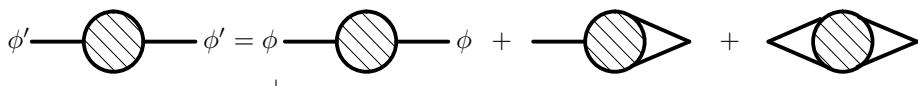


Fig. 6.1 ϕ' propagator.

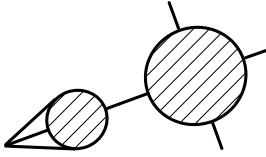


Fig. 6.2 One-line reducible ϕ^3 vertex insertion.

The contributions to n -point functions which have poles on the external lines then have the form shown in Fig. 6.2.

In the mass-shell limit ($\mathbf{p}_i^2 = -m^2$), the ϕ' and ϕ correlation functions become proportional. The S -matrix elements are identical. Here, again, all fields related by transformation (6.50) are equivalent.

In such situations not all parameters of the theory are physical. For example, the field renormalization is obviously non-physical. The same physical theory may have renormalizable and non-renormalizable realizations.

Using the background field method (see Section A7.2), one can avoid the calculation of non-physical quantities (however, to prove renormalizability, the study of correlation functions cannot be avoided (see Chapter 9)).

6.6 Källen–Lehmann representation and field renormalization

To precisely describe how the S -matrix can be calculated in a general theory of scalar particles, we still have to discuss field renormalization. This can be done more conveniently in the *Euclidean formulation* (see Chapter 7). We define the field $\phi(\mathbf{x})$ in such a way that it has zero-expectation value. This can always be achieved by a constant shift $\phi(\mathbf{x}) \mapsto \phi(\mathbf{x}) - \langle 0 | \phi | 0 \rangle$. We then consider the two-point correlation function $W^{(2)}$. It follows from the analysis of Section 2.5 that it is equal to the vacuum expectation value of the time-ordered product of two field operators

$$\delta^{d-1}(p + p') W^{(2)}(t, p) = (2\pi)^{d-1} \langle 0 | \tilde{\phi}(p) e^{-|t|(\mathbf{H} - E_0)} \tilde{\phi}(p') | 0 \rangle. \quad (6.59)$$

where E_0 is the ground state (or vacuum) energy. We adapt the arguments of Section A2.2 to the kinematic properties of an $O(d)$ invariant field theory. We introduce a complete set of projectors P_ε on the eigenspaces corresponding to the positive eigenvalues ε of the Hermitian, positive operator $\mathbf{H} - E_0$, and obtain

$$\delta^{d-1}(p + p') W^{(2)}(t, p) = (2\pi)^{d-1} \int d\varepsilon \langle 0 | \tilde{\phi}(p) P_\varepsilon e^{-\varepsilon|t|} \tilde{\phi}(p') | 0 \rangle.$$

The theory being translation invariant, we know that the result is proportional to $\delta^{d-1}(p + p')$. In this limit, since $\tilde{\phi}(-p) = \tilde{\phi}^\dagger(p)$, the left-hand side is a sum of positive terms (a point that we have already discussed in Section A2.2). We thus find an expression of the form,

$$W^{(2)}(t, p) = (2\pi)^{d-1} \int d\varepsilon \rho(\varepsilon, p) e^{-\varepsilon|t|},$$

where ρ is a positive measure. The Fourier transform with respect to time is given by

$$\tilde{W}^{(2)}(p_0, p) = (2\pi)^{d-1} \int dt e^{-ip_0 t} \int d\varepsilon \rho(\varepsilon, p) e^{-\varepsilon|t|} = (2\pi)^{d-1} \int d\varepsilon \frac{2\varepsilon \rho(\varepsilon, p)}{p_0^2 + \varepsilon^2}.$$

Using the $O(d)$ invariance, we can set $p = 0$ and replace p_0^2 by \mathbf{p}^2 without changing the result. We then obtain the *Källen–Lehmann representation*,

$$\tilde{W}^{(2)}(\mathbf{p}) = \int d\varepsilon \frac{\rho(\varepsilon)}{\mathbf{p}^2 + \varepsilon^2}, \quad (6.60)$$

where $\rho(\varepsilon)$, proportional to $\rho(\varepsilon, 0)$, is a positive measure.

Since, by definition, the physical mass m is the lowest energy eigenstate above the ground state, the domain of integration is $\varepsilon \geq m$. Moreover, at $p = 0$, the state is isolated. Therefore, the measure has an isolated δ -function and then a continuous part starting at the threshold for scattering states (in the simple scalar field theory $2m$, or $3m$ if the parity of the number of particles is conserved):

$$\rho(\varepsilon) = Z\delta(\varepsilon - m) + \rho'(\varepsilon), \quad \text{with } \rho'(\varepsilon) = 0, \text{ for } \varepsilon < 2m. \quad (6.61)$$

Conversely, the Fourier transform with respect to the energy variable p_0 now is

$$\begin{aligned} W^{(2)}(t, p) &= \frac{1}{2\pi} \int dp_0 e^{ip_0 t} \tilde{W}^{(2)}(\mathbf{p}) = \frac{1}{2\pi} \int dp_0 e^{ip_0 t} \int d\varepsilon \frac{\rho(\varepsilon)}{\mathbf{p}^2 + \varepsilon^2} \\ &= \int d\varepsilon \frac{\rho(\varepsilon)}{2\sqrt{p^2 + \varepsilon^2}} e^{-|t|\sqrt{p^2 + \varepsilon^2}}. \end{aligned} \quad (6.62)$$

Returning to the definition (6.59), assuming $t > 0$, taking the derivative with respect to time and the limit $t = 0$, one finds

$$\delta^{d-1}(p + p') \left. \frac{\partial}{\partial t} W^{(2)}(p, t) \right|_{t \rightarrow 0+} = -(2\pi)^{d-1} \langle 0 | \tilde{\phi}(p)(\mathbf{H} - E_0)\tilde{\phi}(p') | 0 \rangle. \quad (6.63)$$

The product $(\mathbf{H} - E_0)\tilde{\phi}$ in the right-hand side can be replaced by the commutator $[\mathbf{H}, \tilde{\phi}]$. Then, using the reality of the left-hand side, one can take the Hermitian part of the operator. Thus,

$$\delta^{d-1}(p + p') \left. \frac{\partial}{\partial t} W^{(2)}(p, t) \right|_{t \rightarrow 0+} = -\frac{1}{2}(2\pi)^{d-1} \langle 0 | [\tilde{\phi}(p), [\mathbf{H}, \tilde{\phi}(p')]] | 0 \rangle.$$

If the action has the form (6.26), the commutator is proportional to the conjugated momentum. Then,

$$[\tilde{\phi}(p), [\mathbf{H}, \tilde{\phi}(p')]] = (2\pi)^{1-d} \delta^{d-1}(p + p').$$

It follows that

$$\left. \frac{\partial}{\partial t} W^{(2)}(t, p) \right|_{t \rightarrow 0+} = -\frac{1}{2}.$$

Calculating the left-hand side from the representation (6.62), one obtains

$$\int d\varepsilon \rho(\varepsilon) = 1.$$

The conclusion is that, except in a free-field theory, the residue Z of the pole at $\mathbf{p}^2 = -m^2$, is strictly smaller than 1:

$$0 < Z < 1. \quad (6.64)$$

Implication for the S-matrix. This result has several implications, one being related to the S -matrix. The state of lowest energy, with momentum p , gives the leading contribution to $W^{(2)}(t, p)$ for t large. From the Källen–Lehmann representation (6.62) and the decomposition (6.61), one then learns that

$$W^{(2)}(t, p) \underset{t \rightarrow \infty}{\sim} \frac{Z}{2\sqrt{p^2 + m^2}} e^{-|t|\sqrt{p^2 + m^2}}.$$

Comparing the result with the contribution of a normalized one-particle eigenstate of the Hamiltonian \mathbf{H} (see equation (6.21)),

$$(2\pi)^{d-1} \langle 1, p | e^{-|t|(\mathbf{H}-E_0)} | 1, p' \rangle = \delta^{d-1}(p + p') \frac{1}{2\sqrt{p^2 + m^2}} e^{-|t|\sqrt{p^2 + m^2}},$$

one notes that the field ϕ has only a component \sqrt{Z} on the one-particle state. Another way to formulate the same result is to verify that, in real time, the Heisenberg field has a free-field large time behaviour with an amplitude \sqrt{Z} on normalized creation or annihilation operators.

After continuation to real time, one finds the two-point function

$$\begin{aligned} \langle 0 | T[\tilde{\phi}(t, p)\tilde{\phi}(p', 0)] | 0 \rangle &= \langle 0 | \tilde{\phi}(p) e^{-i(\mathbf{H}-E_0)|t|} \tilde{\phi}(p') | 0 \rangle \\ &= (2\pi)^{1-d} \delta^{d-1}(p + p') \int d\varepsilon \frac{\rho(\varepsilon)}{2\sqrt{p^2 + \varepsilon^2}} e^{-i|t|\sqrt{p^2 + \varepsilon^2}}. \end{aligned}$$

It can be verified that its large time behaviour is related to the leading singularity of the measure ρ . Since $\rho(\varepsilon)$ is the sum of a δ -function and a continuous function (for $d \geq 2$), one obtains

$$\langle 0 | \tilde{\phi}(p) e^{-i(\mathbf{H}-E_0)|t|} \tilde{\phi}(p') | 0 \rangle \underset{|t| \rightarrow \infty}{=} Z(2\pi)^{1-d} \frac{1}{2\omega(p)} \delta^{d-1}(p + p') e^{-i\omega(p)|t|} + O(1/t).$$

Introducing $\varphi, \bar{\varphi}$, the properly normalized creation and annihilation operators of the one-particle states, one concludes that, for large time, the field $\tilde{\phi}(t, p)$ tends, in a weak sense (not in an operator sense, but in all expectation values), towards

$$\tilde{\phi}(t, p) \underset{|t| \rightarrow \infty}{\sim} \sqrt{Z} \frac{1}{2\omega(p)} [\varphi(-p) + \bar{\varphi}(p)].$$

The constant \sqrt{Z} is the field *renormalization constant*.

Normalized S-matrix elements. Properly normalized S -matrix elements can be obtained by using the action $\mathcal{A}(\phi\sqrt{Z})$. Alternatively, the S -matrix elements calculated in terms of the initial field have to be renormalized.

7 Perturbative quantum field theory (QFT): Algebraic methods

In Chapter 6, we have introduced the quantum theory of the relativistic scalar field, natural relativistic extension of the non-relativistic quantum statistical physics of the Bose gas (Section 4.4.2), in the formalism of Section 4.4.7.

In this chapter, we begin a more systematic study of the algebraic properties of perturbation theory in the example of a *local, relativistic* QFT. We discuss only scalar fields and postpone the study of relativistic fermions to Chapter 12, because this requires first describing the representations of the spin group. However, most algebraic results can be easily generalized.

We consider the Euclidean formulation of QFT based on the density matrix at thermal equilibrium, but we mainly investigate the simpler zero-temperature limit where all d coordinates, Euclidean time and space, can be treated symmetrically.

Our discussion is based on field integrals (for an early textbook see, *e.g.* Ref. [29]). The field integral defines a functional measure to which correspond expectation values of product of fields called *correlation functions*. These functions are analytic continuations to imaginary (Euclidean) time of the vacuum expectation values of time-ordered products of field operators, which *we shall also call correlation functions*, although the functional measure then is complex. As we have already explained in Sections 2.4.2 and 2.5, in the case of the quantum particle, they have also an interpretation as correlation functions in some models of classical statistical physics, in continuum formulations, or at equal time of finite temperature quantum field theory. Finally, they naturally appear in non-relativistic quantum statistical physics in various limits, high or critical temperatures.

We first calculate a Gaussian integral, which includes the example of the free field theory. Adding a source term to the action, we obtain a generating functional of Gaussian correlation functions. The field integral, corresponding to a general action with an interaction polynomial in the field, can then be expressed as an infinite sum of Gaussian expectation values, which can be calculated, for example, with the help of Wick's theorem. Each contribution has a graphical representation in terms of *Feynman diagrams*.

We show, because this may no longer be obvious, that the properties of the perturbative expansion are consistent with the usual manipulations performed on integrals, like integration by parts and change of variables. We define the functional δ -function.

Quite generally, the field integral, corresponding to an action to which a term linear in the field coupled to an external source (or field) J has been added, defines a generating functional $\mathcal{Z}(J)$ of field correlation functions (the partition function in an external field) [9, 30]. The functional $\mathcal{W}(J) = \ln \mathcal{Z}(J)$ (analogous to the free energy of statistical physics) is then the generating functional of *connected correlation functions*, to which contribute only connected Feynman diagrams. In a local field theory (the action is the integral of a function of the field and its derivatives at the same point), connected correlation functions, as a consequence of locality, have *cluster properties*.

The Legendre transform $\Gamma(\varphi)$ of $\mathcal{W}(J)$ is the generating functional of *vertex functions* (sometimes called effective potential and analogous to the thermodynamic potential of statistical physics) [31]. Only one-line irreducible Feynman diagrams (diagrams that cannot be disconnected by cutting only one line) contribute to vertex functions.

These diagrams are also called one-particle irreducible (or *1PI*, an acronym we often use in this work) in particle physics.

Vertex functions play a basic role in the renormalization of local quantum field theories. In Section 7.11, we give a quantum interpretation to $\Gamma(\varphi)$. We also relate it to the partition function at fixed field time average. This relation explains why $\Gamma(\varphi)$ also appears in the discussion of symmetry breaking.

We explain how to calculate these functionals in a reorganized perturbative expansion, called *loop expansion* [32].

In the appendix, we calculate the generating functional of two-loop Feynman diagrams, introduce the background field method, and discuss some properties of Feynman diagrams relevant for cluster properties.

Warning. Many quantities that we meet in the chapter may be infinite in the case of relativistic quantum field theories in their straightforward definition, a problem that is carefully studied in Chapters 8 and 9. Therefore, the theories have to be *regularized*, for example, by modifying the propagator at short distance, or using a space-time lattice. In this chapter, we implicitly assume such a regularization (see Chapter 8).

7.1 Generating functionals of correlation functions

Correlation functions. We consider Euclidean, local action $\mathcal{S}(\phi)$ for a neutral scalar field ϕ in d -Euclidean dimensions, of the type introduced in Section 6.1, and define the n -point ϕ -field correlation function (a generalized moment) by the field integral,

$$Z^{(n)}(x_1, \dots, x_n) \equiv \langle \phi(x_1) \cdots \phi(x_n) \rangle \equiv \frac{1}{\mathcal{Z}_0} \int [d\phi] \phi(x_1) \cdots \phi(x_n) e^{-\mathcal{S}(\phi)}, \quad (7.1)$$

where \mathcal{Z}_0 is the partition function,

$$\mathcal{Z}_0 = \int [d\phi] e^{-\mathcal{S}(\phi)}.$$

The field integral

$$\mathcal{Z}(J) = \int [d\phi] \exp \left[-\mathcal{S}(\phi) + \int d^d x J(x) \phi(x) \right], \quad (7.2)$$

where $J(x)$ is an external field or source, is a generating functional of correlation functions (see also Section 2.5).

Indeed, expanded in powers of J , it yields

$$\mathcal{Z}(J) = \mathcal{Z}_0 \sum_{n=0}^{\infty} \frac{1}{n!} \int d^d x_1 \cdots d^d x_n Z^{(n)}(x_1, \dots, x_n) J(x_1) \cdots J(x_n). \quad (7.3)$$

We now introduce functional differentiation, a differentiation completely defined by

$$\frac{\delta J(x)}{\delta J(y)} = \delta^{(d)}(x - y), \quad (7.4)$$

where $\delta^{(d)}(x)$ is the d -dimensional Dirac function. Then,

$$\frac{\delta}{\delta J(y)} e^{J \cdot \phi} = \phi(y) e^{J \cdot \phi}. \quad (7.5)$$

Applying this identity to the field integral (7.2), one verifies that ϕ -field correlation functions can also be obtained from the *generating functional* (7.2) by functional differentiation, as

$$Z^{(n)}(x_1, \dots, x_n) = \frac{1}{\mathcal{Z}_0} \left[\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{Z}(J) \right] \Big|_{J=0}. \quad (7.6)$$

7.2 Perturbative expansion. Wick's theorem and Feynman diagrams

We assume *translation invariance* in this chapter.

To generate a perturbative expansion, we write $\mathcal{S}(\phi)$ as the sum of a quadratic term and an interaction, of the form,

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x d^d y \phi(x) K(x-y) \phi(y) + \mathcal{V}_I(\phi), \quad (7.7)$$

where $\mathcal{V}_I(\phi)$ is a *local* polynomial, and where the symmetric, positive kernel K is chosen to render all terms of the perturbative expansion in powers of $\mathcal{V}_I(\phi)$ finite (a regularization, see, *e.g.*, Section 8.4.2). For a field with mass m , the kernel K can be chosen as a local operator of the form ($\nabla \equiv (\partial/\partial x_1, \dots, \partial/\partial x_d)$),

$$K(x-y) = \kappa(-\nabla^2) \delta^{(d)}(x-y), \quad (7.8)$$

where κ is an analytic function on the real axis, positive for $z > 0$, of the form

$$\kappa(z) = m^2 + z + O(z^2).$$

In what follows, we use the symbolic notation

$$(\phi K \phi) \equiv \int d^d x d^d y \phi(x) K(x-y) \phi(y), \quad J \cdot \phi \equiv \int d^d x J(x) \phi(x). \quad (7.9)$$

7.2.1 Gaussian integral and free field theory

We consider the general Gaussian field integral,

$$\mathcal{Z}_G(J) = \int [d\phi] \exp \left[-\frac{1}{2} (\phi K \phi) + J \cdot \phi \right]. \quad (7.10)$$

In expression (7.10), a normalization is implied; we choose $\mathcal{Z}_G(0) = 1$. The functional $\mathcal{Z}_G(J)$ is the generating functional of ϕ -field correlation functions corresponding to the weight $e^{-(\phi K \phi)/2}$. The general expression (7.6) yields the Gaussian correlation functions,

$$\langle \phi(x_1) \cdots \phi(x_n) \rangle_0 = \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{Z}_G(J) \Big|_{J=0},$$

where $\langle \bullet \rangle_0$ means Gaussian expectation value.

The kernel Δ , inverse of K , is defined by

$$\int d^d z \Delta(x-z) K(z-y) = \delta^{(d)}(x-y). \quad (7.11)$$

To calculate the field integral (7.10), we shift $\phi(x)$ by $\int d^d y \Delta(x-y) J(y)$, and find after integration,

$$\mathcal{Z}_G(J) = \exp \left[\frac{1}{2} (J \Delta J) \right], \quad (7.12)$$

with the notation

$$(J \Delta J) = \int d^d x d^d y J(x) \Delta(x-y) J(y).$$

Symmetry implies $\langle \phi(x) \rangle_0 = 0$. Then, differentiating expression (7.12), one infers,

$$\langle \phi(x_1) \phi(x_2) \rangle_0 = \frac{\delta^2 \mathcal{Z}_G(J)}{\delta J(x_1) \delta J(x_2)} \Big|_{J=0} = \Delta(x_1 - x_2). \quad (7.13)$$

An example of a Gaussian theory is provided by a free scalar field theory, where $\kappa(z) = m^2 + z$. The kernel Δ then is given by

$$\Delta(x) = \frac{1}{(2\pi)^d} \int d^d p \frac{e^{ip \cdot x}}{p^2 + m^2}. \quad (7.14)$$

7.2.2 Perturbative expansion: A compact expression

In Section 2.6, we show how to calculate a path integral for a Hamiltonian of the form $\frac{1}{2}p^2 + \frac{1}{2}\omega^2q^2 + V_1(q)$ as an expansion in powers of $V_1(q)$, for any function $V_1(q)$ expandable in powers of q . The result is based on the calculation of a reference Gaussian integral (in Chapter 2, the harmonic oscillator). Here, we apply a similar method to field integrals.

Furthermore, although most of the results derived in this chapter are illustrated with examples corresponding to an action of the form (6.45) (but suitably regularized, see Section 8.4.2), the results apply to any boson field theory, and this explains the choice of the abstract notation (7.9).

We consider a general Euclidean action of the form (7.7). We use the compact notation,

$$D_J \equiv \delta/\delta J, \quad (7.15)$$

each time $\delta/\delta J$ appears as an argument. Then, it follows from the property (7.5) that we can express the field integral

$$\mathcal{Z}(J) = \int [d\phi] \exp [-\mathcal{S}(\phi) + J \cdot \phi], \quad (7.16)$$

in terms of $\mathcal{Z}_G(J)$ (equation (7.10)) as

$$\mathcal{Z}(J) = \exp [-V_1(D_J)] \mathcal{Z}_G(J) = \exp [-V_1(D_J)] \exp \left(\frac{1}{2} J \Delta J \right). \quad (7.17)$$

The expression (7.17) expresses, in a compact algebraic form, the result of the calculations of all Gaussian expectation values obtained by expanding the field integral (7.1), in the example of equation (7.7), as a formal series in powers of the interaction potential $V_1(\phi)$. The expansion, then combined with the identity (7.6), generates the perturbative expansion of all ϕ -field correlation functions [34–37].

7.2.3 Wick's theorem

The direct expansion of the field integral (7.1), in the example of equation (7.7), in powers of $V_1(\phi)$ reduces all calculations to Gaussian expectation values of products of fields. From the expression (7.17), and using the arguments of Section 1.1, one obtains a straightforward generalization of equations (1.9–1.14) or (2.61), which expresses Wick's theorem [1] in field theory,

$$\begin{aligned} \left\langle \prod_1^{2s} \phi(z_i) \right\rangle_0 &= \left[\prod_{i=1}^{2s} \frac{\delta}{\delta J(z_i)} \exp \left(\frac{1}{2} J \Delta J \right) \right]_{J=0} \\ &= \sum_{\substack{\text{all possible pairings} \\ \text{of } \{1, 2, \dots, 2s\}}} \Delta(z_{i_1} - z_{i_2}) \cdots \Delta(z_{i_{2s-1}} - z_{i_{2s}}). \end{aligned} \quad (7.18)$$

Perturbation theory involves, as a basic ingredient, the two-point function Δ of the Gaussian theory (equation (7.13)), which we call the *propagator*.

Graphically, each term in the sum can be represented by a set of contractions corresponding to a particular pairing. For example, for $s = 2$, one finds

$$\begin{aligned} \langle \phi(z_1) \phi(z_2) \phi(z_3) \phi(z_4) \rangle_0 &= \overline{\phi(z_1)} \overline{\phi(z_2)} \overline{\phi(z_3)} \overline{\phi(z_4)} + 2 \text{ terms} \\ &= \Delta(z_1 - z_2) \Delta(z_3 - z_4) + \Delta(z_1 - z_3) \Delta(z_2 - z_4) + \Delta(z_1 - z_4) \Delta(z_2 - z_3). \end{aligned}$$

7.2.4 Feynman diagrams

When the interaction terms are local, that is, integrals of polynomials of the field $\phi(x)$ and its derivatives, any perturbative contribution to the n -point correlation function is a Gaussian expectation value of the form (for simplicity, we omit derivative couplings, because they leave the argument unchanged)

$$\left\langle \phi(x_1) \cdots \phi(x_n) \int d^d y_1 \phi^{p_1}(y_1) \int d^d y_2 \phi^{p_2}(y_2) \cdots \int d^d y_k \phi^{p_k}(y_k) \right\rangle_0.$$

Therefore, it is a sum of products of propagators integrated over the points corresponding to interaction vertices. It is then possible to give a graphical representation of each product [37]: a propagator is represented by a line joining the two points which appear as arguments; moreover, any point that is common to more than one line corresponds to an argument that has to be integrated over.

Ultraviolet divergences. Since interacting theories with a propagator of the form (7.14) have large momentum (ultraviolet) or short distance divergences, in what follows we assume that the field theory has been *regularized* by any one of the methods that ensures the convergence of all terms in the perturbative expansion (see Sections 8.4.2, 8.7, and 10.1).

7.3 Connected correlation functions: Generating functional

We now define another generating functional, analogous to the free energy of statistical physics,

$$\mathcal{W}(J) = \ln \mathcal{Z}(J), \quad (7.19)$$

the generating functional of *connected correlation functions*.

We introduce the notation

$$\mathcal{W}(J) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d^d x_1 \cdots d^d x_n W^{(n)}(x_1, x_2, \dots, x_n) J(x_1) J(x_2) \cdots J(x_n). \quad (7.20)$$

The connected diagrams that contribute to the partition function \mathcal{Z}_0 (also called vacuum diagrams in particle physics), contribute only to $\mathcal{W}(J = 0)$.

Then,

$$W^{(n)}(x_1, \dots, x_n) = \left[\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{W}(J) \right] \Big|_{J=0}. \quad (7.21)$$

The functions $W^{(n)}$ are more directly related to physical observables.

The three first relations, as implied by the definition (7.19), between complete correlation functions $Z^{(n)}$ and connected functions $W^{(n)}$ are

$$\begin{aligned} Z^{(1)}(x) &= W^{(1)}(x), \\ Z^{(2)}(x_1, x_2) &= W^{(2)}(x_1, x_2) + W^{(1)}(x_1)W^{(1)}(x_2), \\ Z^{(3)}(x_1, x_2, x_3) &= W^{(3)}(x_1, x_2, x_3) + W^{(1)}(x_1)W^{(2)}(x_2, x_3) + W^{(1)}(x_2)W^{(2)}(x_3, x_1) \\ &\quad + W^{(1)}(x_3)W^{(2)}(x_1, x_2) + W^{(1)}(x_1)W^{(1)}(x_2)W^{(1)}(x_3). \end{aligned}$$

Note that the right-hand side involves all possible products of connected functions with coefficient 1.

7.3.1 An alternative proof of connectivity

In Section 1.2.1, we give a combinatorial proof in a sense of a perturbative diagrammatic expansion: only connected Feynman diagrams contribute to $\mathcal{W}(J)$. As a consequence, the functions $W^{(n)}$, unlike the functions $Z^{(n)}$, contain no contributions that can be factorized into products of the form

$$F_1(x_1, \dots, x_p) F_2(x_{p+1}, \dots, x_n),$$

where the two disjoint sets of arguments are not empty.

This property is the starting point of an alternative global algebraic proof of connectivity which relies, in particular, on linearity and locality: a linear combination of connected functions is still connected.

In Section 7.2.1, we have calculated the generating functional $\mathcal{Z}_G(J)$ for a Gaussian theory, and found (equation (7.12)),

$$\mathcal{Z}_G(J) = \exp \left[\frac{1}{2} \int d^d x d^d y J(x) \Delta(x - y) J(y) \right], \quad (7.22)$$

where $\Delta(x - y)$ is the propagator.

Expanding in powers of J , we immediately verify that, except for the two-point correlation function, all correlation functions are disconnected. By contrast, the expansion of the functional

$$\mathcal{W}_G(J) = \ln \mathcal{Z}_G(J) = \frac{1}{2} \int d^d x d^d y J(x) \Delta(x - y) J(y),$$

only involves the connected contribution. All functions $W^{(n)}$ vanish, except $W^{(2)} = \Delta$, which is connected.

We now assume that, for some action $\mathcal{S}(\phi)$, we have proved that $\mathcal{W}(J)$ generates connected correlation functions, and we add a local perturbation to the action:

$$\mathcal{S}_\varepsilon(\phi) = \mathcal{S}(\phi) + \varepsilon \int d^d x \phi^N(x),$$

(derivative couplings leave the argument unchanged). Then, expanding to first order in ε , we obtain,

$$\begin{aligned} e^{\mathcal{W}_\varepsilon(J)} &= \int [d\phi] \exp \left[-\mathcal{S}_\varepsilon(\phi) + \int d^d x \phi(x) J(x) \right] \\ &= \left[1 - \varepsilon \int d^d x \left(\frac{\delta}{\delta J(x)} \right)^N \right] \exp \mathcal{W}(J) + O(\varepsilon^2), \end{aligned}$$

and, therefore,

$$\mathcal{W}_\varepsilon(J) = \mathcal{W}(J) - \varepsilon e^{-\mathcal{W}(J)} \int d^d x \left(\frac{\delta}{\delta J(x)} \right)^N e^{\mathcal{W}(J)} + O(\varepsilon^2).$$

The contribution of order ε is a linear combination of products of derivatives of $\mathcal{W}(J)$ with respect to a source at a unique point x . For example, for $N = 3$,

$$\mathcal{W}_\varepsilon(J) - \mathcal{W}(J) = -\varepsilon \int d^d x \left[\frac{\delta^3 \mathcal{W}(J)}{[\delta J(x)]^3} + 3 \frac{\delta^2 \mathcal{W}(J)}{[\delta J(x)]^2} \frac{\delta \mathcal{W}(J)}{\delta J(x)} + \left(\frac{\delta \mathcal{W}(J)}{\delta J(x)} \right)^3 \right] + O(\varepsilon^2).$$

Since $\mathcal{W}(J)$ contains only connected contributions, all terms are products of connected contributions linked to the same point, x . Therefore, if $\mathcal{W}(J)$ is connected, all terms of order ε are also connected.

The argument immediately generalizes to any local polynomial $\lambda V_1(\phi)$ (λ is a parameter). Since $\mathcal{W}(J)$ is connected for a general Gaussian theory, it follows, after integration over the parameter λ , that it remains connected for any interaction.

The functions $W^{(n)}$ are thus indeed connected correlation functions. Occasionally, we will emphasize this character by using the notation

$$W^{(n)}(x_1, \dots, x_n) = \langle \phi(x_1) \cdots \phi(x_n) \rangle_c, \quad (7.23)$$

where the symbol $\langle \bullet \rangle_c$ means a connected contribution to the correlation function.

7.3.2 Inversion

The relation between action and generating functional, which has the form of a Laplace transformation, can formally be inverted:

$$e^{-S(\phi)} = \int [dJ] \exp \left[\mathcal{W}(J) - \int d^d x J(x) \phi(x) \right], \quad (7.24)$$

where one integrates over imaginary sources $J(x)$ [38]. A truncated loop expansion (see Section 7.9) of the field integral then yields approximate non-linear equations for correlation functions. It is actually convenient to introduce, in the right-hand side, the generating functional of vertex functions defined in Section 7.7.

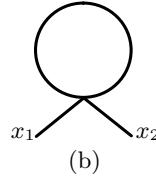


Fig. 7.1 Two-point function at order g

7.4 The example of the ϕ^4 QFT

We illustrate the previous discussion with the second order expansion of the important quartic example,

$$\mathcal{V}_I(\phi) = \frac{g}{4!} \int d^d x \phi^4(x).$$

The two-point function. The two-point function to order g^2 has the expansion

$$\langle \phi(x_1) \phi(x_2) \rangle = (a) - \frac{1}{2} g (b) + \frac{1}{4} g^2 (c) + \frac{1}{4} g^2 (d) + \frac{1}{6} g^2 (e) + O(g^3).$$

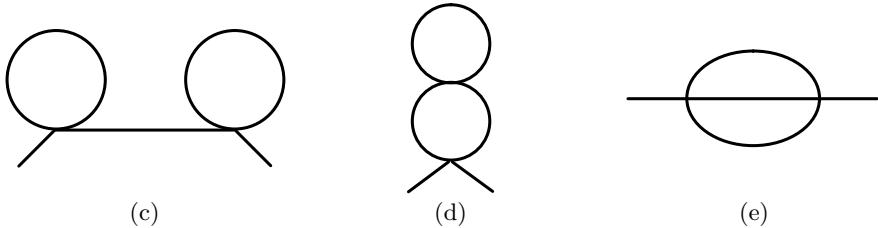


Fig. 7.2 Contributions of order g^2 to the two-point function

Note that, three additional contributions, which factorize into

$$\langle \phi(x_1)\phi(x_2) \rangle_0 \langle \phi^4(y) \rangle_0, \quad \langle \phi(x_1)\phi(x_2)\phi^4(y_1) \rangle_0 \langle \phi^4(y_2) \rangle_0, \text{ and} \\ \langle \phi(x_1)\phi(x_2) \rangle_0 \langle \phi^4(y_1)\phi^4(y_2) \rangle_0,$$

cancel in the division by $\mathcal{Z}(J = 0)$. The diagrams contributing to $\mathcal{Z}(0)$ (the partition function) are called *vacuum diagrams* in particle terminology.

Then,

- (a) is the propagator: x_1 ————— x_2 ;
 (b) is the Feynman diagram that appears at order g and is displayed in Fig. 7.1; (c),
 (d), (e) are the three diagrams displayed in Fig. 7.2.

Let us explain, for example, in detail the weight $1/6$ in front of diagram (e). Expanding the exponential at second order, we have to calculate the Gaussian expectation value of

$$\frac{g^2}{2! \langle 4|} \int d^d y_1 \int d^d y_2 \left\langle \phi(x_1) \phi(x_2) \phi^4(y_1) \phi^4(y_2) \right\rangle_0 .$$

We apply Wick's theorem. First, $\phi(x_1)$ can be associated with any ϕ field of the interaction terms; there are eight choices, and one interaction term is distinguished. Then, $\phi(x_2)$ must be contracted with a field of the remaining interaction term: four choices. Finally, the three remaining fields of the first interaction term can be paired with any permutation of the fields of the second one: $3!$ equivalent possibilities. Multiplying all factors, one finds

$$\frac{1}{2} \frac{1}{(4!)^2} \times 8 \times 4 \times 3! = \frac{1}{6}.$$

Note also that the factor $1/6$ multiplying the diagram can be shown to have an interpretation as $1/3!$, the combinatorial factor in the denominator reflecting the symmetry under permutation of the three lines joining the two vertices. There exist systematic expressions giving the weight factor of Feynman diagrams in terms of the symmetry group of the graph.

A useful practical remark is the following: the sum of all weight factors at a given order can be calculated from the ‘zero-dimensional’ field theory obtained by suppressing the arguments of the field and all derivatives and integration in the action, because the propagator can then be normalized to 1. For example, in the case of the ϕ^4 field considered here, the action becomes

$$\mathcal{S}(\phi) = \frac{1}{2}\phi^2 + \frac{1}{4!}g\phi^4,$$

and the two-point function is given by

$$Z^{(2)} = \frac{\int d\phi \phi^2 \exp[-\mathcal{S}(\phi)]}{\int d\phi \exp[-\mathcal{S}(\phi)]} = 1 - \frac{g}{2} + \frac{2}{3}g^2 + O(g^3),$$

in which the expressions correspond to ordinary one-variable integrals.

The sum rules are satisfied. For example, at order g^2 ,

$$\frac{2}{3} = \frac{1}{4} + \frac{1}{4} + \frac{1}{6}.$$

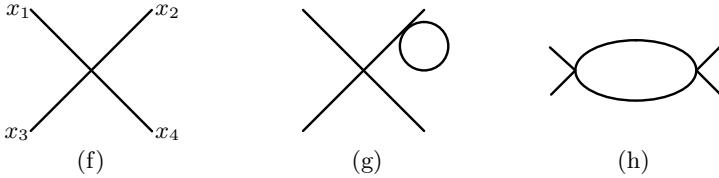


Fig. 7.3 New Feynman diagrams contributing to the four-point function

The four-point function. The expansion of the four-point function to order g^2 is

$$\begin{aligned} \langle \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4) \rangle &= [(a)_{12} (a)_{34} + 2 \text{ terms}] - \frac{1}{2}g [(a)_{12} (b)_{34} + 5 \text{ terms}] \\ &\quad - g (f) + g^2 \{ (a)_{12} [\frac{1}{4}((c)_{34} + (d)_{34}) + \frac{1}{6}(e)_{34}] + 5 \text{ terms} \} \\ &\quad + \frac{1}{4}g^2 [(b)_{12} (b)_{34} + 2 \text{ terms}] + \frac{1}{2}g^2 [(g) + 3 \text{ terms}] + \frac{1}{2}g^2 [(h) + 2 \text{ terms}] + O(g^3). \end{aligned}$$

The notation $(a)_{12}$, for example, means diagram (a) contributing to the two-point function with arguments x_1 and x_2 . Finally, the additional terms are obtained by exchanging the external arguments to restore the permutation symmetry of the four-point function.

Again, as for the two-point function, we have omitted disconnected diagrams in which one factor has no external arguments. As one can check directly here, and as the general arguments in Section 7.3 will confirm, these diagrams are cancelled by the perturbative expansion of $\mathcal{Z}(J = 0)$.

The graphs that involve only contributions to the two-point functions ((a), (b), (c), (d), (e)) are disconnected, that is, factorize into a product of functions depending on disjoint subsets of variables. The origin of this phenomenon has already been indicated in Section 1.2.1, and further discussed in Section 7.3. Finally, the new *connected diagrams*, (f), (g), (h), are displayed in Fig. 7.3.

A final remark: local interaction terms may also involve derivatives of the field $\phi(x)$. Then in expression (7.18) derivatives of the propagator appear. The representation in terms of the Feynman diagrams given previously is no longer faithful, since it does not indicate the positions of the derivatives. A more faithful representation can be obtained by splitting points at the vertices and putting arrows on lines.

7.5 Algebraic properties of field integrals. Quantum field equations

Field integrals: Perturbative definition. Rigorously proving the existence of field integrals is not a simple mathematical problem and is solved only in particular examples. However, some difficulties in the very definition of field integrals simply reflect the problem of ambiguities or divergences in perturbation theory, as the discussion of Section 3.3 has already shown. Since straightforward perturbation theory in QFT exhibits divergences, we assume that the QFT has been regularized (see Chapters 8–10), and the perturbation theory defined. This implies, in particular, that the fields contributing to the field integral are regular enough (at least continuous).

Assuming a suitable regularization, here we want to prove that algebraic properties of perturbation theory derived by field integral techniques are consistent with results obtained by working directly with the perturbative expansion [39, 40]. This is sometimes an issue, especially in the case of *dimensional continuation*.

The proofs rely on simple but powerful techniques, which, simultaneously, lead to a derivation of important identities satisfied by correlation functions, like *Schwinger–Dyson equations* (a form of the quantum field equations expressed in terms of correlation functions). Finally, the formal manipulations involved in these derivations, should familiarize the reader with some algebraic properties of the perturbative expansion.

Therefore, we want to show that some properties derived from the field integral (7.16) with action (7.7),

$$\mathcal{Z}(J) = \int [d\phi] \exp \left[-\frac{1}{2}(\phi K \phi) - \mathcal{V}_I(\phi) + J \cdot \phi \right], \quad (7.25)$$

are consistent with properties derived from the corresponding perturbative expression (7.17) (using the notation (7.15)),

$$\mathcal{Z}(J) = \exp[-\mathcal{V}_I(D_J)] \exp \left[\frac{1}{2}(J \Delta J) \right]. \quad (7.26)$$

We assume that all terms in the expansion of this expression in powers of \mathcal{V}_I exist and are finite (which, as we shall see later, implies some conditions for the kernel K and the interaction term $\mathcal{V}_I(\phi)$). From the field integral representation (7.16), we now derive identities satisfied by the generating functional $\mathcal{Z}(J)$, and then prove algebraically that these identities also follow from the perturbative definition (7.17).

7.5.1 Integration by parts and quantum field equations

The integral of a total derivative vanishes. This simple property implies

$$\int [d\phi] \frac{\delta}{\delta \phi(x)} \exp(-\mathcal{S}(\phi) + J \cdot \phi) = 0, \quad (7.27)$$

and, therefore,

$$\int [d\phi] \left[J(x) - \frac{\delta \mathcal{S}}{\delta \phi(x)} \right] \exp(-\mathcal{S}(\phi) + J \cdot \phi) = 0. \quad (7.28)$$

Using the identity (7.5), one can transform the equation into

$$\left[\frac{\delta \mathcal{S}(D_J)}{\delta \phi(x)} - J(x) \right] \mathcal{Z}(J) = 0. \quad (7.29)$$

Equation (7.29) is a compact form of quantum field or *Schwinger–Dyson equations* [36, 30]. It is equivalent to an infinite set of relations between correlation functions obtained by expanding in powers of the source $J(x)$. These equations, in turn, can be solved perturbatively, and, as we show, determine correlation functions completely.

Solution of the quantum field equations. Conversely, the solution of equation (7.29) is the field integral (7.16). To prove this statement, we express $\mathcal{Z}(J)$, the solution of equation (7.29), as a generalized Laplace transform,

$$\mathcal{Z}(J) = \int [d\phi] \exp[-\Sigma(\phi) + J \cdot \phi]. \quad (7.30)$$

Equation (7.29) implies

$$\int [d\phi] \exp[-\Sigma(\phi) + J \cdot \phi] \left[\frac{\delta \mathcal{S}(\phi)}{\delta \phi(x)} - J(x) \right] = 0. \quad (7.31)$$

Using equation (7.28), but with \mathcal{S} replaced by Σ , we transform the equation into

$$\int [d\phi] \exp [-\Sigma(\phi) + J \cdot \phi] \frac{\delta}{\delta \phi(x)} [\mathcal{S}(\phi) - \Sigma(\phi)] = 0. \quad (7.32)$$

Therefore, $\mathcal{S}(\phi) - \Sigma(\phi)$ is independent of ϕ and the difference only affects the normalization of the field integral.

Schwinger–Dyson equations: An example. We consider the Euclidean action (omitting, for notational simplicity, the required regularization)

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} [\nabla_x \phi(x)]^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{g}{4!} \phi^4(x) \right].$$

Equation (7.29), in this example, reads

$$\left[(-\nabla_x^2 + m^2) \frac{\delta}{\delta J(x)} + \frac{g}{3!} \left(\frac{\delta}{\delta J(x)} \right)^3 - J(x) \right] \mathcal{Z}(J) = 0.$$

Explicit equations can be obtained by expanding in powers of $J(x)$. For example, differentiating once with respect to $J(y)$ and setting $J = 0$, one finds

$$(-\nabla_x^2 + m^2) \langle \phi(x) \phi(y) \rangle + \frac{g}{3!} \langle \phi(x)^3 \phi(y) \rangle = \delta^{(d)}(x - y).$$

Differentiating thrice and setting $J = 0$, one obtains

$$\begin{aligned} & (-\nabla_x^2 + m^2) \langle \phi(x) \phi(y_1) \phi(y_2) \phi(y_3) \rangle + \frac{g}{3!} \langle \phi(x)^3 \phi(y_1) \phi(y_2) \phi(y_3) \rangle \\ &= \delta^{(d)}(x - y_1) \langle \phi(y_2) \phi(y_3) \rangle + 2 \text{ terms obtained by permutation of 123}. \end{aligned}$$

More generally, these equations relate the $2n$ -, $(2n+2)$ - and $(2n+4)$ -point functions. They can be solved by expanding all functions in powers of g , and this leads to perturbation theory. One can also try solving these equations in a non-perturbative, but approximate, way by truncating, in some form, the infinite set.

7.5.2 Direct algebraic proof of the quantum field equations

We now show with purely algebraic transformations that $\mathcal{Z}(J)$ defined by equation (7.17) indeed satisfies equation (7.29) which, for the explicit form (7.7) of $\mathcal{S}(\phi)$, reads

$$\left[\int d^d y K(x - y) \frac{\delta}{\delta J(y)} + \frac{\delta \mathcal{V}_I}{\delta \phi(x)} (D_J) - J(x) \right] \mathcal{Z}(J) = 0. \quad (7.33)$$

We want thus to prove algebraically

$$\left[\int d^d y K(x - y) \frac{\delta}{\delta J(y)} + \frac{\delta \mathcal{V}_I(D_J)}{\delta \phi(x)} - J(x) \right] \exp [-\mathcal{V}_I(D_J)] \exp [\frac{1}{2}(J \Delta J)] = 0. \quad (7.34)$$

We first observe that the identity is true in the Gaussian theory, since, as a consequence of the definition (7.11) of Δ , one infers

$$\left[\int d^d y K(x - y) \frac{\delta}{\delta J(y)} - J(x) \right] \exp [\frac{1}{2}(J \Delta J)] = 0. \quad (7.35)$$

We then act with the operator $\exp(-\mathcal{V}_I(D_J))$ on the left of equation (7.35).

- (i) The operator commutes with KD_J .
- (ii) We use the commutation relation

$$[F(D_J), J(x)] = \frac{\delta F(D_J)}{\delta \phi(x)}, \quad (7.36)$$

which can easily be verified by expanding F in powers of D_J . Applied to the functional

$$F(\phi) = \exp[-\mathcal{V}_I(\phi)],$$

the commutation relation implies

$$\exp(-\mathcal{V}_I(D_J))J(x) = J(x)\exp(-\mathcal{V}_I(D_J)) - \frac{\delta \mathcal{V}_I(D_J)}{\delta \phi(x)}\exp(-\mathcal{V}_I(D_J)). \quad (7.37)$$

This completes the proof.

7.5.3 The infinitesimal change of variables

Various identities satisfied by correlation functions in the case of field theories possessing symmetries can be proved by the method of infinitesimal change of variables.

We change variables $\phi(x) \mapsto \chi(x)$ in a field integral, setting

$$\phi(x) = \chi(x) + \varepsilon F(x; \chi), \quad (7.38)$$

in which ε is an infinitesimal parameter, and $F(x; \chi)$ a general functional of χ ,

$$F(x; \chi) = \sum_1^\infty \frac{1}{n!} \int d^d y_1 \cdots d^d y_n \chi(y_1) \cdots \chi(y_n) F^{(n)}(x; y_1, \dots, y_n). \quad (7.39)$$

The variation of the action $\mathcal{S}(\phi)$ is

$$\mathcal{S}(\phi) - \mathcal{S}(\chi) = \varepsilon \int d^d x \frac{\delta \mathcal{S}}{\delta \chi(x)} F(x; \chi) + O(\varepsilon^2). \quad (7.40)$$

The change of variables (7.38) in the field integral generates the Jacobian

$$\mathcal{J} = \det \frac{\delta \phi(x)}{\delta \chi(y)} = \det \left[\delta^{(d)}(x-y) + \varepsilon \frac{\delta F(x; \chi)}{\delta \chi(y)} \right]. \quad (7.41)$$

As a consequence of identity (A2.4),

$$\det(1 + \varepsilon M) = 1 + \varepsilon \operatorname{tr} M + O(\varepsilon^2),$$

one obtains

$$\mathcal{J} = 1 + \varepsilon \int d^d x \frac{\delta F(x; \chi)}{\delta \chi(x)} + O(\varepsilon^2). \quad (7.42)$$

It follows that

$$\begin{aligned} \mathcal{Z}(J) = & \int [d\chi] \left(1 + \varepsilon \int d^d x \frac{\delta F(x; \chi)}{\delta \chi(x)} \right) \left[1 - \varepsilon \int d^d x \frac{\delta \mathcal{S}}{\delta \chi(x)} F(x; \chi) \right. \\ & \left. + \varepsilon \int d^d x J(x) F(x; \chi) \right] \exp [-\mathcal{S}(\chi) + J \cdot \chi] + O(\varepsilon^2). \end{aligned} \quad (7.43)$$

The term of order ε^0 is $\mathcal{Z}(J)$ itself. The terms of order ε thus must cancel. Collecting them, replacing χ by D_J when appropriate, one obtains the identity

$$\int d^d x \left[F(x; D_J) \frac{\delta \mathcal{S}(D_J)}{\delta \chi(x)} - \frac{\delta F(x; D_J)}{\delta \chi(x)} - J(x) F(x; D_J) \right] \mathcal{Z}(J) = 0. \quad (7.44)$$

Algebraic proof. The algebraic proof of the identity relies on acting with the differential operator $\int d^d x F(x; D_J)$ on the field equation (7.29),

$$\left[\frac{\delta \mathcal{S}(D_J)}{\delta \phi(x)} - J(x) \right] \mathcal{Z}(J) = 0.$$

Equation (7.44) then follows immediately from the commutation relation (7.36) used in the form

$$[F(y; D_J), J(x)] = \frac{\delta F(y; D_J)}{\delta \chi(x)}.$$

Note that we have dealt, in this section, with an infinitesimal change of variables at first order in ε , although the algebraic proof can be extended to all orders in ε .

7.5.4 The choice of the Gaussian measure

We want to show that, provided the sums of some geometric series exist, a part of the quadratic term K can be treated as a perturbation.

We thus decompose the kernel K in expression (7.7) into a sum of two terms:

$$K = K_1 + K_2, \quad \text{with } K_1 > 0.$$

We want to prove algebraically that

$$\begin{aligned} & \exp [-\mathcal{V}_I(D_J) - \frac{1}{2}(D_J K_2 D_J)] \exp [\frac{1}{2}(J K_1^{-1} J)] \\ &= \mathcal{N}(K_1, K_2) \exp [-\mathcal{V}_I(D_J)] \exp \left[\frac{1}{2}(J (K_1 + K_2)^{-1} J) \right], \end{aligned} \quad (7.45)$$

where \mathcal{N} is independent of the source J . Since the operator $\exp [-\mathcal{V}_I(\delta/\delta J)]$ can be factorized in both sides of equation (7.45), it is sufficient to prove the identity for $\mathcal{V}_I = 0$, that is, to calculate the functional

$$\mathcal{Z}_G(J) = \exp [-\frac{1}{2}(D_J K_2 D_J)] \exp [\frac{1}{2}(J K_1^{-1} J)]. \quad (7.46)$$

We act with $K_1 D_J$ on $\mathcal{Z}_G(J)$:

$$\int d^d y K_1(x - y) \frac{\delta}{\delta J(y)} \mathcal{Z}_G(J) = \exp [-\frac{1}{2}(D_J K_2 D_J)] J(x) \exp [\frac{1}{2}(J K_1^{-1} J)]. \quad (7.47)$$

In the right-hand side, we then commute J to bring it to the left (equation (7.36)):

$$\exp\left[-\frac{1}{2}(D_J K_2 D_J)\right] J(x) = J(x) \exp\left[-\frac{1}{2}(D_J K_2 D_J)\right] - (K_2 D_J)(x) \exp\left[-\frac{1}{2}(D_J K_2 D_J)\right].$$

We conclude that \mathcal{Z}_G satisfies the equation

$$\left[\int d^d y (K_1 + K_2)(x-y) \frac{\delta}{\delta J(y)} - J(x) \right] \mathcal{Z}_G(J) = 0. \quad (7.48)$$

Integrating the equation, one finds

$$\mathcal{Z}_G(J) = \mathcal{N}(K_1, K_2) \exp\left[\frac{1}{2}(J(K_1 + K_2)^{-1} J)\right],$$

proving identity (7.45). After some additional algebra, one verifies that $\mathcal{N}^2 = \det(1 + K_2 K_1^{-1})$.

In the same way, one can show that a part of the source term can be treated as an interaction, without changing $\mathcal{Z}(J)$. The result follows from the identity

$$\exp\left[\int d^d x L(x) \frac{\delta}{\delta J(x)}\right] F(J) = F(J + L).$$

All these identities make sense only if both sides exist.

7.5.5 The functional Dirac δ -function

We consider, here, a field $\phi(x)$ with N components $\phi_i(x)$, $i = 1, \dots, N$, satisfying a constraint of the form

$$F(x, \phi) = 0, \quad (7.49)$$

where F is expandable in powers of ϕ . We assume that it is possible to solve the constraint and calculate one component, for example $\phi_N(x)$, as a formal power series in the remaining components.

We then define the functional Dirac δ -function by

$$\delta(F) \equiv \int [d\lambda(x)] \exp\left[\int d^d x \lambda(x) F(x, \phi)\right], \quad (7.50)$$

where the λ -integration runs along the imaginary axis. This is a generalized Fourier representation. For reasons explained before, we have omitted any normalization factor.

We now show that representation (7.50) has the properties expected from a δ -function. We consider a field integral in which the integration is restricted to fields ϕ satisfying the constraint $F(x, \phi) = 0$:

$$\mathcal{Z}_F(\mathbf{J}) = \int \prod_{i=1}^N [d\phi_i(x)] \delta(F) \exp(-\mathcal{S}(\phi) + \mathbf{J} \cdot \phi). \quad (7.51)$$

We use the representation (7.50) to write \mathcal{Z}_F as

$$\mathcal{Z}_F(\mathbf{J}) = \int [d\lambda] \prod_{i=1}^N [d\phi_i] \exp[-\mathcal{S}(\phi) + \lambda \cdot F(\phi) + \mathbf{J} \cdot \phi]. \quad (7.52)$$

We now assume that the quadratic part in the fields (ϕ, λ) in the total action is not singular in such a way that, after adding a source term for $\lambda(x)$, we can use the equivalent of equation (7.17) to define, algebraically, integral (7.52) (it is sufficient, in particular, for the quadratic part of $\mathcal{S}(\phi)$ to be non-singular). Then, if we add a term proportional to $F(x, \phi)$ to $\mathcal{S}(\phi)$,

$$\mathcal{S}(\phi) \mapsto \mathcal{S}(\phi) + \int \mu(x) F(x, \phi) d^d x, \quad (7.53)$$

we can cancel the change by the change of variables $\lambda(x) \mapsto \lambda(x) + \mu(x)$. We have actually proven invariance by change of variables only for the infinitesimal change of variables, but for translations the property can be generalized to finite changes.

Finally, note that with the field $\lambda(x)$ is associated a simple field equation. From the identity

$$\int \prod_i [d\phi_i] [d\lambda] \frac{\delta}{\delta \lambda(x)} \exp [\lambda \cdot F + \mathbf{J} \cdot \phi - \mathcal{S}(\phi)] = 0,$$

one derives

$$F(x; D_J) \mathcal{Z}_F(\mathbf{J}) = 0. \quad (7.54)$$

This equation expresses the constraint $F = 0$ on the generating functional of correlation functions $\mathcal{Z}_F(J)$.

Examples. In Section 19.9, a functional δ -function in the form (7.50) is used to express the non-linear σ model in terms of an N -vector field ϕ satisfying an $O(N)$ invariant constraint $\phi^2(x) = 1$.

In Chapter 21, one covariant gauge corresponds to constraining the gauge field A_μ to satisfy the condition $\sum_\mu \partial_\mu A_\mu = 0$. Again, the representation (7.50) is useful. The equation (7.54) then becomes $\sum_\mu \partial_\mu^\alpha \delta \mathcal{Z} / \delta J_\mu(x) = 0$, where J_μ is the gauge field source.

7.6 Connected correlation functions. Cluster properties

We now consider *local* Euclidean actions $\mathcal{S}(\phi)$ functions of a scalar field $\phi(x)$ of the form

$$\mathcal{S}(\phi) = \int d^d x \mathcal{L}(\phi; x), \quad (7.55)$$

where the Euclidean Lagrangian density $\mathcal{L}(\phi; x)$ is a function of the field $\phi(x)$ and its derivatives and does not depend on space explicitly, but only through ϕ (as in the example (6.1) after continuation to Euclidean time).

We have already observed in quantum mechanics that the partition function $\text{tr } e^{-\beta H}$ in the large volume limit, that is, for $\beta \rightarrow \infty$, has an exponential behaviour; more precisely (equation (2.1)),

$$\lim_{\beta \rightarrow \infty} -\frac{1}{\beta} \ln \text{tr } e^{-\beta H} = E_0,$$

in which E_0 is the ground state energy. Moreover, the convergence towards the limit is exponential when the ground state is isolated (equation (2.2)).

Here, we generalize the quantum mechanics result to local field theories. The discussion that follows is somewhat intuitive and tries only to motivate results that can be proven rigorously. Some peculiarities found in massless theories or due to ground state degeneracy are ignored here and discussed later, starting with Chapter 13.

The functional $\mathcal{Z}(J)$ is a partition function in the presence of an external source $J(x)$. We take for source $J(x)$ the sum of two terms $J_1(x)$ and $J_2(x)$, in which $J_1(x)$ and $J_2(x)$ have disjoint supports consisting of two domains Ω_1 and Ω_2 of large volumes V_1 and V_2 :

$$J(x) = J_1(x) + J_2(x), \quad \begin{cases} J_1(x) = 0 & \text{for } x \notin \Omega_1, \\ J_2(x) = 0 & \text{for } x \notin \Omega_2, \end{cases} \quad \Omega_1 \cap \Omega_2 = \emptyset. \quad (7.56)$$

We also assume that $J_1(x)$ and $J_2(x)$ fluctuate around an arbitrarily small but non-vanishing constant. Then the *locality* of $\mathcal{S}(\phi)$ implies the decomposition

$$\begin{aligned} \mathcal{S}(\phi) - \int d^d x J(x)\phi(x) &= \int_{x \in \Omega_1} d^d x [\mathcal{L}(\phi; x) - J_1(x)\phi(x)] \\ &+ \int_{x \in \Omega_2} d^d x [\mathcal{L}(\phi; x) - J_2(x)\phi(x)] + \int_{x \notin \Omega_1 \cup \Omega_2} d^d x \mathcal{L}(\phi; x) \\ &+ \text{contributions from boundaries}. \end{aligned} \quad (7.57)$$

We then write $\mathcal{Z}(J)$ as

$$\mathcal{Z}(J) = \mathcal{Z}_1(J_1)\mathcal{Z}_2(J_2)\mathcal{Z}_{12}(J_1, J_2), \quad (7.58)$$

with the definitions,

$$\begin{aligned} \mathcal{Z}_1(J_1) &= \int_{x \in \Omega_1} [d\phi(x)] \exp \left[-\mathcal{S}(\phi) + \int d^d x J_1(x)\phi(x) \right], \\ \mathcal{Z}_2(J_2) &= \int_{x \in \Omega_2} [d\phi(x)] \exp \left[-\mathcal{S}(\phi) + \int d^d x J_2(x)\phi(x) \right]. \end{aligned}$$

Both functionals \mathcal{Z}_1 and \mathcal{Z}_2 are normalized to 1 for $J_1 = 0$ or $J_2 = 0$, respectively. The functional \mathcal{Z}_{12} is defined by equation (7.58). Its dependence in J_1 and J_2 comes entirely from the existence of boundary terms in equation (7.57). When we scale up Ω_1 and Ω_2 , these boundary terms grow like surfaces while the first two terms in equation (7.57) grow like the volumes V_1 and V_2 . Therefore, $\ln \mathcal{Z}_{12}(J_1, J_2)$ becomes asymptotically negligible compared to $\ln \mathcal{Z}_1(J_1)$ and $\ln \mathcal{Z}_2(J_2)$ when V_1 and $V_2 \rightarrow \infty$.

To express this property, it is natural to introduce the functional $\mathcal{W}(J) = \ln \mathcal{Z}(J)$ (equation (7.19)), which then satisfies

$$\mathcal{W}(J_1 + J_2) = \mathcal{W}_1(J_1) + \mathcal{W}_2(J_2) + \text{negligible}. \quad (7.59)$$

In particular, if $\mathcal{S}(\phi) - J\phi$ is translation invariant (which implies that J is a constant), $\mathcal{W}(J)$ is extensive, that is, proportional to the total volume. This property generalizes property (2.1).

Cluster properties. After the infinite volume limit has been taken, one can expand $\mathcal{W}(J_1 + J_2)$ in powers of J_1 and J_2 :

$$\begin{aligned} \mathcal{W}(J_1 + J_2) &= \sum_{0 \leq p \leq n}^{\infty} \frac{1}{p!(n-p)!} \int d^d x_1 \cdots d^d x_p d^d y_{p+1} \cdots d^d y_n \\ &\times W^{(n)}(x_1, \dots, x_p, y_{p+1}, \dots, y_n) J_1(x_1) \cdots J_1(x_p) J_2(y_{p+1}) \cdots J_2(y_n), \end{aligned} \quad (7.60)$$

with $x_i \in \Omega_1$, $y_j \in \Omega_2$.

The property (7.59) implies that all terms with $p \neq 0$ or $p \neq n$ are negligible for V_1 and V_2 large. Considering expression (7.60), we note that this implies that the functions $W^{(n)}$ must decrease rapidly enough when the two non-empty sets of points $\{x_1, \dots, x_p\}$ and $\{y_{p+1}, \dots, y_n\}$ have large separations. More precisely,

$$W^{(n)}(x_1, \dots, x_p, y_{p+1}, \dots, y_n) \rightarrow 0 \quad \text{for} \quad \min_{\substack{i=1 \dots p \\ j=p+1 \dots n}} |x_i - y_j| \rightarrow \infty. \quad (7.61)$$

This property, which we describe here only qualitatively, is called the *cluster property*, and is a *characteristic property of the connected correlation functions* generated by the functional $\mathcal{W}(J)$ (see Section A7.3 for details).

Feynman diagrams. We have seen that a Feynman diagram that is disconnected in the sense of graphs can be factorized into a product of the form

$$F_1(x_1, \dots, x_p) F_2(y_1, \dots, y_q).$$

In a translation invariant theory, when the two set of points $\{x_i\}$ and the set $\{y_i\}$ are not empty, we can separate the two sets in a way which leaves such a product invariant: a disconnected diagram cannot satisfy the cluster property. We recover the property that the Feynman diagrams which contribute to the perturbative expansion of $\mathcal{W}(J)$ are all connected. Furthermore, it can be verified that, in a field theory containing only massive fields, connected Feynman diagrams decrease exponentially, when points are separated, with a minimal rate which is the inverse of the smallest mass in the theory. This property is a consequence of the exponential decrease of the propagator (Section A7.3).

7.7 Legendre transformation. Vertex functions

We introduce a new generating functional, $\Gamma(\varphi)$, Legendre transform (see Section 1.8) of the generating functional of connected correlation functions $\mathcal{W}(J)$, called the generating functional of *vertex functions* also called proper vertices or 1PI correlation functions for reasons that are explained in Section 7.10.

Legendre transformation. In the context of statistical physics and phase transitions, it is natural to consider the thermodynamic potential, Legendre transform of the *free energy* $\mathcal{W}(J)$ and generator of vertex functions. Here, vertex functions are introduced, because they have special properties from the point of view of perturbation theory, as we demonstrate in Sections 7.9 and 7.10. The generating functional $\Gamma(\varphi)$ of vertex functions, in which $\varphi(x)$ is a classical field argument of Γ , is related to $\mathcal{W}(J)$ by

$$\Gamma(\varphi) + \mathcal{W}(J) - \int d^d x J(x) \varphi(x) = 0, \quad (7.62)$$

with

$$\varphi(x) = \frac{\delta \mathcal{W}}{\delta J(x)}. \quad (7.63)$$

We have described some properties of the Legendre transformation in Section 1.8 and shown that it is involutive. In particular,

$$J(x) = \frac{\delta \Gamma}{\delta \varphi(x)}. \quad (7.64)$$

Moreover, if $\mathcal{W}(J)$ depends on a parameter v , then,

$$\frac{\partial \mathcal{W}}{\partial v} + \frac{\partial \Gamma}{\partial v} = 0. \quad (7.65)$$

This identity, derived here for one external parameter, obviously applies also for an external field or source. It will be used frequently.

Expansion of $\Gamma(\varphi)$. If we set $J = 0$ in equation (7.63), we obtain

$$\varphi(x)|_{J=0} = \left. \frac{\delta \mathcal{W}}{\delta J(x)} \right|_{J=0} \equiv W^{(1)}(x) = \langle \phi(x) \rangle,$$

that is, that $\varphi(x)$ for a vanishing source is the expectation value of the field ϕ . Moreover, equation (7.64) then implies that the expectation value of $\phi(x)$ is an extremum of $\Gamma(\varphi)$.

Inverting the relation (7.63), we can expand the source $J(x)$ as a series of powers of

$$\xi(x) = \varphi(x) - W^{(1)}(x) = \varphi(x) - \langle \phi(x) \rangle. \quad (7.66)$$

The classical field $\xi(x)$ is related to the correlation functions of the field

$$\Xi(x) = \phi(x) - \langle \phi(x) \rangle, \quad (7.67)$$

which has a vanishing expectation value, $\langle \Xi(x) \rangle = 0$.

The first terms of the expansion of equation (7.63) are

$$\xi(x) = \int d^d x_1 W^{(2)}(x, x_1) J(x_1) + \frac{1}{2!} \int d^d x_1 d^d x_2 W^{(3)}(x, x_1, x_2) J(x_1) J(x_2) + \dots \quad (7.68)$$

We introduce the inverse $S(x, y)$ of the connected two-point function:

$$\int d^d z S(x, z) W^{(2)}(z, y) = \delta^{(d)}(x - y). \quad (7.69)$$

We can then write the expansion of $J(x)$ as

$$\begin{aligned} J(x) &= \int d^d x_1 S(x, x_1) \xi(x_1) - \frac{1}{2!} \int d^d y_1 d^d y_2 d^d y_3 d^d x_1 d^d x_2 S(x, y_3) \\ &\quad \times S(x_1, y_1) S(x_2, y_2) W^{(3)}(y_1, y_2, y_3) \xi(x_1) \xi(x_2) + \dots \end{aligned} \quad (7.70)$$

The expansion can be expressed in terms of so-called amputated correlation functions

$$W_{\text{amp.}}^{(n)}(x_1, \dots, x_n) = \int \left[\prod_{i=1}^n d^d y_i S(x_i, y_i) \right] W^{(n)}(y_1, \dots, y_n). \quad (7.71)$$

In terms of Feynman diagrams, this means that propagators and contributions to the two-point function on external lines are cancelled.

One can use equation (7.64) to calculate $\Gamma(\varphi)$. We set

$$\Gamma(\varphi) = \sum_1^\infty \frac{1}{n!} \int d^d x_1 \dots d^d x_n \Gamma^{(n)}(x_1, \dots, x_n) \xi(x_1) \dots \xi(x_n), \quad (7.72)$$

where the $\Gamma^{(n)}$ are associated with the field Ξ (defined by equation (7.67)).

One finds

$$\begin{aligned}\Gamma^{(1)}(x) &= 0, \\ \Gamma^{(2)}(x_1, x_2) &= S(x_1, x_2) = [W^{(2)}]^{-1}(x_1, x_2), \\ \Gamma^{(3)}(x_1, x_2, x_3) &= -W_{\text{amp.}}^{(3)}(x_1, x_2, x_3), \\ \Gamma^{(4)}(x_1, x_2, x_3, x_4) &= -W_{\text{amp.}}^{(4)}(x_1, x_2, x_3, x_4) \\ &\quad + \int d^d y d^d z W_{\text{amp.}}^{(3)}(x_1, x_2, y) W^{(2)}(y, z) W_{\text{amp.}}^{(3)}(z, x_3, x_4) + 2 \text{ terms} \\ &\quad \dots\end{aligned}$$

The inverse relations are even more useful, because, as we show in Section 7.10, $\Gamma(\varphi)$ has simpler properties than $\mathcal{W}(J)$. For example,

$$\begin{aligned}W^{(2)}(x_1, x_2) &= [\Gamma^{(2)}]^{-1}(x_1, x_2), \\ W_{\text{amp.}}^{(3)}(x_1, x_2, x_3) &= -\Gamma^{(3)}(x_1, x_2, x_3), \\ W_{\text{amp.}}^{(4)}(x_1, x_2, x_3, x_4) &= -\Gamma^{(4)}(x_1, x_2, x_3, x_4), \\ &\quad + \int d^d y d^d z \Gamma^{(3)}(x_1, x_2, y) W^{(2)}(y, z) \Gamma^{(3)}(z, x_3, x_4) + 2 \text{ terms}, \\ &\quad \dots\end{aligned}$$



Fig. 7.4 Graphical representation of the two-point function and the n -point vertex function

With the graphical definitions of Fig. 7.4, we can give a graphical representation of the first equations. The correlation functions $W^{(3)}$ and $W^{(4)}$ are then be represented as shown in Figs. 7.5 and 7.6, respectively.

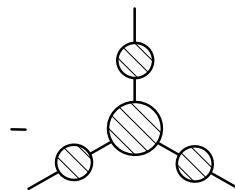


Fig. 7.5 The connected three-point function $W^{(3)}$

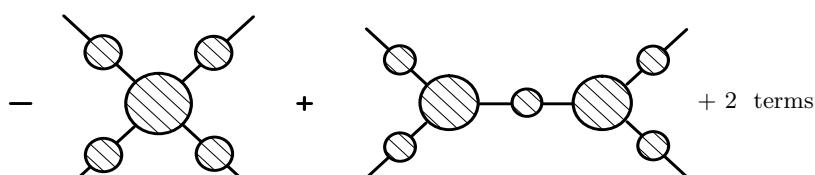


Fig. 7.6 The connected four-point function $W^{(4)}$

Mass operator. It follows from the set of relations between connected correlation functions and vertex functions that Feynman diagrams that contribute to vertex functions appear in the expansion of connected functions with the opposite sign except in the case of the two-point function. Indeed, if we set (in the notation (7.7))

$$\Gamma^{(2)}(x, y) = K(x - y) + \Sigma(x, y),$$

where K is the two-point function $\Gamma^{(2)}$ in the free (Gaussian) theory, then $\Sigma(x, y)$, called the mass operator, contains all the perturbative corrections. The connected two-point function $W^{(2)}(x, y)$ then takes the form of a geometric series of the form,

$$W^{(2)}(x, y) = \Delta(x - y) - \int d^d z_1 d^d z_2 \Delta(x - z_1) \Sigma(z_1, z_2) \Delta(z_2 - y) + \dots,$$

where Δ is the propagator (7.11). In this case, K , the first term in perturbation theory contributing to $\Gamma^{(2)}$, is special from the point of view of Feynman graph expansion, since it has the same sign as the propagator in $W^{(2)}$.

Convexity. The connected two-point function can be written as

$$W^{(2)}(x, y) = \langle [\phi(x) - \langle \phi(x) \rangle] [\phi(y) - \langle \phi(y) \rangle] \rangle = \langle \Xi(x) \Xi(y) \rangle.$$

Setting $(J \cdot \Xi) \equiv \int d^d x J(x) \Xi(x)$, we immediately obtain

$$\int d^d x d^d y J(x) J(y) W^{(2)}(x, y) = \langle (J \cdot \Xi)^2 \rangle \geq 0.$$

It follows that the two-point function $W^{(2)}(x, y)$ is the kernel of a positive operator. Since $\Gamma^{(2)}(x, y)$ is the inverse of $W^{(2)}(x, y)$ in the sense of kernels, it is also the kernel of a positive operator,

$$\int d^d x d^d y \varphi(x) \Gamma^{(2)}(x, y) \varphi(y) \geq 0. \quad (7.73)$$

For the same reasons, $\delta^2 \mathcal{W}(J)/\delta J(x)\delta J(y)$ and $\delta^2 \Gamma(\varphi)/\delta \varphi(x)\delta \varphi(y)$, which are the two-point functions in an external source, are positive operators. In particular, in the case of constant fields J or φ , $\mathcal{W}(J)$ and $\Gamma(\varphi)$, both divided by the total space volume, are convex functions of the sources. We shall recall this property when we examine the physics of spontaneous symmetry breaking and meet functions $\Gamma(\varphi)$ that do not seem obviously convex (see Section 7.11).

7.8 Momentum representation

In this work, we mainly study theories invariant under space translations. Correlation functions then depend only on differences of space arguments. It is thus natural to introduce their Fourier representation. To establish a consistent set of conventions, we start from the generating functional of vertex functions (in d space dimensions)

$$\Gamma(\varphi) = \sum_n \frac{1}{n!} \int \prod_{i=1}^n d^d x_i \varphi(x_i) \Gamma^{(n)}(x_1, \dots, x_n).$$

We introduce the Fourier components of the field,

$$\varphi(x) = \int e^{ipx} \tilde{\varphi}(p) d^d p. \quad (7.74)$$

We define the vertex function $\tilde{\Gamma}^{(n)}(p_1, \dots, p_n)$ in the Fourier (or momentum) representation in terms of the coefficient of $\tilde{\varphi}(p_1) \cdots \tilde{\varphi}(p_n)$ in $\Gamma(\varphi)$, assuming translation invariance, which implies total momentum conservation,

$$(2\pi)^d \delta^{(d)}(p_1 + \cdots + p_n) \tilde{\Gamma}^{(n)}(p_1, \dots, p_n) = \int \left(\prod_k d^d x_k e^{ix_k p_k} \right) \Gamma^{(n)}(x_1, \dots, x_n). \quad (7.75)$$

In the same way, we introduce the Fourier components $\tilde{J}(p)$ of the source,

$$J(x) = \int e^{ipx} \tilde{J}(p) d^d p.$$

We insert this representation into the generating functional $\mathcal{W}(J)$ of connected correlation functions and define $\tilde{W}^{(n)}(p_1, \dots, p_n)$ in terms of the coefficient of $\tilde{J}(p_1) \cdots \tilde{J}(p_n)$:

$$(2\pi)^d \delta^{(d)}(p_1 + \cdots + p_n) \tilde{W}^{(n)}(p_1, \dots, p_n) = \int \left(\prod_k d^d x_k e^{ix_k p_k} \right) W^{(n)}(x_1, \dots, x_n). \quad (7.76)$$

Inverting the Fourier transformation, we find for the various two-point functions,

$$\begin{aligned} \Gamma^{(2)}(x, y) &= \frac{1}{(2\pi)^d} \int d^d p e^{ip(x-y)} \tilde{\Gamma}^{(2)}(-p, p), \\ W^{(2)}(x, y) &= \frac{1}{(2\pi)^d} \int d^d p e^{ip(x-y)} \tilde{W}^{(2)}(-p, p). \end{aligned}$$

In what follows, we denote simply by $\tilde{\Gamma}^{(2)}(p)$ and $\tilde{W}^{(2)}(p)$, the functions

$$\tilde{\Gamma}^{(2)}(p) = \tilde{\Gamma}^{(2)}(-p, p), \quad \tilde{W}^{(2)}(p) = \tilde{W}^{(2)}(-p, p).$$

The Legendre transformation takes the simple form

$$\tilde{\Gamma}^{(2)}(p) \tilde{W}^{(2)}(p) = 1. \quad (7.77)$$

The explicit expressions of Section 7.7 (*e.g.* equation (7.71)) then show that amputation and Legendre transformation become in the Fourier representation purely algebraic operations in the sense that no momentum integration is involved, because the two-point function is proportional to $\delta^{(d)}(p_1 + p_2)$. For example:

$$\tilde{W}_{\text{amp.}}^{(n)}(p_1, \dots, p_n) = \tilde{W}(p_1, \dots, p_n) \prod_{i=1}^n \left[\tilde{W}^{(2)}(p_i) \right]^{-1},$$

and also

$$\begin{aligned} \tilde{\Gamma}^{(3)}(p_1, p_2, p_3) &= -\tilde{W}_{\text{amp.}}^{(3)}(p_1, p_2, p_3), \\ \tilde{\Gamma}^{(4)}(p_1, p_2, p_3, p_4) &= -\tilde{W}_{\text{amp.}}^{(4)}(p_1, p_2, p_3, p_4) + \left[\tilde{W}_{\text{amp.}}^{(3)}(p_1, p_2, -p_1 - p_2) \tilde{W}^{(2)}(p_1 + p_2) \right. \\ &\quad \times \left. \tilde{W}_{\text{amp.}}^{(3)}(p_3, p_4, -p_3 - p_4) + \text{cyclic permutation of } \{p_2, p_3, p_4\} \right], \\ &\quad \dots \end{aligned}$$

This remark will be used in the discussion of divergences in perturbation theory.

7.9 Loop or semi-classical expansion

It is sometimes useful to reorganize perturbation theory by grouping some classes of Feynman diagrams. An important example is provided by the *loop expansion* [32]. We set \hbar in front of the classical action and the source term and consider the generating functional

$$\mathcal{Z}(J) = \mathcal{N} \int [d\phi] \exp \left[-\frac{1}{\hbar} (\mathcal{S}(\phi) - J \cdot \phi) \right], \quad (7.78)$$

($\mathcal{Z}(0) = 1$), in the symbolic notation of Section 7.2.2. For $\hbar \rightarrow 0$, the field integral can be calculated by the *steepest descent method* (see Section 1.3). The successive corrections to the leading order approximation generate an expansion in powers of \hbar . We show below that this expansion gathers Feynman diagrams according to the number of loops.

Because it is an expansion in powers of \hbar , the expansion is also called *semi-classical*, although classical field equations correspond to a true classical limit only for massless theories like quantum electrodynamics.

7.9.1 Loop expansion at leading order

The saddle point equation is

$$\frac{\delta \mathcal{S}}{\delta \phi(x)} [\phi_c(J)] = J(x). \quad (7.79)$$

We assume in the following calculations that the field ϕ has been defined in such a way that $\phi_c(J=0) = 0$ and that $S(\phi=0) = 0$.

Substituting the solution $\phi_c(J)$ into the classical action, one obtains $\mathcal{Z}(J)$ at leading order,

$$\ln \mathcal{Z}(J) \sim \ln \mathcal{Z}_0(J) \equiv \frac{1}{\hbar} [-\mathcal{S}(\phi_c) + J \cdot \phi_c]. \quad (7.80)$$

When \hbar is explicit, it is convenient to define $\mathcal{W}(J)$ by

$$\mathcal{W}(J) = \hbar \ln \mathcal{Z}(J). \quad (7.81)$$

Then, at leading order,

$$\mathcal{W}(J) = \mathcal{W}_0(J) \equiv -\mathcal{S}(\phi_c) + J \cdot \phi_c. \quad (7.82)$$

Together, the two equations (7.79) and (7.82) imply that the two functionals $\mathcal{S}(\phi)$ and $\mathcal{W}_0(J)$ are related by a Legendre transformation (see also Section 1.8).

Perturbation theory. Ordinary perturbation theory is recovered by expanding the solution $\phi_c(J)$ in powers of J .

To show this, we decompose $\mathcal{S}(\phi)$ into a sum of a quadratic part and interaction terms (expression (7.7)). Moreover, we assume

$$\mathcal{V}_I(\phi) = O(\phi^3) \quad \text{for } \phi \rightarrow 0.$$

Equation (7.79) then takes the symbolic form

$$K\phi_c + \frac{\delta \mathcal{V}_I(\phi_c)}{\delta \phi} = J. \quad (7.83)$$

It can be solved iteratively as ($K\Delta = 1$)

$$\phi_c = \Delta J - \Delta \frac{\delta \mathcal{V}_I(\phi_c)}{\delta \phi} = \Delta J - \Delta \frac{\delta \mathcal{V}_I(\Delta J)}{\delta \phi} + \dots. \quad (7.84)$$

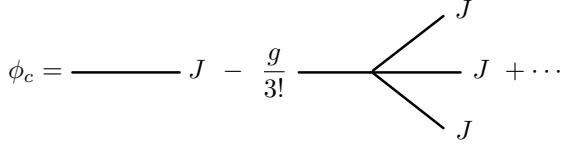


Fig. 7.7 Feynman-diagram expansion of ϕ_c in the ϕ^4 field theory

If, for example,

$$\mathcal{V}_I(\phi) = \frac{g}{4!} \int d^d x \phi^4(x),$$

ϕ_c has the Feynman diagram expansion displayed in Fig. 7.7. We observe that only trees [33], that is, diagrams without loops, are generated. Substituting the expansion into equation (7.82), we note that the perturbative expansion of $\mathcal{W}_0(J)$ in powers of J also contains only connected tree Feynman diagrams. The functional $\mathcal{W}_0(J)$ is the generating functional of connected tree diagrams.

Vertex functions. Since the action $\mathcal{S}(\phi)$ and $\mathcal{W}_0(J)$, the generating functional of connected tree diagrams are related by a Legendre transformation, we conclude immediately

$$\Gamma_0(\varphi) = \mathcal{S}(\varphi). \quad (7.85)$$

At leading order, $\Gamma(\varphi)$, the generating functional of vertex functions, is identical to the classical action. From the point of view of Feynman diagrams, it reduces to the vertices of the field theory.

7.9.2 Order \hbar or one-loop contributions

The Gaussian integral obtained by expanding around the saddle point generates the order \hbar corrections (Appendix B of Ref. [41] and Ref. [42]). We set

$$\phi(x) = \phi_c(J, x) + \sqrt{\hbar} \chi(x), \quad (7.86)$$

$$S^{(2)}(x_1, x_2; \phi) = \frac{\delta^2 \mathcal{S}}{\delta \phi(x_1) \delta \phi(x_2)}. \quad (7.87)$$

Expanding the action in powers of \hbar , one finds

$$\mathcal{S}(\phi) - J \cdot \phi = \mathcal{S}(\phi_c) - J \cdot \phi_c + \frac{\hbar}{2} \int d^d x_1 d^d x_2 S^{(2)}(x_1, x_2; \phi_c) \chi(x_1) \chi(x_2) + O(\hbar^{3/2}).$$

The field integral at this order becomes

$$\mathcal{Z}(J) \sim \mathcal{Z}_0(J) \int [d\chi] \exp \left[-\frac{1}{2} \int d^d x_1 d^d x_2 S^{(2)}(x_1, x_2; \phi_c) \chi(x_1) \chi(x_2) \right],$$

and, therefore,

$$\mathcal{Z}(J) \propto \mathcal{Z}_0(J) \left[\det S^{(2)}(x_1, x_2; \phi_c) / \det S^{(2)}(x_1, x_2; 0) \right]^{-1/2}, \quad (7.88)$$

where the normalization follows from the conditions $\phi_c(0) = 0$ and $\mathcal{Z}(0) = 1$.

At this order, the connected generating functional $\mathcal{W}(J)$ is then ($\ln \det = \text{tr} \ln$)

$$\mathcal{W}(J) = \mathcal{W}_0(J) + \hbar \mathcal{W}_1(J) + O(\hbar^2), \quad (7.89)$$

with

$$\mathcal{W}_1(J) = -\frac{1}{2} \left[\text{tr} \ln S^{(2)}(x_1, x_2; \phi_c) - \text{tr} \ln S^{(2)}(x_1, x_2; 0) \right]. \quad (7.90)$$

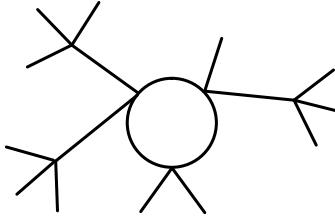


Fig. 7.8 One-loop connected contribution in the ϕ^4 field theory

We again illustrate the result with the example of the ϕ^4 interaction, where

$$S^{(2)}(x_1, x_2; \phi) = K(x_1 - x_2) + \frac{1}{2}g\phi^2(x_1)\delta^{(d)}(x_1 - x_2).$$

Then,

$$\text{tr ln } S^{(2)}(x_1, x_2; \phi_c) - \text{tr ln } S^{(2)}(x_1, x_2; 0) = \text{tr ln} \left[\delta^{(d)}(x_1 - x_2) + \frac{1}{2}g\Delta(x_1 - x_2)\phi_c^2(x_2) \right].$$

The expansion of $\mathcal{W}_1(J)$ in powers of ϕ_c takes the form

$$\begin{aligned} \mathcal{W}_1(J) = & -\frac{1}{2} \left[\frac{g}{2} \int d^d x_1 \Delta(x_1 - x_1) \phi_c^2(x_1) \right. \\ & \left. - \frac{g^2}{8} \int d^d x_1 d^d x_2 \Delta(x_1 - x_2) \phi_c^2(x_2) \Delta(x_2 - x_1) \phi_c^2(x_1) + \dots \right]. \end{aligned}$$

Note that the trace operation has generated a set of *one-loop Feynman diagrams*.

To recover perturbation theory, one still has to expand $\phi_c(J)$ in powers of J . A typical contribution to $\mathcal{W}_1(J)$ has the representation displayed in Fig. 7.8.

At order \hbar , the generating functional $\Gamma(\varphi)$ can be obtained from the relation

$$\frac{\partial \Gamma}{\partial \hbar} + \frac{\partial \mathcal{W}}{\partial \hbar} = 0,$$

for $\hbar \rightarrow 0$. Thus,

$$\Gamma(\varphi) = \mathcal{S}(\varphi) - \hbar \mathcal{W}_1(J) = \mathcal{S}(\varphi) + \frac{1}{2}\hbar \left[\text{tr ln } S^{(2)}(x_1, x_2; \phi_c) - \text{tr ln } S^{(2)}(x_1, x_2; 0) \right]. \quad (7.91)$$

At this order, we can replace $\phi_c(J)$ by φ . Setting

$$\Gamma(\varphi) = \mathcal{S}(\varphi) + \hbar \Gamma_1(\varphi) + O(\hbar^2), \quad (7.92)$$

we obtain the order \hbar correction,

$$\Gamma_1(\varphi) = \frac{1}{2} \text{tr} \left[\ln S^{(2)}(x_1, x_2; \varphi) - \ln S^{(2)}(x_1, x_2; 0) \right]. \quad (7.93)$$

If $\mathcal{S}(\phi)$ has the decomposition (7.7), and $\mathcal{V}_I(\phi)$ the local form,

$$\mathcal{V}_I(\phi) = \int d^d x V_I(\phi(x)),$$

the expansion of $\Gamma_1(\varphi)$ in powers of V_I takes the form,

$$\begin{aligned} \Gamma_1(\varphi) = & \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \int d^d x_1 \cdots d^d x_n V_I''(\varphi(x_1)) \Delta(x_1 - x_2) \\ & \times V_I''(\varphi(x_2)) \Delta(x_2 - x_3) \cdots V_I''(\varphi(x_n)) \Delta(x_n - x_1). \end{aligned} \quad (7.94)$$

In terms of Feynman diagrams, it is a sum of one-loop diagrams (see Fig. 7.9).

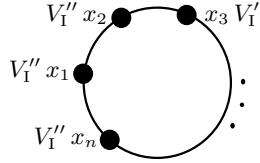


Fig. 7.9 Vertex functions: One-loop diagrams

The remarkable property of the Feynman graph expansion of the functional $\Gamma_1(\varphi)$ at this order is that all diagrams are one-line irreducible (1PI in the terminology of particle physics): they cannot be disconnected by cutting only one line. The functional $\Gamma_1(\varphi)$ is *the generating functional of vertex functions, which sum the one-line or 1PI one-loop Feynman diagrams.*

7.9.3 Loop expansion at higher orders

We now expand the field integral

$$\mathcal{Z}(J) = \int [d\phi] \exp \left[-\frac{1}{\hbar} (\mathcal{S}(\phi) - J \cdot \phi) \right], \quad (7.95)$$

with $\mathcal{S}(\phi)$ of the form (7.7), to all orders in perturbation theory. The propagator is $\hbar\Delta$. Any vertex generated by $\mathcal{V}_I(\phi)$ is multiplied by $1/\hbar$. In the same way, at the end of all external lines is attached a factor J which also yields a factor $1/\hbar$. Denoting by I the number of internal lines of a Feynman diagram (propagators which join two vertices), by E the number of external lines (propagator joining a vertex to a source J), by V the number of vertices, we find that the power of \hbar that multiplies a *connected* diagram (a contribution to $\mathcal{W}(J)$) is

$$\hbar^{I+E-(V+E)+1} = \hbar^{I-V+1},$$

the last factor \hbar coming from our normalization of $\mathcal{W}(J)$ (equation (7.81)).

We prove in the next section that the same result is obtained for 1PI Feynman diagrams (*i.e.* contributions to $\Gamma(\varphi)$), because the factor \hbar coming from the source cancels the factor coming from the external propagator.

We now show that the power of \hbar that we have found counts the number of loops of a diagram [43]. The number of loops is defined in the following way: if, by cutting a line of a connected diagram γ we obtain a new connected diagram γ' , then,

$$\# \text{ loops } \gamma = \# \text{ loops } \gamma' + 1.$$

From this definition follows a relation between the number of loops L , the number of internal lines I and the number of vertices V :

$$L = I - V + 1. \quad (7.96)$$

Indeed, we note that each time we can remove an internal line without disconnecting the diagram, we decrease I by 1 and L by 1. Eventually, we obtain a tree diagram, that is, a diagram in which no internal line can be cut without disconnecting the diagram. We then have to show that

$$I - V + 1 = 0.$$

From a tree diagram, we can remove systematically a vertex at the boundary with the line connecting it to the diagram until we obtain the simplest diagram, composed of a line joining two vertices, which satisfies the equation.

We have thus shown that the expansion in powers of \hbar reorganizes perturbation theory according to the number of loops of Feynman diagrams.

The number of loops is also the number of independent internal intensities in the corresponding electric circuit, the current being conserved at each vertex, the intensities flowing into the diagram being fixed.

Indeed, the number L of independent intensities is equal to the total number of intensities I minus the number of conservation equations ($V - 1$) (because one equation gives the total conservation of the current) and thus equation (7.96) is again satisfied. We will eventually use this remark to relate the number of loops to the number of independent momentum integration variables.

Higher order calculations. We now indicate how successive terms in the loop expansion can be calculated by applying the steepest descent method to the field integral (7.95).

The saddle point $\phi_c(J)$ is the solution of the equation

$$\frac{\delta \mathcal{S}}{\delta \phi(x)}(\phi_c) = J(x). \quad (7.97)$$

We change variables in the field integral (7.78), $\phi \mapsto \chi$, setting

$$\phi(x) = \phi_c(x) + \sqrt{\hbar}\chi(x). \quad (7.98)$$

We expand $\mathcal{S}(\phi)$ in powers of χ ,

$$\mathcal{S}(\phi_c + \chi) = \mathcal{S}(\phi_c) + \sum_{n=2}^{\infty} \frac{\hbar^{n/2}}{n!} \int \left(\prod_{i=1}^n d^d x_i \chi(x_i) \right) S^{(n)}(x_1, x_2, \dots, x_n; \phi_c). \quad (7.99)$$

where we have introduced the notation

$$\left. \frac{\delta^n \mathcal{S}(\phi)}{\delta \phi(x_1) \cdots \delta \phi(x_n)} \right|_{\phi=\phi_c} = S^{(n)}(x_1, \dots, x_n; \phi_c). \quad (7.100)$$

The expansion in powers of \hbar and the integration over χ generate vacuum Feynman diagrams with a ϕ_c -dependent propagator Δ solution of,

$$\int S^{(2)}(x, y; \phi_c) \Delta(y, z; \phi_c) d^d y = \delta^{(d)}(x - z), \quad (7.101)$$

and ϕ_c -dependent vertices $S^{(n)}$. At a finite order in \hbar , only a finite number of vertices contribute.

In this expansion, $\mathcal{W}(J)$ is the sum of all connected vacuum diagrams. To calculate $\Gamma(\varphi)$ we have to express ϕ_c and J in terms of φ . At leading order, $\phi_c = \varphi + O(\hbar)$. Since the Legendre transformation removes only the 1PI diagrams, we conclude that $\Gamma(\varphi) - \mathcal{S}(\varphi)$ is given by the sum of the 1PI vacuum diagrams, in which ϕ_c has been replaced by φ . We verify this property at the two-loop order in Section A7.1.

7.10 Vertex functions: One-line irreducibility

We have shown that the two first orders of the loop expansion of $\Gamma(\varphi)$ contain only 1PI Feynman diagrams. We now suspect that the coefficient of \hbar^L is the generating functional of all L -loop 1PI diagrams. By realizing that vertex functions and connected correlation functions are affected by global combinatorial factors identical, respectively, to those of vertices in the action and connected tree diagrams, it is possible to give a general proof that $\Gamma(\varphi)$ is the generating functional of 1PI Feynman diagrams: one takes $\Gamma(\phi)$ as the action, one calculates at leading order in \hbar , and recovers $\mathcal{W}(J)$ as the generating functional of the corresponding connected tree diagrams. The result then follows from the considerations of the beginning of the section (equation (7.82)) (for a discussion, see Ref. [44]). However, we give here a more powerful and completely algebraic proof.

To prove that vertex functions are given in perturbation theory by a sum of 1PI Feynman diagrams, we directly use the definition and prove that by cutting one line in all possible ways in all diagrams contributing to $\Gamma(\varphi)$, the diagrams remain connected.

We consider the modified action,

$$\mathcal{S}_\varepsilon(\phi) = \mathcal{S}(\phi) + \frac{\varepsilon}{2} \left(\int d^d x \phi(x) \right)^2 = \frac{1}{2} \int d^d x d^d y \phi(x) \phi(y) [K(x - y) + \varepsilon] + \mathcal{V}_1(\phi), \quad (7.102)$$

where we have introduced a parameter ε in which we will expand at first order. The corresponding propagator $\Delta_\varepsilon(x - y)$ given by

$$\int \Delta_\varepsilon(x - z) [K(z - y) + \varepsilon] d^d z = \delta^{(d)}(x - y),$$

can be written as

$$\Delta_\varepsilon(x, y) = \Delta(x - y) - \varepsilon \eta(x) \eta(y) + O(\varepsilon^2), \quad (7.103)$$

with the definition

$$\eta(x) = \int \Delta(x - z) d^d z.$$

If we now expand in ε a Feynman diagram with the new propagator $\Delta_\varepsilon(x - y)$, we obtain, at first order, a sum of terms which consist of all possible ways in which a propagator $\Delta(x - y)$ has been replaced by the product $-\eta(x)\eta(y)$. Since, in this product, the dependence in x and y is factorized, this means topologically that, in the Feynman diagram, the corresponding line has been cut. A necessary and sufficient condition for a diagram to be 1PI is that all terms at order ε are connected.

Higher orders in ε make it possible to study irreducibility with respect to cutting two, three, or n lines.

The partition function $\mathcal{Z}_\varepsilon(J)$, at first order in ε , is given by

$$\mathcal{Z}_\varepsilon(J) = \int [d\phi] \left(1 - \frac{\varepsilon}{2} \int d^d x d^d y \phi(x) \phi(y) \right) \exp [-\mathcal{S}(\phi) + J \cdot \phi] + O(\varepsilon^2) \quad (7.104)$$

and, therefore,

$$\mathcal{Z}_\varepsilon(J) = \left[1 - \frac{\varepsilon}{2} \int d^d x d^d y \frac{\delta^2}{\delta J(x) \delta J(y)} + O(\varepsilon^2) \right] \mathcal{Z}(J). \quad (7.105)$$

One infers for the generating functional $\mathcal{W}_\varepsilon(J) = \ln \mathcal{Z}_\varepsilon(J)$,

$$\mathcal{W}_\varepsilon(J) = \mathcal{W}(J) - \frac{\varepsilon}{2} \left\{ \left[\int d^d x \frac{\delta \mathcal{W}}{\delta J(x)} \right]^2 + \int d^d x d^d y \frac{\delta^2 \mathcal{W}}{\delta J(y) \delta J(x)} \right\} + O(\varepsilon^2). \quad (7.106)$$

It contains at order ε a contribution of the form

$$\left[\int d^d x \frac{\delta \mathcal{W}}{\delta J(x)} \right]^2,$$

which is disconnected, as expected.

In the Legendre transformation, we use the relation (7.65) in the form

$$\frac{\partial \Gamma}{\partial \varepsilon} = -\frac{\partial \mathcal{W}}{\partial \varepsilon},$$

for $\varepsilon = 0$. Therefore,

$$\Gamma_\varepsilon(\varphi) = \Gamma(\varphi) + \frac{\varepsilon}{2} \left\{ \left[\int d^d x \varphi(x) \right]^2 + \int d^d x d^d y \left[\frac{\delta^2 \Gamma}{\delta \varphi(x) \delta \varphi(y)} \right]^{-1} \right\} + O(\varepsilon^2). \quad (7.107)$$

The first order in ε contains two terms, the term $\frac{1}{2}\varepsilon \left[\int d^d x \varphi(x) \right]^2$, which we have added explicitly to the action, and a second term which is the connected two-point function in the presence of an external field. The disconnected terms have been removed by the Legendre transformation. Since the variation of $\Gamma(\varphi)$ is connected, $\Gamma(\varphi)$ is indeed one line or, in the terminology of particle physics, one particle irreducible. In the chapters that follow, we use the term of vertex functions but recall occasionally that the functions are 1PI. Finally, note that in Section 7.7 we have shown that

$$W_{\text{amp}}^{(n)} = -\Gamma^{(n)} + \text{reducible terms} \quad (n > 2),$$

because the difference contains only lower correlation functions related by propagators. The diagrams contributing to $\Gamma^{(n)}$ thus differ from the 1PI amputated diagrams of $W^{(n)}$ only by a sign. In the same way, the diagrams contributing to $\Gamma^{(2)}$ beyond tree level are, up to a change of sign, the amputated diagrams of the mass operator.

7.11 Statistical and quantum interpretation of the vertex functional

The functional $\mathcal{Z}(J)$ can be considered as the classical partition function in an external field (or source) $J(x)$, and then $\mathcal{W}(J)$ is proportional to the free energy. We now provide a statistical interpretation to $\Gamma(\varphi)$.

7.11.1 Interpretation and variational principle

Statistical physics interpretation. We consider the free energy $\mathcal{W}(J)$ given by

$$e^{\mathcal{W}(J)} = \int [d\phi] e^{-S(\phi) + J \cdot \phi}.$$

We introduce a field $\varphi(x)$ and constrain the source $J(x)$ to satisfy

$$\varphi(x) = \frac{\delta \mathcal{W}(J)}{\delta J(x)} = \langle \phi(x) \rangle_J,$$

where, by $\langle \bullet \rangle_J$, we denote the expectation value with the weight $e^{-S(\phi) + J \cdot \phi}$.

The exponential function being convex, we infer the inequality,

$$\ln \langle e^{-J \cdot \phi} \rangle_J = \mathcal{W}(0) - \mathcal{W}(J) \geq \langle -J \cdot \phi \rangle_J = -J \cdot \varphi,$$

or introducing the Legendre transform $\Gamma(\varphi)$ of $\mathcal{W}(J)$,

$$\mathcal{W}(0) \geq -\Gamma(\varphi). \quad (7.108)$$

The inequality is the starting point of a variational principle. Moreover, $\mathcal{W}(0)$ is obtained in the limit $J = 0$ by taking the solution of

$$\frac{\delta \Gamma(\varphi)}{\delta \varphi(x)} = 0, \quad (7.109)$$

which minimizes $\Gamma(\varphi)$. In the framework of statistical physics, the Legendre transform $\Gamma(\varphi)$ is the *thermodynamic potential*, a quantity which plays a central role in the discussion of critical phenomena.

Finally, note that we have uncovered a general property of the Legendre transformation and, therefore, if we introduce a source J for any function of $\phi(x)$, a similar arguments will apply.

Quantum interpretation [45]. We now distinguish in the d -dimensional space \mathbb{R}^d Euclidean time direction and denote by t, x the time and space arguments, respectively. We explicitly assume time translation invariance: the quantum Hamiltonian $\hat{H}(\phi)$ corresponding to the action $\mathcal{S}(\phi)$ is time independent. We assume that t varies in a finite interval $[0, \beta]$ and impose periodic boundary conditions on the fields in the time direction. Moreover, we restrict the analysis to time-independent sources

$$J(t, x) = J(x),$$

and, therefore, also to functionals $\Gamma(\varphi)$ where the field φ is time independent, $\varphi(t, x) = \varphi(x)$. As a consequence, the functional $\mathcal{Z}(J)$ becomes, from the point of view of quantum statistical physics, the partition function at temperature $T = 1/\beta$:

$$\mathcal{Z} = \text{tr } e^{-\beta \hat{H}}.$$

The inequality (7.108) becomes

$$\ln \text{tr } e^{-\beta \hat{H}} \geq -\Gamma(\varphi), \quad (7.110)$$

where φ is the thermal expectation value corresponding to the quantum Hamiltonian

$$\hat{H}(\phi, J) = \hat{H}(\phi) - \int d^{d-1}x J(x) \hat{\phi}(x), \quad (7.111)$$

$\hat{\phi}(x)$ being the field operator.

In the large β limit, the partition function is dominated by the ground state energy E_0 of \hat{H} , and, thus,

$$\mathcal{W}(0) \underset{\beta \rightarrow \infty}{\sim} -\beta E_0. \quad (7.112)$$

In this limit, the inequality (7.110) becomes

$$E_0 \leq \frac{1}{\beta} \Gamma(\varphi), \quad (7.113)$$

where again E_0 is obtained by looking for the solution of equation (7.109) that minimizes $\Gamma(\varphi)$.

7.11.2 Vertex functional and free energy at fixed field time average

We now present a related, but slightly different, interpretation. We calculate the partition function with the same periodic boundary condition in time but restricted to fields satisfying the constraint

$$\frac{1}{\beta} \int_0^\beta dt \phi(t, x) = \varphi(x), \quad (7.114)$$

Time translation invariance then implies $\varphi(x) = \langle \phi(t, x) \rangle$. We denote by $-\beta\mathcal{G}(\varphi)$ the corresponding free energy

$$e^{-\beta\mathcal{G}(\varphi)} = \int [d\phi(t, x)] \exp[-\mathcal{S}(\phi)]. \quad (7.115)$$

We have written the free energy in the form $-\beta\mathcal{G}(\varphi)$, because we anticipate that, in the large β limit, the free energy is proportional to β .

Then, the free energy corresponding to the sum over all field configurations in the presence of a time-independent source $J(x)$ is given by

$$e^{\mathcal{W}(J)} = \int [d\varphi(x)] \exp \left[-\beta\mathcal{G}(\varphi) + \beta \int d^d x J(x)\varphi(x) \right]. \quad (7.116)$$

In the $\beta \rightarrow \infty$ limit, the field integral can be calculated by the steepest descent method. The saddle point equation is

$$J(x) = \frac{\delta\mathcal{G}}{\delta\varphi(x)}. \quad (7.117)$$

When the equation has several solutions, one has to take the solution that corresponds to a local maximum and yields the largest contribution to the free energy. Then,

$$\mathcal{W}(J) \sim -\beta\mathcal{G}(\varphi) + \beta \int d^d x J(x)\varphi(x). \quad (7.118)$$

After Legendre transformation, one finds

$$\beta\mathcal{G}(\varphi) = \Gamma(\varphi),$$

where again $\Gamma(\varphi)$ is the vertex functional restricted to time-independent fields. However, note that $\mathcal{G}(\varphi)$ has in general no reasons to be convex. One may find field configurations such that the operator

$$\frac{\delta^2\mathcal{G}(\varphi)}{\delta\varphi(x)\delta\varphi(y)}$$

is not positive. On the other hand, because $\Gamma(\varphi)$ is the result of a steepest descent calculation, it may coincide with $\beta\mathcal{G}(\varphi)$ only in regions of field space where the operator is positive. In general, in perturbation theory, one calculates a quantity which, restricted to time-independent fields, coincides with \mathcal{G} rather than Γ . This explains an apparent paradox: in the several phase region, one often pretends to be discussing the minima of $\Gamma(\varphi)$, that is, the minima of a quantity which has convexity properties and can have only one minimum. In fact, one discusses the properties of \mathcal{G} .

A7 Additional results and methods

A7.1 Generating functional at two loops

We illustrate the discussion of Section 7.9.3 by an explicit two-loop calculation. Expanding the interaction terms $S^{(3)}$ and $S^{(4)}$ (equation (7.100)), integrating term by term over χ (equation (7.98)) and choosing a simple normalization of the field integral, we expand $\mathcal{W}(J) = \hbar \ln \mathcal{Z}(J)$ as

$$\mathcal{W}(J) = -\mathcal{S}(\phi_c) + J \cdot \phi_c - \frac{1}{2}\hbar \text{tr} \ln S^{(2)}(x_1, x_2; \phi_c) + \hbar^2 \mathcal{W}_2(J) + O(\hbar^3), \quad (A7.1)$$

where the two-loop contribution is given by

$$\begin{aligned} \mathcal{W}_2(J) &= -\frac{1}{8} \int d^d x_1 \cdots d^d x_4 S^{(4)}(x_1, x_2, x_3, x_4) \Delta(x_1, x_2; \phi_c) \Delta(x_3, x_4; \phi_c) \\ &\quad + \int d^d x_1 \cdots d^d y_3 S^{(3)}(x_1, x_2, x_3; \phi_c) S^{(3)}(y_1, y_2, y_3; \phi_c) \left[\frac{1}{8} \Delta(x_1, x_2; \phi_c) \right. \\ &\quad \times \Delta(y_1, y_2; \phi_c) \Delta(x_3, y_3; \phi_c) + \frac{1}{12} \Delta(x_1, y_1; \phi_c) \Delta(x_2, y_2; \phi_c) \Delta(x_3, y_3; \phi_c) \left. \right]. \end{aligned} \quad (A7.2)$$

We now perform the Legendre transformation (7.62). We need $\varphi(x)$ only up to order \hbar ,

$$\varphi(x) = \phi_c(J; x) - \frac{\hbar}{2} \frac{\delta}{\delta J(x)} \text{tr} \ln S^{(2)}(x_1, x_2; \phi_c) + O(\hbar^2), \quad (A7.3)$$

because expression $\Gamma(\varphi) - J \cdot \varphi$ is stationary in φ . Using $\delta \text{tr} \ln X = \text{tr} \delta X X^{-1}$, valid for any matrix or operator X , and applying the chain rule, we rewrite the order \hbar correction as

$$\frac{\delta}{\delta J(x)} \text{tr} \ln S^{(2)}(x_1, x_2; \phi_c) = \int d^d y d^d z_1 d^d z_2 \frac{\delta \phi_c(y)}{\delta J(x)} S^{(3)}(y, z_1, z_2; \phi_c) \Delta(z_2, z_1, \phi_c).$$

Using equations (7.97), (A7.3), and the definition (7.101), we express ϕ_c in terms of φ . At order \hbar ,

$$\varphi(x) = \phi_c(J, x) - \frac{\hbar}{2} \int d^d y d^d y_1 d^d y_2 S^{(3)}(y, y_1, y_2; \varphi) \Delta(x, y; \varphi) \Delta(y_1, y_2; \varphi). \quad (A7.4)$$

We still need $J(x)$ at order \hbar . Using equations (7.97) and (A7.4), we find that

$$J(x) = \frac{\delta \mathcal{S}(\varphi)}{\delta \phi(x)} + \frac{\hbar}{2} \int d^d y_1 d^d y_2 S^{(3)}(x, y_1, y_2; \varphi) \Delta(y_1, y_2; \varphi) + O(\hbar^2). \quad (A7.5)$$

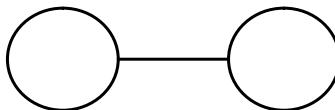


Fig. 7.10 The reducible part at two-loop order

Equation (7.62) then yields $\Gamma(\varphi)$ at two-loop order. As expected the reducible part in expression (A7.2) (Fig. 7.10) cancels, and we obtain

$$\Gamma(\varphi) = \mathcal{S}(\varphi) + \frac{1}{2}\hbar \text{tr} \ln S^{(2)}(x_1, x_2; \varphi) + \hbar^2 \Gamma_2(\varphi) + O(\hbar^3), \quad (\text{A7.6})$$

where the two-loop contribution is

$$\begin{aligned} \Gamma_2(\varphi) &= \frac{1}{8} \int d^d x_1 d^d x_2 d^d x_3 d^d x_4 \Delta(x_1 - x_2; \varphi) S^{(4)}(x_1, x_2, x_3, x_4; \varphi) \Delta(x_3 - x_4; \varphi) \\ &\quad - \frac{1}{12} \int d^d x_1 d^d x_2 d^d x_3 d^d y_1 d^d y_2 d^d y_3 S^{(3)}(x_1, x_2, x_3; \varphi) \Delta(x_1 - y_1; \varphi) \Delta(x_2 - y_2; \varphi) \\ &\quad \times \Delta(x_3 - y_3; \varphi) S^{(3)}(y_1, y_2, y_3; \varphi). \end{aligned} \quad (\text{A7.7})$$

Fig. 7.11 gives a diagrammatic representation of the two-loop contributions.



Fig. 7.11 The two-loop contributions to $\Gamma(\varphi)$

A7.2 The background field method

In most of our work, we regard correlation functions of the field as fundamental physical objects. However, in some cases all or some local functionals of the field, which have a non-trivial linear part in ϕ , are equivalent. Important examples are:

- (i) Some models are defined on Riemannian manifolds and fields correspond to a particular choice of coordinates on the manifold. For some issues, only quantities intrinsic to the manifold are physical.
- (ii) In gauge theories, only gauge independent quantities are physical. Change of gauges correspond to field redefinitions.
- (iii) In particle physics, normalized S -matrix elements are invariant under a change of field variables (see Section 6.5.4).

Here, we introduce a method, the background field method [46], which has among its main merits, the power to make a more direct calculation possible of quantities that are rather insensitive to a change of field variables.

The generating functional of vertex functions $\Gamma(\varphi)$, corresponding to an action $\mathcal{S}(\phi)$, can be derived from

$$e^{-\Gamma(\varphi)+J\cdot\varphi} = \int [d\phi] e^{-\mathcal{S}(\phi)+J\cdot\phi}, \quad (\text{A7.8})$$

or, using

$$J(x) = \frac{\delta\Gamma}{\delta\varphi(x)},$$

from

$$e^{-\Gamma(\varphi)} = \int [d\phi] \exp \left[-\mathcal{S}(\phi) + \int d^d x (\phi(x) - \varphi(x)) \frac{\delta\Gamma}{\delta\varphi(x)} \right], \quad (\text{A7.9})$$

or, equivalently, translating $\phi(x)$, $\phi \mapsto \varphi + \chi$:

$$e^{-\Gamma(\varphi)} = \int [d\chi] \exp \left[-\mathcal{S}(\chi + \varphi) + \int d^d x \chi(x) \frac{\delta\Gamma}{\delta\varphi(x)} \right]. \quad (\text{A7.10})$$

We now assume that the equation

$$\frac{\delta \Gamma}{\delta \varphi(x)} = 0 \quad (A7.11)$$

has a non-trivial solution $\varphi_c(x)$, which, at leading order in perturbation theory, is a solution $\varphi_c^{(0)}(x)$ of the classical equation of motion:

$$\frac{\delta \mathcal{S}}{\delta \varphi(x)} \left[\varphi_c^{(0)} \right] = 0. \quad (A7.12)$$

Then equation (A7.10) becomes (see also the discussion at the end of Section 7.9.3)

$$e^{-\Gamma(\varphi_c)} = \int [d\chi] e^{-\mathcal{S}(\chi + \varphi_c)}. \quad (A7.13)$$

The quantity $\Gamma(\varphi_c)$ is clearly independent of the representation of the field ϕ , and contains, therefore, only physical information in the sense defined at the beginning of the section. We introduce Γ_r , the renormalized generating functional of vertex functions (see Chapter 9), given by

$$\Gamma_r(\varphi_c) = -\ln \int [d\chi] \exp [-\mathcal{S}_0(\chi + \varphi_c) + \text{counter-terms}],$$

in which $\mathcal{S}_0(\phi)$ is the tree order action. The solution φ_c of

$$\frac{\delta \Gamma_r}{\delta \varphi_c(x)} = 0$$

is expanded around the solution of $\varphi_c^{(0)}(x)$ of

$$\frac{\delta \mathcal{S}_0}{\delta \varphi_c(x)} = 0.$$

It can be inferred from the discussion of Section 6.5.3 that, in real time, the background field method yields the S -matrix. We shall provide other examples of calculations involving the background field method in the coming chapters.

A7.3 Connected Feynman diagrams: Cluster properties

We briefly describe the cluster properties of connected Feynman diagrams contributing to Euclidean correlation functions, in a massive field theory. We restrict ourselves to a theory with one massive scalar field, but the generalization is straightforward. We then discuss the influence of threshold effects for real-time diagrams at large time separation.

A7.3.1 Decay of connected Feynman diagrams in Euclidean space

The propagator can be written (equation (7.14)) as

$$\Delta(x) = \frac{1}{(2\pi)^d} \int d^d p \frac{e^{ip \cdot x}}{p^2 + m^2}. \quad (A7.14)$$

We assume in the following that m is the physical mass, and this requires a mass renormalization of Feynman diagrams (see Chapter 9).

To determine the behaviour of Δ for $|x|$ large, we rewrite the integral as

$$\Delta(x) = \frac{1}{(2\pi)^d} \int d^d p \int_0^{+\infty} dt e^{ip \cdot x} e^{-t(p^2 + m^2)}. \quad (A7.15)$$

We then perform the Gaussian integration over the momentum p :

$$\Delta(x) = \frac{\pi^{d/2}}{(2\pi)^d} \int \frac{dt}{t^{d/2}} \exp \left(-tm^2 - \frac{1}{4t}x^2 \right). \quad (A7.16)$$

The behaviour of Δ for large separation is given by the method of steepest descent. The saddle point is

$$t = \frac{|x|}{2m}. \quad (A7.17)$$

Finally, the Gaussian integration over fluctuations around the saddle point yields

$$\Delta(x) \sim \frac{1}{2m} \left(\frac{m}{2\pi|x|} \right)^{(d-1)/2} e^{-m|x|}. \quad (A7.18)$$

The large distance decay of the propagator is governed by the inverse of the particle mass m (for $\hbar = c = 1$), which, in the theory of phase transitions, is also the *correlation length*. This asymptotic estimate remains valid even when the propagator is replaced by a momentum-regularized propagator, as one can verify more easily using the regularization scheme of Section 8.4.3.

Using this asymptotic estimate, one can derive the following property: if in a connected diagram we separate two sets of points by a distance l , then at large l the diagram decreases as $\exp(-nml)$. In this expression, n is the smallest number of lines it is necessary to cut in order to disconnect the diagram, the two sets of points being attached to different connected components (see Fig. 7.12).

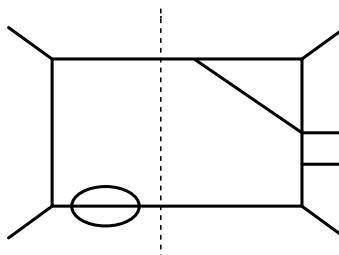


Fig. 7.12 Example with $n = 2$

By contrast, in a massless theory ($m = 0$) the decay is only algebraic when the propagator exists (this implies, in perturbation theory, $d > 2$).

A7.3.2 Threshold effects

The precise large distance behaviour of diagrams is related to the strength of the leading singularity in momentum space (or Fourier representation). For example, if, in momentum space, a contribution to the two-point function has the algebraic singularity,

$$\tilde{K}^{(2)}(p) \propto (m^2 + p^2)^{-\alpha},$$

a generalization of the preceding calculation yields the large distance behaviour

$$K^{(2)}(r) \propto r^{\alpha - d/2 - 1/2} e^{-mr}.$$

If we now consider an 1PI diagram with n internal lines, it yields, for r large, the contribution,

$$K^{(2)}(r) \propto r^{-n(d-1)/2} e^{-nmr}.$$

This, in turn, corresponds to a singularity

$$\tilde{K}^{(2)}(p) \propto (m^2 + p^2)^{-\alpha}, \quad \alpha = \frac{1}{2}[(n-1)d - n - 1].$$

In particular, for $d > 1$, the singularity softens when n increases. The two-particle threshold yields the strongest singularity $(p^2 + 4m^2)^{(d-3)/2}$. The nature of the singularity is important for the large-time behaviour of correlation functions in real time: the leading large-time behaviour is then related to the leading singularity in the energy variable (a property of the Fourier transformation). Therefore, if we consider the two-point function, its large-time behaviour is given by the one-particle pole, then the next to leading term is related to the two-particle threshold and so on.

8 Ultraviolet divergences: Effective field theory (EFT)

In Chapter 7, we have explained how, in quantum field theory (QFT), physical quantities can be calculated as power series in the various interactions. We consider now specifically *local* relativistic QFTs: the action that appears in the field integral is the space-time integral of a classical Lagrangian density, function of fields, and their derivatives, taken at the same point. We show that, as a consequence of locality (an essential property), infinities appear in *perturbative calculations*, due to short distance singularities, or after Fourier transformation, to integrals diverging at large momenta: one speaks of ultraviolet (UV) divergences. These divergences are peculiar to local QFT: in contrast to classical mechanics or non-relativistic quantum mechanics with a finite number of particles, a straightforward construction of a relativistic quantum theory of point-like objects with contact interactions is impossible.

Therefore, a local QFT, in a straightforward formulation, is an *incomplete theory*. It must, eventually (perhaps at the Planck's scale?), be embedded in some non-local theory, which renders the full theory finite, but where the non-local effects affect only short distance properties (an operation sometimes called UV completion). The impossibility to define a QFT without an explicit reference to an external short scale is an indication of a *non-decoupling* between short and long-distance physics.

From the more technical viewpoint of field integration, the Gaussian measure corresponding to a free field theory does not constrain the fields to be even continuous and, therefore, expectation values of products of fields taken at the same point as they appear in the perturbative expansion are not defined.

In Section 8.2, we investigate the form of divergences in a general local QFT, to all orders in perturbation theory. The analysis is based on *power counting* arguments. It leads to a classification of interactions as being super-renormalizable, renormalizable, and non-renormalizable. The relevance of this classification will appear in Chapter 9, where we introduce renormalization and renormalization group (RG).

To characterize the nature of divergences, one modifies the QFT at large momentum (the cut-off scale), or equivalently at short distance, in such a way that a finite perturbative expansion (based on Feynman diagrams) can be defined. The modifications must be such that when the momentum cut-off is sent to infinity, the original action is formally recovered. This procedure is called *regularization* (the dimensional regularization is of different nature, and has no direct physics interpretation, see Chapter 10). The regularization that renders the perturbative expansion finite is a substitute for the necessary short-distance structure, which, either is too complicated (as in statistical physics), or is unknown, as in particle physics. It makes it possible to isolate well-defined divergent parts of diagrams and deal with them with renormalization, as we explain in Chapter 9.

The physical interpretation of the divergences of QFT leads to the concept of *EFT* (see also Section 9.11): QFT is an EFT, local approximation, valid at large distance or low energy, of a more fundamental non-local finite theory.

Note that *one also calls EFT a QFT that is a low energy approximation of an embedding QFT with heavier masses*, which act as cut-offs [47].

Remark. We discuss deeply quantum and relativistic physics (or statistical physics) , and thus set everywhere $\hbar = c = 1$.

8.1 Gaussian expectation values and divergences: The scalar field

The n -point correlation function for a scalar field ϕ is given by a field integral of the form,

$$\langle \phi(x_1) \cdots \phi(x_n) \rangle = \mathcal{Z}^{-1} \int [d\phi] \phi(x_1) \cdots \phi(x_n) e^{-S(\phi)}, \quad \text{with } \mathcal{Z} = \int [d\phi] e^{-S(\phi)}, \quad (8.1)$$

where \mathcal{Z} is the partition function, or vacuum amplitude.

One could have expected that, as in non-relativistic quantum mechanics, $S(\phi)$ can be chosen as a *local* action, space integral of Euclidean *classical Lagrangian density*, function of $\phi(x)$ and its derivatives. Examples are translation invariant QFTs, in d space dimensions (including Euclidean time), with an $O(d)$ invariant action, of the form,

$$S(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + V_1(\phi(x)) \right], \quad (8.2)$$

$(\nabla \equiv (\partial/\partial x_1, \dots, \partial/\partial x_d))$ where m is the ϕ mass at leading order, and $V_1(\phi)$ a polynomial.

Locality and unavoidable UV divergences. As we have explained in Section 7.2, the field integral can be calculated, *perturbatively*, by keeping a quadratic term, here a free field action of the form

$$S_0(\phi) = \frac{1}{2} \int d^d x \left[(\nabla \phi(x))^2 + m^2 \phi^2(x) \right], \quad (8.3)$$

in the exponential and expanding the partition function in a power series of $V_1(\phi)$, reducing the calculation to an infinite sum of Gaussian integrals (see Section 7.2.2).

The Gaussian expectation value of a monomial of degree n contributing to $V_1(\phi)$,

$$\langle \phi^n(x) \rangle_0 \propto \int [d\phi] \phi^n(x) e^{-S_0(\phi)},$$

can be evaluated, using Wick's theorem (7.18), in terms of the Gaussian two-point correlation function corresponding to the action (8.3),

$$\Delta(x) = \frac{1}{(2\pi)^d} \int \frac{d^d p e^{ipx}}{p^2 + m^2}.$$

It thus involves contributions of the form

$$\langle \phi^2(x) \rangle_0 = \Delta(x=0) = \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 + m^2}, \quad (8.4)$$

which are infinite for $d \geq 2$. One calls such divergences *UV divergences*, because they follow from the large momentum behaviour of perturbative contributions.

This divergence is related to the property that the typical fields that contribute to the Gaussian field integral are not continuous. Indeed, in d dimensions,

$$\frac{1}{2} \left\langle (\phi(x) - \phi(y))^2 \right\rangle_0 = \frac{1}{(2\pi)^d} \int \frac{d^d p (1 - e^{ip(x-y)})}{p^2 + m^2} \underset{|x-y| \rightarrow 0}{\propto} |x-y|^{2-d},$$

and thus, the fields are continuous only for $d < 2$ (quantum mechanics). Expectation values of product of fields at coinciding points (local monomials) are always divergent.

Note that, in the case of relativistic derivative couplings, divergences of the form,

$$\left\langle (\nabla \phi(x))^2 \right\rangle_0 = \frac{1}{(2\pi)^d} \int \frac{d^d p p^2}{p^2 + m^2},$$

may appear. In order to be defined, these quantities require the fields contributing to the field integral to be differentiable.

8.2 Divergences of Feynman diagrams: Power counting

We have shown that a straightforward perturbation expansion is plagued by UV divergences, revealing that *any local, interacting QFT is an incomplete theory*, because *short-and large-distance physics do not decouple*. However, before discussing the methods to deal with this essential problem (regularization, renormalization, and RG), we first characterize the general structure and nature of the UV divergences of all Feynman diagrams generated by polynomial local interactions [25].

8.2.1 UV dimension of fields and interaction vertices

Propagator: Large momentum behaviour. Quite generally, we assume that the propagator (after continuation to Euclidean time) of the field ϕ , in the Fourier representation,

$$\langle \tilde{\phi}(p)\tilde{\phi}(-p) \rangle_0 = \tilde{\Delta}(p),$$

behaves for large momenta as,

$$\left| \tilde{\Delta}(\lambda p) \right|_{\lambda \rightarrow +\infty} \sim C \lambda^{-\sigma}, \quad \text{with } \sigma, C > 0. \quad (8.5)$$

For the action (8.2) $\sigma = 2$, but, for later purpose, we need to generalize this behaviour.

Then, from the viewpoint of the functional Gaussian measure, the space of fields contributing to the field integral can be inferred from the short-distance behaviour of the Gaussian two-point function. One finds

$$\frac{1}{2} \left\langle (\phi(x) - \phi(y))^2 \right\rangle_0 = \frac{1}{(2\pi)^d} \int d^d p \left(1 - e^{ip(x-y)} \right) \tilde{\Delta}(p) \Big|_{|x-y| \rightarrow 0} |x-y|^{\sigma-d},$$

for $\sigma > d$. Typical fields contributing to the field integral are continuous only for $\sigma > d$ and are k times continuously differentiable if

$$\sigma > d + 2k. \quad (8.6)$$

In particular, as we have shown, the expectation value of any derivative-free interaction $V_I(\phi)$ is finite only if the fields contributing to the field integral are continuous.

UV field dimension. We now define the dimensions $[x]$ of space and, correspondingly, $[\nabla]$ of derivative or momentum, and the canonical, or UV dimension $[\phi]$ of the field ϕ , appropriated to the large momentum behaviour (equation (8.5)) by

$$[x] = -1, \quad [\nabla] = 1, \quad [\phi] = \frac{1}{2}(d - \sigma). \quad (8.7)$$

Then, the leading order contribution from $V_I(\phi)$ is finite only if $[\phi]$ is negative.

Interaction vertex UV dimensions. Quite generally, we assume that the perturbation V_I is a linear combination of local monomials of degree n in the field ϕ of the form,

$$\mathcal{V}_{k,n}^\alpha(\phi) = \int d^d x V_{k,n}^\alpha(\phi, x), \quad (8.8)$$

where $V_{k,n}^\alpha$ involves k differentiations (the index α reflects the property that, at k, n fixed, one can find sometimes several monomials), for example,

$$\begin{aligned} \phi^3(x) (n=3, k=0), \quad &\phi^4(x) \ (n=4, k=0), \quad \phi^6(x) \ (n=6, k=0), \\ \phi^2(x) \nabla^2 \phi^2(x) \ (n=4, k=2) \dots . \end{aligned}$$

We call a monomial of the form $\mathcal{V}_{k,n}^\alpha(\phi)$ a vertex, because, for $n \geq 3$, it is represented by a vertex in Feynman diagrams.

It is natural then to define the dimension $[\mathcal{V}]$ of a vertex $\mathcal{V}(\phi)$ by

$$[\mathcal{V}] = -d + k + n[\phi]. \quad (8.9)$$

In terms of the Fourier components $\tilde{\phi}(p)$ of the fields $\phi(x)$, and taking into account translation invariance, we write the vertex $\mathcal{V}(\phi)$ (in symbolic notation) as

$$\mathcal{V}(\phi) \propto \int \prod_{i=1}^n d^d p_i \delta^{(d)}(p_1 + p_2 + \cdots + p_n) p_{i_1} p_{i_2} \cdots p_{i_k} \tilde{\phi}(p_1) \cdots \tilde{\phi}(p_n),$$

where $[\tilde{\phi}] = [\phi] - d$. For simplicity, we have written all expressions for only one field ϕ , but the same divergence analysis applies to several scalar fields that have propagators with the same large momentum behaviour.

8.2.2 Vertex functions: Power counting, superficial degree of divergence,

We consider only 1PI diagrams (see Section 7.10) contributing to vertex functions. In the Fourier representation, each vertex contains a δ -function of momentum conservation (translation invariance). The number of independent integration momenta in a Feynman diagram, taking into account momentum conservation at vertices, thus equals the number of loops. This follows directly from one of the definitions of the number of loops L in a diagram given in Section 7.9. Finally, a vertex multiplies the numerator of a Feynman diagram by the product of k momenta.

Therefore, if all integration momenta in a diagram γ are scaled by a factor λ , for $\lambda \rightarrow \infty$, the diagram is scaled by a factor $\lambda^{\delta(\gamma)}$ with

$$\delta(\gamma) = dL - I\sigma + \sum_a v_a k_a, \quad (8.10)$$

in which v_a is the number of vertices of type a with k_a derivatives, and I the number of internal lines corresponding to propagators Δ joining the different vertices.

The calculation of $\delta(\gamma)$ is called *power counting*, and $\delta(\gamma)$ is called the *superficial degree of divergence* of the diagram γ . For $\delta(\gamma) \geq 0$, the Feynman diagram diverges. If $\delta(\gamma)$ is negative, the diagram is superficially convergent but, beyond one-loop, it may still have divergences coming from subdiagrams.

Example. In the example of the ϕ^3 field theory in $d = 6$ dimensions, since $\sigma = 2$ and $k = 0$, expression (8.10) yields $\delta(\gamma) = 6L - 2I$.

At one-loop $L = 1$, $\delta(\gamma) = 6 - 2I$ and, for $I = 1, 2, 3$, the diagrams are divergent in agreement with equation (8.32). For $I > 3$, all one-loop diagrams are convergent.

Fig. 8.1 exhibits a diagram in the same theory that is superficially convergent ($L = 2$, $I = 7$), but which, inside the dotted box, contains a divergent subdiagram.

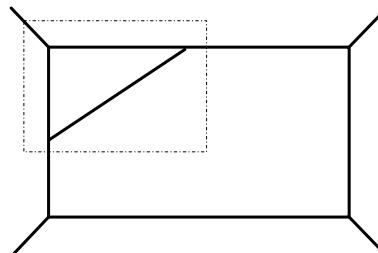


Fig. 8.1 Divergent subdiagram inside the dotted box

Other expressions. By using several topological relations on graphs, it is possible to express $\delta(\gamma)$ in different forms.

Combining equation (8.10) with the relation (7.96) written in the form,

$$L = I - \sum_a v_a + 1, \quad (8.11)$$

one can eliminate L and finds

$$\delta(\gamma) = d + 2I[\phi] + \sum_a (k_a - d)v_a. \quad (8.12)$$

We now consider a diagram γ contributing to a vertex function with E (for external line) fields ϕ . We denote by n^a the number of fields ϕ belonging to a vertex a , inside a diagram γ . They satisfy the relation

$$E + 2I = \sum_a n^a v_a. \quad (8.13)$$

The relation has a simple interpretation: each internal line connects two vertices while an external line is attached only to one vertex. Fig. 8.2 gives an example.

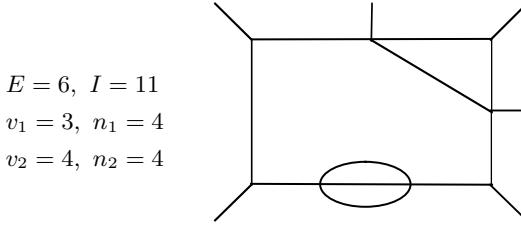


Fig. 8.2 Diagram illustrating equation (8.13)

Combining equation (8.13) with the relation (8.12), one can eliminate I in $\delta(\gamma)$, and obtain

$$\delta(\gamma) = d - E[\phi] + \sum_a v_a [\mathcal{V}_a], \quad (8.14)$$

where $[\mathcal{V}_a]$ is the dimension of the vertex a (equation (8.9)):

$$[\mathcal{V}_a] = -d + k_a + n_a [\phi]$$

and $[\phi]$ the dimension of ϕ (equation (8.7)). Equation (8.14) leads directly to a classification of scalar field theories, according to the degree of divergence of Feynman diagrams.

8.3 Classification of interactions in scalar quantum field theories

We classify vertices from the point of view of power counting. We use the word *renormalizable*, a concept that will be justified in Chapter 9. For a scalar field theory, $\sigma = 2$ and, thus, $[\phi] = \frac{1}{2}(d - 2)$. Then, the dimension of a vertex is (equation (8.9)),

$$[\mathcal{V}] = -d + k + \frac{1}{2}n(d - 2), \quad (8.15)$$

where k is even in scalar QFTs.

8.3.1 Classification of vertices

(i) *Super-renormalizable vertices (or interactions).* If the dimension (8.9) of a vertex is negative, that is,

$$-d + k + n(d - 2)/2 < 0, \quad (8.16)$$

the vertex is called *super-renormalizable* or, in the RG terminology, *relevant* (see Section 15.2). Adding such a vertex to a Feynman diagram decreases the superficial degree of divergence (8.14). The condition (8.16) then implies:

- for $d = 2$, $k = 0$, and the interaction ϕ^n is super-renormalizable for all n ;
- for $d = 3$, the condition becomes $k + n/2 - 3 < 0$, and then $k = 0$ and $n < 6$;
- for $d = 4$, the condition is $k + n - 4 < 0$, and then again $k = 0$ and $n < 4$;
- for $d = 5$, $k = 0$ and $3n/2 < 5$, and only ϕ^3 is super-renormalizable. For $d \geq 6$, no interaction is super-renormalizable.

(ii) *Renormalizable vertices.* This is the RG marginal case where the dimension of the vertex vanishes. Adding such a vertex to a Feynman diagram does not change the superficial degree of divergence. The condition (8.16) then implies:

- for $d = 2$, $k = 2$ and all interactions with two derivatives are renormalizable;
- for $d = 3$, only the ϕ^6 interaction is renormalizable;
- for $d = 4$, the ϕ^4 interaction is renormalizable;
- for $d > 4$, only the ϕ^3 interaction is renormalizable in $d = 6$ dimensions.

(iii) *Non-renormalizable (or RG irrelevant) vertices.* The remaining vertices, when added to a diagram, increase the superficial degree of divergence.

8.3.2 Classification of field theories

We now classify field theories based on the form (8.14) of the superficial degree of divergence. The classification follows directly from the classification of vertices.

Non-renormalizable theories. If at least one vertex \mathcal{V} has a positive dimension, $[\mathcal{V}] > 0$, then the degree of divergence of diagrams contributing to any vertex function can be rendered arbitrarily large by increasing the number v of vertices of this type. A field theory with such a vertex is not renormalizable, as will be discussed in Chapter 9.

Super-renormalizable theories. When all vertices have strictly negative dimensions, only a finite number of Feynman diagrams are superficially divergent. The corresponding field theory is called super-renormalizable.

Example. In the ϕ^4 field theory, in dimension $d = 3$,

$$\delta(\gamma) = 3 - \frac{1}{2}E - v.$$

The superficially divergent diagrams are listed in Fig. 8.3.

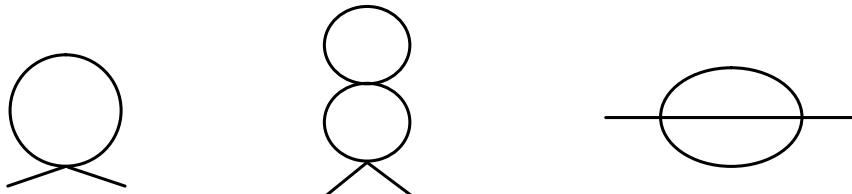


Fig. 8.3 Superficially divergent diagrams in the $\phi^4_{d=3}$ QFT

Renormalizable theories. Renormalizable QFTs are characterized by the property that at least one vertex has dimension 0, and no vertex has a positive dimension. Then, an infinite number of diagrams has a positive superficial degree of divergence, but the maximal degree of divergence at E fixed is bounded and is independent of the number of insertions of the vertices of dimension 0. In addition, if all dimensions of fields $[\phi]$ are strictly positive, only a finite number of vertex functions are superficially divergent.

If at least one field has dimension 0 (this happens with a scalar field in dimension 2), the situation is more complicated: the degree of divergence is bounded, but an infinite number of vertex functions are superficially divergent. Generically, this leads to field theories depending on an infinite number of parameters, except when some symmetry relates them (see Chapters 19 and 29).

Finally, no stable field theory with polynomial interaction exists above dimension 4. For particle physics, dimension 4 is also the physical dimension (at least for macroscopic dimensions). It is not known whether this is a mere coincidence or has a deeper meaning.

8.4 Momentum regularization

The perturbative expansion in all local, relativistic QFTs is plagued by UV divergences. These divergences signal a non-decoupling of short and large distance scales, a situation totally unusual in physics. Therefore, a QFT is an incomplete theory, which has eventually to be embedded within a finite non-local microscopic theory. However, in particle physics, such a theory is unknown. In macroscopic phase transitions, the microscopic structure is often known, but may be too complicated, and contain irrelevant features.

The momentum *regularization method* consists in *artificially* modifying the short distance, or large Euclidean momentum structure of the QFT to render the perturbative expansion finite. The modification is largely arbitrary and, in particle physics, not physical, since it violates unitarity. Momentum regularization works in the continuum, compared to lattice methods (see Section 8.7), and at fixed dimension, unlike dimensional regularization (which, moreover, has no physics interpretation, see Chapter 10). From the viewpoint of field integration, the regularized action selects a space of more regular fields (at least continuous to render theories with derivative-free interactions finite). Such modified theories make physical sense only if their large distance (or low energy-momentum) properties are, at least to a great extent, insensitive to the detailed form of the modification (a property also called *universality*). Renormalization theory and RG are the relevant tools to investigate this problem.

8.4.1 Effective field theory: Regularization

In order to limit the number of parameters, we consider only QFTs symmetric in the reflection $\phi \rightarrow -\phi$ (but we exhibit a different example in Section 8.5).

To give a meaning to perturbation theory, we modify the action, which we write now, quite generally, as

$$\mathcal{S}(\phi) = \mathcal{S}_G(\phi) + \mathcal{V}_I(\phi), \quad (8.17)$$

with

$$\mathcal{S}_G(\phi) = \frac{1}{2} \int d^d x d^d y \phi(x) K(x-y) \phi(y), \quad \mathcal{V}_I(\phi) = \int d^d x [\frac{1}{2} v_0 \phi^2(x) + V_I(\phi, x)], \quad (8.18)$$

where v_0 is a parameter, and K can be a positive local differential operator, polynomial in ∇^2 or even, if required, a non-local kernel (see Section A8.3), which cuts the momentum integration in Feynman diagrams, and $V_I(\phi, x)$ a polynomial (for simplicity) in ϕ and its derivatives at point x , with degree $n \geq 4$.

For generic values of the parameter v_0 , the physical mass (or the inverse correlation length equation defined by (A7.18)) is of the order of the momentum cut-off and, thus, non-physical. Therefore, the parameter v_0 has to be tuned in such a way that the physical mass becomes much smaller than the cut-off. In macroscopic phase transitions, this amounts to adjusting the temperature close to the critical temperature (see Chapter 14). This problem is specific to scalar QFTs.

The kernel K is an artificial substitute for the real short-distance structure that renders the theory finite. We assume that the kernel K has the Fourier representation,

$$K(x) = \frac{1}{(2\pi)^d} \int d^d p e^{-ipx} \tilde{K}(p), \quad \tilde{K}(p) = m^2 + p^2 + O(p^4). \quad (8.19)$$

Moreover, we assume that $\tilde{K}(p)$ is an entire function of p^2 , positive for $p^2 > 0$, which has zeros only on the p^2 negative semi-axis $p^2 \leq 0$, and which increases fast enough for $|p| \rightarrow \infty$. Indeed, singularities in momentum variables generate, after Fourier transformation, contributions to the large-distance behaviour of the propagator and regularization should modify the QFT only at short distance. Examples are described in Sections A8.2 (based on Schwinger's proper-time representation [53], see Section A8.1) and A8.3.

We define the perturbative expansion by keeping S_G in the exponential and expanding the remainder, reducing the calculation of the field integral (8.1) to an infinite sum of Gaussian expectation values. As a consequence of Wick's theorem, all perturbative contributions can be expressed in terms of the Gaussian two-point function.

Power counting. The Gaussian two-point function, or propagator, in the Fourier representation, is now

$$\tilde{\Delta}(p) = 1/\tilde{K}(p). \quad (8.20)$$

If we assume that $\tilde{K}(p) \propto |p|^\sigma$, we can use the power counting arguments of Section 8.2 and investigate under which condition the perturbative expansion is finite.

In the specific example of a QFT with derivative-free interactions, finiteness of $\langle \phi^2(0) \rangle$ (see also the integral (8.4)) then yields the necessary condition $\sigma > d$ and, therefore, $[\phi] < 0$. The condition is also sufficient since equation (8.12) implies,

$$\delta(\gamma) = d - d \sum_a v_a + 2I[\phi] < 0,$$

because $[\phi] < 0$ and $v_a \geq 1$.

In the case of interactions with derivatives, using equation (8.10) and the topological relation (8.11), it is convenient to rewrite the superficial (or global) degree of divergence of a Feynman diagram γ (equation (8.14)) as

$$\delta(\gamma) = (d - \sigma)L + \sum_a v_a (k_a - \sigma) + \sigma,$$

where L is the number of loops and v_a the number of vertices of type a and k_a the number of derivatives at the vertex (Section 8.2).

To render all diagrams finite, it is thus necessary to choose $\sigma > d + \sup_a k_a$, since both L and v_a can increase indefinitely. Moreover, since $L \geq 1$, and one at least of the v_a is positive, one can then always satisfy $\delta(\gamma) < 0$ for all diagrams. One recovers the regularity condition (8.6) for the fields contributing to the field integral.

UV divergences from quantization. Momentum cut-offs deal with divergences caused by the infinite number of degrees of freedom of fields, a property that leads also to an RG. With some care they preserve, beyond space–time symmetries, all linear symmetries of the initial action. However, they do not remove, in general, divergences related to quantization problems and order of quantum operators in local products (*e.g.* see Chapter 19).

8.4.2 Terms quadratic in the fields with higher derivatives

We now exhibit a special class of regularizations that render perturbation theory finite when, in the action, all non-renormalizable interactions are omitted (see Section 8.3.1). It is characterized by the polynomial form of $\tilde{K}(p)$ in the expression (8.19).

We consider again the action (8.17),

$$\mathcal{S}(\phi) = \mathcal{S}_G(\phi) + \mathcal{V}_I(\phi). \quad (8.21)$$

To improve the convergence of Feynman diagrams at large momentum, we choose

$$\mathcal{S}_G(\phi) = \frac{1}{2} \int d^d x \phi(x) (-\nabla^2 + m^2) \prod_{i=1}^s (1 - \nabla^2/M_i^2) \phi(x), \quad (8.22)$$

with $M_i^2 \gg m^2$. In the limit where all M_i 's become infinite, the unregularized propagator is recovered. This is the spirit of Pauli–Villars's regularization scheme [48].

The choice corresponds to the kernel (equation (8.55)),

$$\tilde{K}(p) = (p^2 + m^2) \prod_{i=1}^s (1 + p^2/M_i^2). \quad (8.23)$$

The degree s must be chosen large enough to render all Feynman diagrams convergent. In the example of derivative-free interactions, we have shown that this implies $2s+2 > d$.

The corresponding propagator (8.20) cannot be derived from a Hermitian Hamiltonian and, thus, is non-physical for particle physics. Indeed, the Hermiticity of the Hamiltonian leads to Källen–Lehmann's representation representation (6.60) for the two-point function. If the propagator is, as above, a rational fraction, it must be a sum of poles with positive residues and thus, cannot decrease faster than $1/p^2$.

However, the modification (8.23), unlike more general modifications (Section A8.3), can be implemented also in Minkowski space (*i.e.* in real time), because the regularized propagators decrease in all complex p^2 directions.

8.4.3 Regulator fields

Still in the framework of renormalizable theories, momentum regularization has another implementation, which, in the simplest cases equivalent (but not for gauge theories), based on the introduction of regulator fields.

To regularize the action (8.21), one introduces additional dynamical scalar fields ϕ_i , $i = 1, \dots, s$, and considers the modified action

$$\begin{aligned} \mathcal{S}(\phi, \phi_i) = & \frac{1}{2} \int d^d x \left[\phi(x) (-\nabla^2 + m^2) \phi(x) + \sum_{i=1}^s \frac{1}{z_i} \phi_i(x) (-\nabla^2 + M_i^2) \phi_i(x) \right] \\ & + \mathcal{V}(\phi + \sum_i \phi_i). \end{aligned} \quad (8.24)$$

With the action (8.24), any internal ϕ propagator is replaced by the sum of the ϕ propagator and all the ϕ_i propagators $z_i/(p^2 + M_i^2)$. For an appropriate choice of the constants z_k , after integration over the regulator fields, the form (8.23) is recovered, as a simple calculation shows.

In the field integral

$$\int [d\phi] \prod_{i=1}^s [d\phi_i] \exp [-\mathcal{S}(\phi, \phi_i)], \quad (8.25)$$

we first change variables, setting $\phi(x) = \phi'(x) - \sum_{i=1}^s \phi_i(x)$. The action then becomes

$$\begin{aligned} \mathcal{S}(\phi', \phi_i) &= \frac{1}{2} \int d^d x \left[(\phi'(x) - \sum_i \phi_i(x)) (-\nabla^2 + m^2) (\phi'(x) - \sum_i \phi_i(x)) \right. \\ &\quad \left. + \sum_{i=1}^s \frac{1}{z_i} \phi_i(x) (-\nabla^2 + M_i^2) \phi_i(x) \right] + \mathcal{V}(\phi'). \end{aligned}$$

The Gaussian integration over the fields ϕ_i can be performed explicitly. A straightforward calculation leads, as expected, to

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x \phi(x) K(-\nabla^2) \phi(x) + \mathcal{V}(\phi),$$

with

$$[K]^{-1} = (-\nabla^2 + m^2)^{-1} + \sum_i z_i (-\nabla^2 + M_i^2)^{-1}.$$

It is then possible to choose the coefficients z_i in such a way that

$$\left[\frac{1}{-\nabla^2 + m^2} + \sum_i \frac{z_i}{-\nabla^2 + M_i^2} \right]^{-1} = (-\nabla^2 + m^2) \prod_i \frac{(-\nabla^2 + M_i^2)}{(-m^2 + M_i^2)}. \quad (8.26)$$

This corresponds to the choice

$$\tilde{K}(p) = (p^2 + m^2) \prod_i \frac{(p^2 + M_i^2)}{(-m^2 + M_i^2)}.$$

Note that the first condition, $1 + \sum_i z_i = 0$ shows that some z_i must be negative, and thus correspond to non-physical fields, as Källen–Lehmann’s representation implies (see Section 6.6). One then integrates over imaginary values of the corresponding fields.

8.5 Example: The $\phi_{d=6}^3$ field theory at one-loop order

In most of the chapter, we consider only actions even in the field. However, a simple illustration of the analysis of the nature of divergences of Feynman diagrams is provided by the renormalizable ϕ^3 scalar QFT in dimension 6. The ϕ^3 QFT is non-physical because the potential is not bounded from below. However, it has a well-defined perturbative expansion where this non-perturbative pathology is not visible. Moreover, it provides a simplified model of vacuum metastability (see Chapter 38). For g imaginary, it makes sense beyond perturbation theory in lower dimensions and describes the universal properties of the Yang–Lee edge singularity of the Ising model [49] (see Section 8.5.3).

In terms of a scalar field ϕ , the regularized action has the form

$$\begin{aligned} \mathcal{S}(\phi) &= \int d^6 x \left[\frac{1}{2} \phi(x) K(-\nabla^2) \phi(x) + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{3!} g \phi^3(x) \right. \\ &\quad \left. + v_1 \phi(x) + \frac{1}{2} v_2 \phi^2(x) \right], \end{aligned} \quad (8.27)$$

where m , v_1 , v_2 , and g , which is dimensionless, are constants.

At leading order $v_1 = v_2 = 0$, and $\phi = 0$ corresponds to the local minimum of the potential. Then, m is the physical mass in the tree approximation.

However, at higher orders, the ϕ^3 interaction shifts the mass and field expectation value to large non-physical values and the parameters v_1 and v_2 have to be tuned to maintain, for example, $\langle \phi \rangle = 0$ and the physical mass much smaller than the cut-off.

Tree approximation. In the tree approximation (Section 7.9.1), the generating functional of vertex (or 1PI) functions $\Gamma(\varphi)$ reduces to $\mathcal{S}(\varphi)$. At this order, in the infinite cut-off limit, the vertex or inverse two-point function is

$$\Gamma_{\text{tree}}^{(2)}(x, y) = (-\nabla^2 + m^2) \delta^{(6)}(x - y),$$

and, after Fourier transformation,

$$\tilde{\Gamma}_{\text{tree}}^{(2)}(p) = p^2 + m^2. \quad (8.28)$$

More generally, the Fourier components of the n -point vertex functions are

$$\tilde{\Gamma}_{\text{tree}}^{(3)}(p_1, p_2, -p_1 - p_2) = g, \quad \tilde{\Gamma}_{\text{tree}}^{(n)}(p_1, \dots, p_n) = 0, \quad \text{for } n > 3. \quad (8.29)$$

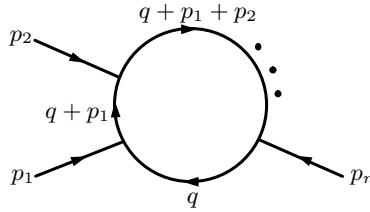


Fig. 8.4 The one-loop contribution to the n -point vertex function

8.5.1 Perturbation theory at one-loop order

The one-loop contribution $\Gamma_1(\varphi)$ to the functional $\Gamma(\varphi)$ has been derived in Section 7.9.2 (see also Section 7.7). From equation (7.93), and before regularization, one infers,

$$\Gamma_1(\varphi) = \frac{1}{2} \text{tr} \ln \left[1 + g (-\nabla^2 + m^2)^{-1} \varphi \right]. \quad (8.30)$$

The expansion of $\Gamma_1(\varphi)$ in powers of φ generates the one-loop contributions to the vertex functions $\Gamma^{(n)}$. After Fourier transformation, one finds

$$\begin{aligned} \tilde{\Gamma}_{\text{1 loop}}^{(n)}(p_1, \dots, p_n) &= -\frac{(n-1)!}{2} (-g)^n \int \frac{d^6 q}{(2\pi)^d} \frac{1}{q^2 + m^2} \frac{1}{(q + p_1)^2 + m^2} \dots \\ &\times \frac{1}{(q + p_1 + \dots + p_{n-1})^2 + m^2}, \end{aligned} \quad (8.31)$$

an expression represented by the Feynman diagram in Fig. 8.4.

For large-momentum q , the integrand in expression (8.31) behaves like $1/q^{2n}$, and the integral thus diverges for $2n \leq 6$, that is, for the one, two and three-point functions. To determine the divergent contributions explicitly, we expand the integrand in a Taylor series in the external momenta. It is easy to verify, using dimensional analysis, that the coefficients of the terms of global degree k in the momenta are given by integrals which diverge only for $6 \geq k + 2n$.

Therefore, the divergent part of a one-loop contribution to the n -point function is a polynomial of degree $(6 - 2n)$.

The essential observation is that, since the divergences are polynomials in the external momenta, the divergent contribution $\Gamma_1^{\text{div.}}(\varphi)$ to the functional $\Gamma(\varphi)$ is *local*, that is, it takes the form of the space integral of a function of the field and its derivatives, like the action itself (see Section A8.2 for a direct calculation).

To isolate more precisely a divergent part, we use one of the regularization methods discussed in Section 8.4.2. Cutting the momentum integral according to a regularization (8.23) with $M_i = \Lambda$, $s = 3$, one finds

$$\begin{aligned}\tilde{\Gamma}_{1\text{ loop}}^{(1)} &= \frac{g}{27\pi^3} \left[\frac{1}{4}\Lambda^4 - \frac{1}{4}m^2\Lambda^2 + m^4 \ln(\Lambda/m) + O(1) \right], \\ \tilde{\Gamma}_{1\text{ loop}}^{(2)} &= -\frac{g^2}{27\pi^3} \left[\frac{1}{10}\Lambda^2 - (2m^2 + p^2/3) \ln(\Lambda/m) + O(1) \right], \\ \tilde{\Gamma}_{1\text{ loop}}^{(3)} &= \frac{g^3}{26\pi^3} \ln(\Lambda/m) + O(1).\end{aligned}\quad (8.32)$$

The three divergent one-loop diagrams are displayed in Fig. 8.5. One verifies that the dimension $d = 6$ is special in the following sense: the vertex functions that diverge are all those that are already non-vanishing in the tree approximation (a term linear in ϕ can be added to the action (8.27) by translating ϕ by a constant). Moreover, the divergent terms and the tree approximation have the same momentum dependence.

By contrast, for dimensions $d \geq 8$, for example, the four-point function, which vanishes in the tree approximation, is also divergent.

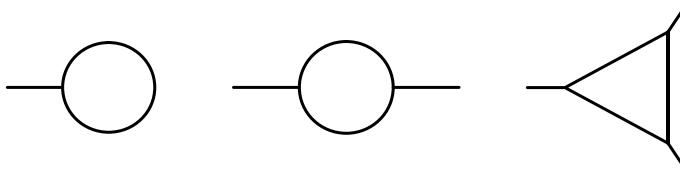


Fig. 8.5 Vertex functions: Divergent one-loop diagrams in the $\phi_{d=6}^3$ QFT

8.5.2 Analysis of the divergences at one-loop order

For the dimension $d = 6$, the divergent parts of the one-loop vertex functions have the structure of the initial action: $\Gamma_1^{\text{div.}}(\varphi)$, the divergent part at one-loop of $\Gamma(\varphi)$ has the structure

$$\begin{aligned}\Gamma_1^{\text{div.}}(\varphi) &= \int d^6x \left[\frac{1}{2}g^2a_0(\Lambda)(\nabla\varphi(x))^2 + ga_1(\Lambda)\varphi(x) + \frac{1}{2}g^2a_2(\Lambda)\varphi^2(x) \right. \\ &\quad \left. + \frac{1}{3!}g^3a_3(\Lambda)\varphi^3(x) \right].\end{aligned}\quad (8.33)$$

The functions $a_i(\Lambda)$ follow from equations (8.32) and are, therefore, defined only up to additive finite parts.

For later purpose (the minimal subtraction scheme), it is convenient to give a canonical definition of the divergent part of a Feynman diagram as the sum of the divergent terms in the asymptotic expansion in a dimensionless parameter. Choosing here Λ/m , one finds

$$\begin{aligned}2^7\pi^3a_0(\Lambda) &= \frac{1}{3}\ln(\Lambda/m), \\ 2^7\pi^3a_1(\Lambda) &= \frac{1}{4}\Lambda^4 - \frac{1}{4}m^2\Lambda^2 + m^4 \ln(\Lambda/m), \\ 2^7\pi^3a_2(\Lambda) &= -\frac{1}{10}\Lambda^2 + 2m^2 \ln(\Lambda/m), \\ 2^7\pi^3a_3(\Lambda) &= 2\ln(\Lambda/m).\end{aligned}\quad (8.34)$$

We first note that the field expectation value is now divergent and given by

$$\frac{\delta\Gamma}{\delta\varphi(x)} = m^2\varphi + ga_1 + O(g^3) = 0.$$

This leads to shift the field to cancel this expectation value, by setting

$$\varphi(x) = \varphi_r(x) - ga_1(\Lambda)/m^2,$$

or, alternatively, to cancel the linear term by choosing $v_1 = -ga_1(\Lambda)$.

Then, in order for the field to have finite correlations, it is necessary to renormalize it to cancel the divergent coefficient a_0 . One sets

$$\varphi(x) = \sqrt{Z}\varphi_r(x), \quad \text{with } Z = 1 - g^2a_0(\Lambda).$$

The next problem is the physical mass: if $v_2 = 0$, the mass is of order Λ and, thus, non-physical. This provides an example of the fine-tuning problem. It is necessary to cancel this divergence by adjusting the coefficient of φ^2 , for example, by setting

$$v_2 = -g^2a_2(\Lambda).$$

After the affine transformation,

$$\varphi(x) = Z^{1/2}\varphi_r(x) - ga_1(\Lambda),$$

and the tuning the mass parameter, only the logarithmic divergence of the contribution to the interaction remains. We define,

$$g_r = g + g^3a_3(\Lambda), \tag{8.35}$$

and then, the new vertex functional $\Gamma(\varphi_r)$, at one-loop order and for Λ large, is

$$\begin{aligned} \Gamma(\varphi_r) &= \int d^6x \left[\frac{1}{2}\varphi_r(x)(-\nabla^2)\phi(x) + \frac{1}{2}m^2\varphi_r^2(x) + \frac{1}{3!}g_r\varphi_r^3(x) \right] + \Gamma_1(\varphi_r) - \Gamma_1^{\text{div.}}(\varphi_r) \\ &\quad + O(\text{two loops}), \end{aligned}$$

an expression that now has a limit, *at g_r fixed*, when the cut-off becomes infinite.

Fixing g_r when the cut-off varies (this is the viewpoint adopted in the renormalization theory) amounts to an *additional fine tuning*. By contrast, if the bare coupling is fixed, g_r becomes a logarithmically running coupling constant, the running being a consequence of the non-decoupling of scales. This running is described by RG equations (see Section 9.11).

Finally, note that a change in the definition of the divergent part changes $\Gamma_1^{\text{div.}}(\varphi_r)$ by a finite local polynomial, and the conclusions are unchanged.

8.5.3 Universal properties of Yang–Lee’s edge singularity

In classical statistical physics, the Yang–Lee edge singularity is a singularity of the partition function of the Ising model for a small imaginary magnetic field when the temperature approaches the critical temperature from above [49]. Its universal properties are described by an $i\phi^3$ field theory corresponding to the unregularized action

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2}(\nabla\phi(x))^2 + \frac{1}{2}r\phi^2(x) + \frac{i}{3!}g\phi^3(x) \right], \tag{8.36}$$

where g and r are real constants. The action has a reflection symmetry $\mathcal{S}^*(\phi) = \mathcal{S}(-\phi)$. It leads to a well-defined field integral. The $d = 1$ example (quantum mechanics) has been thoroughly investigated. The perturbative expansion is related by analytic continuation to the one discussed in Section 8.5.

8.6 Operator insertions: Generating functionals, power counting

So far, we have analysed the divergences of field correlation functions. However, various physical problems involve correlation functions of non-linear local polynomials of the field, called hereafter composite fields or for historical reasons composite operators (this terminology comes from the operator formulation of QFT). Typical examples include

$$\mathcal{O}(\phi; x) \equiv \phi^2(x), \phi^4(x), [\nabla\phi(x)]^2 \dots .$$

One insertion of an operator $\mathcal{O}(\phi)$ yields the correlation functions

$$\langle \mathcal{O}(\phi; y)\phi(x_1) \cdots \phi(x_n) \rangle .$$

Such correlation functions can, in principle, be obtained from the field correlation functions by letting various points coincide. However, since with the initial functional measure fields are not continuous (or if \mathcal{O} involves derivatives, sufficiently differentiable), this limit is singular and leads, in momentum space, to additional integrations that generate new divergences. Therefore, it is necessary to analyse the vertex functions with operator insertions, from the point of view of power counting, separately.

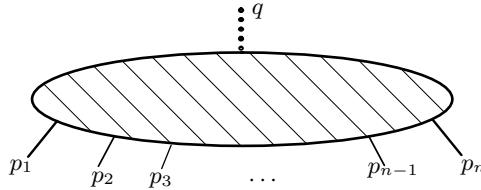


Fig. 8.6 The vertex function $\langle \mathcal{O}(q)\tilde{\phi}(p_1) \cdots \tilde{\phi}(p_n) \rangle$ (Fourier representation)

Generating functional. We consider Euclidean action $\mathcal{S}(\phi)$, to which we add a source (which is a space-dependent coupling constant) for the local operator $\mathcal{O}(\phi; x)$:

$$\mathcal{S}_g(\phi) = \mathcal{S}(\phi) + \int d^d x g(x) \mathcal{O}(\phi; x). \quad (8.37)$$

With the action $\mathcal{S}_g(\phi)$,

$$\mathcal{Z}(J, g) = \int [d\phi] \exp \left[-\mathcal{S}_g(\phi) + \int J(x) \phi(x) d^d x \right] \quad (8.38)$$

becomes the generating functional for correlation functions with \mathcal{O} insertions. The correlation functions with one operator $\mathcal{O}(\phi; x)$ insertion can be derived from the generating functional $\delta \mathcal{Z}/\delta g(x)$, taken at $g = 0$, by functional differentiation:

$$\langle \mathcal{O}(\phi; y)\phi(x_1) \cdots \phi(x_n) \rangle = - \left[\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \frac{\delta}{\delta g(y)} \mathcal{Z}(J, g) \right] \Big|_{J=g=0}. \quad (8.39)$$

More generally, successive differentiations with respect to $g(x)$ yield generating functionals of correlation functions with multiple operator insertions.

After Legendre transformation of $\mathcal{W}(J, g) = \ln \mathcal{Z}(J, g)$ with respect to $J(x)$, one obtains the vertex functional $\Gamma(\varphi, g)$. The generating functional of vertex functions with one $\mathcal{O}(\phi; y)$ insertion, $\Gamma_{\mathcal{O}}^{(n)}$, is then

$$\frac{\delta \Gamma(\varphi, g)}{\delta g(y)} \Big|_{g=0} = \sum_{n=1}^{\infty} \frac{1}{n!} \int d^d x_1 \cdots d^d x_n \varphi(x_1) \cdots \varphi(x_n) \Gamma_{\mathcal{O}}^{(n)}(y; x_1, \dots, x_n).$$

The corresponding Feynman diagrams have the structure displayed in Fig. 8.6. After Fourier transformation, they are just ordinary diagrams with one additional vertex $\mathcal{O}(\phi)$, except that an additional momentum enters the diagram at the vertex so that total momentum there is no longer conserved.

Power counting. When all integration momenta are scaled by a factor λ , in the limit $\lambda \rightarrow \infty$, all external momenta become negligible. Therefore, asymptotically, momentum is conserved even at the vertex corresponding to the operator insertion: the power counting of vertex functions with one operator insertion is the same as with one vertex insertion. We assign to an operator $\mathcal{O}(\phi)$ the dimension

$$[\mathcal{O}] = k + \sum n[\phi], \quad (8.40)$$

in which k is the number of derivatives in the operator, and n the number of fields. This definition differs from the definition of the dimension of the corresponding vertex (equation (8.9)) by d .

The expression (8.14) of the superficial degree of divergence δ_γ of an 1PI diagram γ is then modified in the case of the insertion of the product of operators $\mathcal{O}_1(x_1) \cdots \mathcal{O}_r(x_r)$, and becomes

$$\delta_\gamma(\mathcal{O}_1 \cdots \mathcal{O}_r) = d - E[\phi] + \sum_a v_a [\mathcal{V}_a] + [\mathcal{O}_1] + \cdots + [\mathcal{O}_r] - rd. \quad (8.41)$$

For example, for $d = 4$, one insertion of $\phi^m(x)$ in the ϕ^4 QFT with $\sigma = 2$ yields

$$[\phi^m] = m \Rightarrow \delta_\gamma = 4 - E + m - 4. \quad (8.42)$$

In the same theory, the n -point vertex function with $l \phi^2$ insertions,

$$\Gamma^{(n,l)}(x_1, \dots, x_n; y_1, \dots, y_l) = \langle \phi(x_1) \cdots \phi(x_n) \phi^2(y_1) \cdots \phi^2(y_l) \rangle_{\text{1PI}},$$

has the degree of divergence

$$\delta = 4 - n - 2l. \quad (8.43)$$

The new divergent correlation functions are displayed in Fig. 8.7.

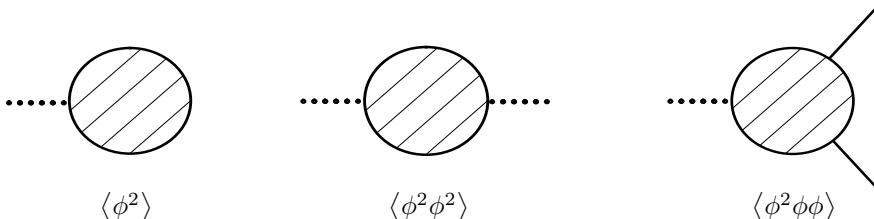


Fig. 8.7 Divergent vertex functions with ϕ^2 insertions (dotted lines)

Change of field parametrization. We show in Section 6.5.4 that the S -matrix of particle physics is formally invariant under local field transformations of the form

$$\phi(x) \mapsto \phi(x) + u_1\phi^3(x) + u_2\phi^5(x) \dots \quad (8.44)$$

Similarly, in the theory of critical phenomena, a change of this nature in the parametrization of the order parameter does not affect universal properties. One may then wonder about the meaning of correlation functions of a field in a specific parametrization. However, the coefficients u_n have a dimension $u_n \propto \Lambda^{-2n}$, and are thus very small. Therefore, the effect of a change of parametrization, like the addition of non-renormalizable interactions, simply induces a change in the parametrization of the renormalizable action (see Section 11.1).

8.7 Lattice regularization. Classical statistical physics

We shall meet examples where Pauli–Villars’s regularization does not work: field theories where quantization problems occur due products of non-commuting operators. This includes actions that have a definite geometric character like models on homogeneous spaces (for example, the non-linear σ -model) or gauge theories. As we have shown in Chapter 3, these problems already appear in path integrals of simple quantum mechanics. In particular, the forms of the propagator and of the interaction terms may not be independent. In this case, when the propagator is regularized, new more singular interactions have to be added to the action to preserve the symmetry and, as a few examples will illustrate, a class of one-loop diagrams cannot be regularized. Other regularization methods are needed. In many cases, lattice regularization of the Euclidean field theory (imaginary time) can be used [51, 52]. The advantages are the following:

- (i) Lattice regularization indeed corresponds to a specific choice of quantization.
- (ii) It is the only established regularization that always has a meaning beyond perturbation theory. Moreover, the regularized theory has the form of a statistical lattice model, interesting in its own right, because it can be studied by various standard methods of statistical physics.
- (iii) It preserves most global or local symmetries, with the exception of the space $O(d)$ symmetry which is replaced by a hypercubic symmetry, but this is not a major issue, as we argue in Chapters 14.4 and 15.

One obvious drawback is that it leads to very cumbersome perturbative calculations.

Scalar field theories. As a regularization, we may define the field integral (6.44) as the formal limit of an integral in which both time and space are discretized. We introduce a d -dimensional hypercubic lattice, with lattice spacing a , and use as dynamical variables the values of the field $\phi(x)$ at all lattice sites. To regularize the scalar field theory with field and action (8.2), one replaces the continuum by a discrete space lattice, which we choose here, for simplicity, to be hypercubic.

Derivatives $\partial_\mu \phi$ are replaced by finite differences, for example,

$$\partial_\mu \phi(x) \mapsto \nabla_\mu^{\text{lat}} \phi = \frac{1}{a} [\phi(x + an_\mu) - \phi(x)],$$

where x is a lattice site and n_μ the unit vector in the μ direction.

The regularized lattice action (8.2) then takes the form

$$\mathcal{S}_{\text{reg.}}(\phi) = a^d \sum_{x \in (a\mathbb{Z})^d} \left[\frac{1}{2} \sum_{\mu=1}^d [\nabla_\mu^{\text{lat}} \phi(x)]^2 + V(\phi(x)) \right]. \quad (8.45)$$

Correlation functions of lattice variables have also, as a formal limit, the continuum ϕ -field correlation functions.

The regularized theory can be considered as a statistical lattice model and the discretized action as the corresponding configuration energy. For $d \geq 2$, such a model does not generate, in general, a continuum limit. However, if the model has a continuous phase transition, and if one parameter, which plays the role of the temperature, is chosen asymptotically close to the transition value (which corresponds to the critical temperature), then a continuum limit can be defined.

The study of the continuum limit leads to establish relations between renormalization theory, as explained in the present chapter and Chapter 9, and the statistical theory of continuous phase transitions, which is discussed in Chapters 14–18, devoted to critical phenomena.

Fourier representation. The Fourier transform $\tilde{\phi}(p)$ of the field is a periodic function of cyclic momentum variables p_μ :

$$\tilde{\phi}(p) = \left(\frac{a}{2\pi}\right)^d \sum_{x \in (a\mathbb{Z})^d} \phi(x) e^{-ip \cdot x} \Leftrightarrow \phi(x) = \int d^d p \tilde{\phi}(p) e^{ip \cdot x}, \quad (8.46)$$

which can, therefore, be restricted to a Brillouin zone, for example,

$$-\pi/a \leq p_\mu < \pi/a, \quad \text{for } \mu = 1, \dots, d.$$

The corresponding propagator $\tilde{\Delta}_B(p)$ is then given by

$$\tilde{\Delta}_B^{-1}(p) = m^2 + \frac{2}{a^2} \sum_{\mu=1}^d (1 - \cos(ap_\mu)). \quad (8.47)$$

In the momentum representation, Feynman diagrams become periodic functions of the momentum components, with period $2\pi/a$. In the small lattice spacing limit, the continuum propagator is recovered since

$$\tilde{\Delta}_B^{-1}(p) = m^2 + p^2 - \frac{1}{12} \sum_{\mu} a^2 p_\mu^4 + O(a^4 p_\mu^6). \quad (8.48)$$

We note that hypercubic symmetry implies $O(d)$ symmetry at order p^2 and, thus, at large distance.

8.8 Effective QFT. The fine-tuning problem

Since in a QFT defined by a field integral involving a local classical action, divergences are unavoidable, we are forced to completely reconsider the very notion of local QFT.

First, we have to assume that a QFT is an incomplete theory, necessarily embedded in another finite non-local theory, but whose non-locality is confined to very short distances (the assumption of short range interactions for macroscopic phase transitions).

The problem of small masses, or large correlation lengths. For reasons that must be understood, the initial non-local theory generates particles with very low mass compared to the inverse microscopic scale. For example, the microscopic theory could imply gauge invariance and some form of chiral symmetry, which would explain low mass vector particles and fermions.

However, the problem of scalar particles is more complex. Goldstone bosons, which result from spontaneous symmetry breaking, are massless, but no such fundamental scalar particle has been observed. Supersymmetry (see Chapter 27) has been proposed, because it relates bosons to fermions but, so far, no sign of supersymmetry has been found in experiments (by 2020). Therefore, a problem of fine tuning remains, which involves adjusting precisely one parameter of the QFT to generate a small scalar mass. For macroscopic phase transitions, the divergence of the correlation length (equivalent to an inverse mass) is obtained by tuning the temperature at the critical temperature.

At the scale of the physical masses (the *physical scale*), the non-locality is only visible through the presence of UV divergences, which indicate a *non-decoupling of scales*. Therefore, at the physical scale, physics can be derived from *regularized local EFT*, also called Landau–Ginzburg–Wilson theory in the context of critical phenomena (see Section 15.1). We discuss the example of scalar field theories, but the concept extends to QFTs with fermions and gauge fields, relevant for particle physics.

The EFT is an approximation to a real microscopic theory, suitable to describe only large-distance or low-energy–momentum physics. It has the form of a QFT but where, in equation (8.19), $\tilde{K}(p)$ increases faster than any power for $|p| \rightarrow \infty$, and in expression (8.17), $\mathcal{V}_I(\phi)$ is a *linear combination of all local monomials* of the form (8.8), only restricted by symmetries (for simplicity, we consider only QFTs symmetrical in $\phi \mapsto -\phi$).

8.8.1 Effective action and perturbative assumption

We thus assume that the effective action has the general form (we include the contributions of the regularization terms),

$$\mathcal{S}(\phi) = \sum_{n \geq 2, k \geq 0, \alpha} \frac{1}{n!} a^{k-d} g_{k,n}^\alpha \mathcal{V}_{k,n}^\alpha(\phi), \quad (8.49)$$

where a is the microscopic scale (the lattice spacing for lattice models), and the powers of a are such that ϕ , $\mathcal{S}(\phi)$ and all coefficients $g_{k,n}^\alpha$ are dimensionless. The field ϕ is normalized such that $g_{2,2} = 1$. Generically, the coefficients are of order 1, and this, to some extent, characterizes the microscopic scale.

The perturbative assumption. The concept of EFT is based on the assumption of the relevance of some form of perturbation theory, an assumption for which there is empirical evidence. In particular, this means that the deviations from a free massless theory (the RG Gaussian fixed point, see Section 15.2), with the action

$$\mathcal{S}^*(\phi) = \frac{1}{2} a^{2-d} \int d^d x (\nabla_x \phi(x))^2, \quad (8.50)$$

are small (massless because masses are small with respect to the inverse microscopic length $1/a$), in some qualitative sense, which soon will become clearer. Since $\mathcal{S}^*(\phi)$ is the leading term, we change the field normalization, setting

$$\phi(x) = a^{(d-2)/2} \phi'(x),$$

giving the field a momentum dimension $(d-2)/2$. Then, the Gaussian action becomes

$$\mathcal{S}^*(\phi) = \frac{1}{2} \int d^d x (\nabla_x \phi(x))^2, \quad (8.51)$$

and the coefficient of $\mathcal{V}_{k,n}^\alpha(\phi)$ is changed to $a^{k-d+n(d-2)/2} g_{k,n}^\alpha / n!$.

In the theory of continuous phase transitions, this general scheme can be justified, to a large extent. In particle physics, it is an educated guess justified by its consequences.

8.8.2 Gaussian renormalization, dimensional analysis

To describe only large-distance physics, it is more convenient to take the physical scale as a reference, instead of the microscopic scale. We thus rescale distances $x \mapsto x'$, with

$$x = \lambda x', \quad \phi(x) = \lambda^{(2-d)/2} \phi'(x), \quad \text{with } \lambda \gg 1, \quad (8.52)$$

where the ϕ renormalization is such that \mathcal{S}^* remains invariant. Therefore, this transformation can be called a *Gaussian renormalization*. It has the form of an RG transformation, which has the Gaussian action $\mathcal{S}^*(\phi)$ as a fixed point (see Section 15.2).

Gaussian scaling of interactions and cut-off. The coefficient of $\mathcal{V}_{k,n}^\alpha(\phi)$ becomes

$$\lambda^{d-k-n(d-2)/2} a^{k-d+n(d-2)/2} g_{k,n}^\alpha / n!. \quad (8.53)$$

In the physical scale, $\Lambda = \lambda/a$ is a large momentum, characteristic of the microscopic scale. Instead of studying physics at large distances, or small momenta, one now studies physics for Λ large.

The scale parameter Λ is also the *cut-off* scale (in the QFT terminology), the scale at which the momentum integrals, in the Fourier representation of the perturbative expansion, must be cut, because the local expansion breaks down below the scale $1/\Lambda$ at which non-localities, which render the initial theory finite, become important.

The coefficient of the monomial $\mathcal{V}_{k,n}^\alpha$ (equation (8.8)) is transformed into

$$\Lambda^{d-n(d-2)/2-k} g_{k,n}^\alpha / n!. \quad (8.54)$$

8.8.3 The quadratic action and the fine-tuning problem

We collect all terms of the action quadratic in ϕ . The sum can be rewritten as

$$\mathcal{S}_2(\phi) = \frac{1}{2} \int d^d x \phi(x) \mathbf{K} \phi(x), \quad (8.55)$$

where \mathbf{K} is an operator, which has a derivative expansion and, therefore, in the momentum basis is diagonal. Its Fourier representation has a Fourier expansion of the form ($g_{0,2} > 0$ is assumed),

$$\mathbf{K} \mapsto \tilde{K}(p) = g_{0,2} \Lambda^2 + p^2 + g_{4,2} p^4 / \Lambda^2 - g_{6,2} p^6 / \Lambda^4 + \dots \quad (8.56)$$

For Λ large, $\tilde{K}(p)$ is dominated by the first term, $g_{0,2} \Lambda^2$. At leading order, in the absence of interactions, the physical mass m is of order $\sqrt{g_{0,2}} \Lambda$ and, for $g_{0,2} = O(1)$, non-physical. This is a first example of the *fine-tuning problem*. Indeed, a mass much smaller than the cut-off requires,

$$m = \sqrt{g_{0,2}} \Lambda \ll \Lambda, \quad \text{or } g_{0,2} = m^2 / \Lambda^2 \ll 1.$$

If this condition is satisfied, for $|p| = O(m)$ (the physical domain), all terms of degree p^2 and higher in the expansion vanish for $\Lambda \rightarrow \infty$.

In the presence of interactions, the fine-tuning issue remains, but $g_{0,2}$ must be tuned to be close to a different value (in renormalization theory, this is related to mass renormalization, see Section 9.2). In the same way, in macroscopic phase transitions, the correlation length is large with respect to the microscopic scale only for temperatures close to the critical temperature (see Chapter 14). This implies a tuning of the temperature.

8.9 The emergence of renormalizable field theories

The cut-off behaviour (8.54) of the monomial $\mathcal{V}_{k,n}^\alpha$ is directly opposite to its momentum or mass (infrared) dimension,

$$[\mathcal{V}_{k,n}^\alpha] = -d + n(d-2)/2 + k. \quad (8.57)$$

The study of the relative strength of the different terms in the action, at the physical scale, is thus reduced again to dimensional analysis or *power counting*. The analysis can also be formulated in terms of the *stability with respect to local perturbations of the free massless theory* [51], or Gaussian fixed point, in the RG terminology. Terms with a positive power of Λ , like $\Lambda^2 \int d^d x \phi^2(x)$ (negative dimension), which contribute to S_2 (equation (8.55) and have already been discussed, are the most important ones; they are called *relevant*. Dimensionless terms are called *marginal*, and terms that vanish for large Λ (positive dimension) are called *irrelevant*. Since the analysis depends on space dimension, we deal with different dimensions separately.

Dimension $d = 4$. Only the $\int d^d x \phi^2(x)$ term has a negative dimension and this leads to the issue of fine tuning. Then, the ϕ^4 interaction is dimensionless (the $\int d^d x (\nabla \phi)^2$ term is dimensionless by construction) and, thus, survives at a large distance. In the terminology of renormalization theory, it is called *renormalizable* (see Chapter 9), and in the RG terminology *marginal*. All other contributions have positive dimensions and thus vanish for $\Lambda \rightarrow \infty$. Therefore, one expects to be able to describe, at leading order, the large-distance, or equivalently small-mass and small-momentum physics with the *renormalizable* action, which we parametrize now as

$$\mathcal{S}(\phi) = \int d^4 x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \phi^4(x) \right], \quad (8.58)$$

where r and $g \geq 0$ are two constants. However, as we have already discussed in Section 8.1, one has to retain at least the few first terms of the expansion of $\tilde{K}(p)$ in powers of p^2 to render perturbation theory finite and thus to satisfy the condition (8.6), $\sigma > 4$.

In this parametrization, the physical mass vanishes for a special value r_c of the parameter r and the condition that the mass should be much smaller than Λ is equivalent to the condition $|r - r_c| \ll \Lambda^2$, which involves a fine tuning of r .

The analysis shows that the *condition of renormalizability*, discovered empirically in particle physics, *is not a law of nature*, but is simply the consequence of a few general assumptions that we have described. Moreover, these assumptions are known to be satisfied for a large class of macroscopic critical phenomena (see Chapters 14 and 15).

Dimension $d = 3$. For $d = 3$, the field has dimension $\frac{1}{2}$. The leading interaction is ϕ^4 , which has dimension -1 and thus is proportional to Λ . The interaction ϕ^6 is dimensionless. All other contributions have positive dimensions and vanish for $\Lambda \rightarrow \infty$.

If only the ϕ^2 coefficient is tuned, in order to generate a small mass, generically, the interaction diverges with the cut-off Λ . Retaining the leading interaction, one finds the action

$$\mathcal{S}(\phi) = \int d^3 x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \Lambda \phi^4(x) \right],$$

where g is dimensionless (and a cut-off is implied). This strong coupling situation is typical of critical phenomena when only the temperature can be adjusted.

By contrast, by tuning the ϕ^4 coupling such that $g = O(m/\Lambda)$, where m is the physical mass, and, in addition, cancelling the ϕ^6 interaction, from the viewpoint of renormalization theory, one obtains a *super-renormalizable* QFT. Then, after tuning of r and at $g\Lambda$ fixed, the resulting theory is finite in the infinite cut-off ($\Lambda \rightarrow \infty$) limit.

Otherwise, by tuning only the coefficient of ϕ^4 (in addition to the coefficient of ϕ^2) to a critical value, one obtains a theory in which the interaction is dominated by the marginal ϕ^6 term and the field theory is renormalizable. This field theory describes a *tricritical behaviour* in critical phenomena.

Dimension $d = 2$. This dimension is peculiar because the dimension of the field vanishes. All derivative-free terms are proportional to Λ^2 and terms with two derivatives are dimensionless, corresponding to renormalizable interactions. This case requires a specific discussion.

8.9.1 Non-renormalizable interactions: The example of four dimensions

Non-renormalizable interactions, like the ϕ^6 interaction, which cannot be dealt with by renormalization theory, appear in the EFT framework as quite innocuous, because they are multiplied by negative powers of Λ . They lead to very weak interactions. However, when added as a perturbation in the ϕ^4 QFT, they generate increasing divergences (this is the reason why they have been excluded in the construction of renormalizable QFTs). These divergences are partially cancelled by the powers of Λ (8.54). Moreover, a study of the *renormalization of local monomials of the fields*, also called composite operators [51], shows that the contributions that do not vanish for $\Lambda \rightarrow \infty$, or infinite cut-off, simply renormalize (like counter-terms, Section 9.2.1) the parameters of the renormalizable part of the action. For example, the operators of dimension 6 generate divergent contributions of dimension 2 and 4 that renormalize the coefficients of terms already present in the action, and a contribution decreasing as $1/\Lambda^2$, up to powers of logarithms, with the cut-off (see Section 11.1.4).

Field parametrization. It is quite possible that, in the microscopic theory, the field is just a parametrization of some manifold (see also Chapters 19, 21–29). Therefore, it could be replaced by any other field ϕ' related to ϕ by

$$\phi'(x) = \phi(x) + c_1\phi^2(x) + c_2\phi^3(x) + \dots, \quad (8.59)$$

where the c_n are constants of order 1 (note that this *strictly local* change of variables is defined only with lattice regularization). However, after the rescaling (8.52), at the physical scale c_n transforms into $c_n\Lambda^{-n(d-2)/2}$ and leads to changes only in the parametrization of the renormalizable and non-renormalizable parts of the action. Therefore, the renormalized perturbation theory is not affected.

Beyond perturbation theory. The analysis of the hierarchy of interactions is based here on power counting, or Gaussian renormalization, equivalent to mean-field theory (see Chapter 14). Its validity, beyond leading order perturbation theory, relies on the assumption that a global RG analysis, which takes into account the scale non-decoupling, would not qualitatively modify the hierarchy of interactions. Such a global study requires solving, to some extent, *functional RG equations* [61–63]. However, there is empirical evidence that, for four-dimensional QFT as relevant for particle physics, as well as for a large class of two- and three-dimensional QFTs, relevant to the theory of macroscopic phase transitions, the general hierarchy of interactions is not modified, the irrelevant or non-renormalizable interactions remain irrelevant, and a much simpler perturbative RG is sufficient. In four dimensions, the Gaussian scaling is then only modified by logarithms. Still, the triviality issue emerges (see Sections 9.11 and 9.12).

A8 Technical details

A8.1 Schwinger's proper-time representation

We first establish a formal representation of the one-loop contribution to the generating functional of vertex functions $\Gamma(\varphi)$. In Section 7.9, we have shown that $\Gamma_{\text{1 loop}}(\varphi)$ is given by

$$\Gamma_{\text{1 loop}}(\varphi) = \frac{1}{2} \text{tr} \left[\ln \frac{\delta^2 \mathcal{S}}{\delta \varphi(x_1) \delta \varphi(x_2)} - \ln \frac{\delta^2 \mathcal{S}}{\delta \varphi(x_1) \delta \varphi(x_2)} \Big|_{\varphi=0} \right]. \quad (\text{A8.1})$$

For example, if $\mathcal{S}(\phi)$ is

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + U(\phi(x)) \right], \quad (\text{A8.2})$$

where $U(\phi)$ is a polynomial, the second functional derivative $\delta^2 \mathcal{S}/\delta \varphi(x_1) \delta \varphi(x_2)$ takes the form of a quantum Hamiltonian:

$$M(x_1, x_2) \equiv \frac{\delta^2 \mathcal{S}}{\delta \varphi(x_1) \delta \varphi(x_2)} = [-\nabla^2 + m^2 + U''(\varphi(x_1))] \delta^{(d)}(x_1 - x_2). \quad (\text{A8.3})$$

We also define

$$M_0(x_1, x_2) \equiv \frac{\delta^2 \mathcal{S}}{\delta \varphi \delta \varphi} \Big|_{\varphi=0} = (-\nabla^2 + m^2) \delta^{(d)}(x_1 - x_2). \quad (\text{A8.4})$$

The general identity,

$$\text{tr} (\ln M - \ln M_0) = - \int_0^\infty \frac{dt}{t} \text{tr} (e^{-tM} - e^{-tM_0}), \quad (\text{A8.5})$$

then leads to a compact representation of the one-loop functional $\Gamma_{\text{1 loop}}(\varphi)$ as an integral over Schwinger's proper time:

$$\Gamma_{\text{1 loop}}(\varphi) = -\frac{1}{2} \int_0^\infty \frac{dt}{t} \text{tr} (e^{-tM} - e^{-tM_0}). \quad (\text{A8.6})$$

The methods that are used in quantum mechanics to calculate the statistical operator $e^{-\beta H}$, can again be used here.

A8.2 Regularization and one-loop divergences

In the representation (A8.6), large momentum divergences appear as divergences at $t = 0$ and, therefore, the determination of one-loop divergences is reduced to the small t expansion of the diagonal matrix elements $\langle x | e^{-tM} | x \rangle$ (we use the bra–ket notation of quantum mechanics) for a Schrödinger-like operator M . This is a problem we have already faced in Section 2.2, and which can be solved, for example, by using Schrödinger's equation.

The expression (A8.6) can be regularized by the various methods explained in the chapter. For instance, we can multiply it by a cutting factor $\rho(t\Lambda^2)$ and we then recover the regularization defined by equation (A8.23).

Schwinger's proper-time regularization. Schwinger's proper-time regularization consists in simply cutting the t integral at a small value ε . Setting, for convenience,

$$M = H + m^2, \quad M_0 = H_0 + m^2,$$

we obtain the regularized expression

$$\Gamma_{1\text{-loop}}^{\text{reg.}}(\varphi) = -\frac{1}{2} \int_{\varepsilon}^{\infty} \frac{dt}{t} e^{-m^2 t} \text{tr} (e^{-tH} - e^{-tH_0}). \quad (\text{A8.7})$$

For illustration purpose, we expand for t small the integrand in the case of the action (A8.2). Setting

$$U''(\phi(x)) = V(x),$$

we first expand the solution of the Schrödinger equation,

$$[-\nabla_x^2 + V(x)] \langle x | e^{-tH} | x' \rangle = -\frac{\partial}{\partial t} \langle x | e^{-tH} | x' \rangle. \quad (\text{A8.8})$$

We set

$$\langle x | e^{-tH} | x' \rangle = e^{-\sigma(x, x'; t)}. \quad (\text{A8.9})$$

The Schrödinger equation then leads to

$$\nabla^2 \sigma(t, x) - (\nabla \sigma(t, x))^2 + V(x) = \frac{\partial \sigma(t, x)}{\partial t}. \quad (\text{A8.10})$$

The function σ has, for $t \rightarrow 0$, an expansion of the form

$$\sigma(t, x) = \frac{1}{4t} (x - x')^2 + \frac{d}{2} \ln 4\pi t + A(x, x')t + B(x, x')t^2 + C(x, x')t^3 + O(t^4). \quad (\text{A8.11})$$

We obtain for the coefficients A , B , and C the equations,

$$\begin{aligned} A + (\mathbf{x} - \mathbf{x}') \cdot \nabla A &= V(x), \\ 2B + (\mathbf{x} - \mathbf{x}') \cdot \nabla B &= \nabla^2 A, \\ 3C + (\mathbf{x} - \mathbf{x}') \cdot \nabla C &= \nabla^2 B - (\nabla A)^2. \end{aligned} \quad (\text{A8.12})$$

The solutions for A and B are

$$\begin{aligned} A(x, x') &= \int_0^1 ds V(x' + s(x - x')), \\ B(x, x') &= \int_0^1 ds s(1-s) \nabla^2 V(x' + s(x - x')). \end{aligned} \quad (\text{A8.13})$$

It follows that

$$A(x, x) = V(x), \quad B(x, x) = \frac{1}{6} \nabla^2 V(x), \quad C(x, x) = -\frac{1}{2} (\nabla V(x))^2 + \frac{1}{20} \nabla^4 V(x).$$

The divergent part of the regularized expression (A8.7) comes from the contribution I_{ε} of the lower bound, near which we can use the small t expansion. Total derivatives disappear in the trace. After an integration by parts, one obtains

$$\begin{aligned} I_{\varepsilon} &= -\frac{1}{2} \frac{1}{(4\pi)^{d/2}} \int_{\varepsilon}^{\infty} \frac{dt e^{-m^2 t}}{t^{1+d/2}} \int d^d x \left\{ -V(x)t + \frac{1}{2} V^2(x)t^2 - \frac{1}{6} \left[V^3(x) + \frac{1}{2} (\nabla V(x))^2 \right] t^3 \right\} \\ &\quad + O(t^4). \end{aligned} \quad (\text{A8.14})$$

Finally, keeping only the contribution of the lower bound of the t integration, one finds

$$I_\varepsilon = \frac{1}{2} \frac{1}{(4\pi)^{d/2}} \left\{ -\frac{\varepsilon^{1-d/2}}{1-d/2} \int d^d x V(x) + \frac{1}{2} \frac{\varepsilon^{2-d/2}}{2-d/2} \int d^d x (V^2(x) + 2m^2 V(x)) \right. \\ \left. - \frac{1}{6} \frac{\varepsilon^{3-d/2}}{3-d/2} \int d^d x [V^3(x) + 3m^2 V^2(x) + 3m^4 V(x) + \frac{1}{2} (\nabla V(x))^2] \right\} + \dots \quad (A8.15)$$

When d is an even integer, $\varepsilon^0/0$ has to be replaced by $\ln(1/\varepsilon)$. This expression gives all divergences for $d \leq 6$.

For example, we can apply this result to the interaction $U(\phi) = g\phi^3/3!$ in six dimensions. To compare with other regularizations, we set $\varepsilon = 1/\Lambda^2$. Then,

$$\Gamma_{1\text{loop}}^{\text{div.}}(\varphi) = \frac{1}{2^7 \pi^3} \int d^6 x \left\{ \frac{\Lambda^4}{2} g \varphi(x) - \frac{\Lambda^2}{2} [g^2 \varphi^2(x) + gm^2 \varphi(x)] \right. \\ \left. + \frac{1}{3} \ln \frac{\Lambda}{m} \left[g^3 \varphi^3(x) + 3g^2 m^2 \varphi^2(x) + 3gm^4 \varphi(x) + \frac{g^2}{2} (\nabla \varphi(x))^2 \right] \right\}. \quad (A8.16)$$

This leads to the results of equations (8.33) and (8.34).

For the interaction $U(\phi) = g\phi^4/4!$ in four dimensions, the result is

$$\Gamma_{1\text{loop}}^{\text{div.}} = \frac{1}{32\pi^2} \left\{ \frac{\Lambda^2}{2} g \int d^4 x \varphi^2(x) - \ln \frac{\Lambda}{m} \int d^4 x \left[\frac{g^2}{4} \varphi^4(x) + gm^2 \varphi^2(x) \right] \right\}. \quad (A8.17)$$

An identical expression will be recovered in Section 9.3, equation (9.29). Both in equations (A8.16) and (A8.17), we have defined the divergent part of $\Gamma(\varphi)$ as the sum of the divergent terms in the asymptotic expansion for Λ/m large.

An application of equation (A8.15) to $d = 2$ and a general interaction $U(\phi)$ yields

$$\Gamma_{1\text{loop}}^{\text{div.}} = \frac{1}{4\pi} \ln \frac{\Lambda}{m} \int d^2 x U''(\varphi(x)). \quad (A8.18)$$

Although, in these examples, the results can easily be recovered from the Feynman graph expansion, in more complicated cases, in which symmetries play an essential role, this method can be quite useful to evaluate divergences of one-loop diagrams.

ζ -function regularization. A variant of the preceding regularization method is to replace expression (A8.6) by

$$\Gamma_{1\text{loop}}^{\text{reg.}}(\varphi) = -\frac{1}{2\Gamma(1+\mu)} \int_0^\infty dt t^{\mu-1} \text{tr} (e^{-tM} - e^{-tM_0}), \quad (A8.19)$$

and to take, after analytic continuation in μ , the limit $\mu = 0$, in the spirit of the dimensional regularization.

We again consider the example of the ϕ^4 field theory in four dimensions and calculate $\Gamma_{1\text{loop}}$ per unit volume for a constant field φ (V is the volume and $V \rightarrow \infty$):

$$\frac{1}{V} \Gamma_{1\text{loop}}^{\text{reg.}}(\varphi) = -\frac{1}{2\Gamma(1+\mu)} \int_0^\infty dt t^{\mu-1} \int \frac{d^4 p}{(2\pi)^4} [e^{-t(p^2+m^2+g\varphi^2/2)} - (\varphi=0)].$$

The integration over the momentum p yields

$$\frac{1}{V} \Gamma_{\text{1 loop}}^{\text{reg}}(\varphi) = -\frac{1}{32\pi^2\Gamma(1+\mu)} \int_0^\infty dt t^{\mu-3} \left[e^{-t(m^2+g\varphi^2/2)} - (\varphi=0) \right]. \quad (A8.20)$$

The integration over t can then also be performed:

$$\frac{1}{V} \Gamma_{\text{1 loop}}^{\text{reg.}}(\varphi) = -\frac{1}{32\pi^2\mu(\mu-1)(\mu-2)} \left[(m^2 + g\varphi^2/2)^{(2-\mu)} - (\varphi=0) \right]. \quad (A8.21)$$

Finally, expanding the expression for μ small and keeping only the divergent and finite parts, one obtains (to be compared with expression (A8.17))

$$\frac{1}{V} \Gamma_{\text{1 loop}}^{\text{reg}}(\varphi) = \frac{1}{64\pi^2} \left(m^2 + \frac{g}{2}\varphi^2 \right)^2 \left[-\frac{1}{\mu} + \ln \left(m^2 + \frac{g}{2}\varphi^2 \right) - \frac{3}{2} \right] - (\varphi=0). \quad (A8.22)$$

The coefficient of the divergent part differs by a factor 2 from the one obtained in dimensional regularization: at leading order $\mu \sim \varepsilon/2$ (see, *e.g.*, equations (10.32)).

A8.3 More general momentum regularizations

Regularizations based on Schwinger's proper-time representation (see equations (A8.7, A8.19)) suggest the more general form of a regularized propagator,

$$\Delta_B(p) = \int_0^\infty dt \rho(t\Lambda^2) e^{-t(p^2+m^2)}, \quad (A8.23)$$

in which the function $\rho(t)$ is positive and satisfies the condition

$$|1 - \rho(t)| < C e^{-\sigma t} \quad (\sigma > 0) \text{ for } t \rightarrow +\infty.$$

By choosing a function $\rho(t)$ that decreases fast enough for $t \rightarrow 0$, the behaviour of the propagator can be arbitrarily improved. If $\rho(t) = O(t^n)$, the behaviour (8.23) is recovered. Another example is

$$\rho(t) = \theta(t-1), \quad (A8.24)$$

$\theta(t)$ being the step function, which leads to exponential decrease:

$$\Delta_B(p) = \frac{e^{-(p^2+m^2)/\Lambda^2}}{p^2 + m^2}. \quad (A8.25)$$

As the example (A8.25) shows, it is possible to find propagators without non-physical singularities in this more general class, but they do not follow from a Hamiltonian formalism because continuation to real time becomes impossible. On the other hand, arbitrary local interaction monomials can be regularized.

9 Introduction to renormalization theory and renormalization group (RG)

In Chapter 8, we have shown that a straightforward construction of a local, relativistic quantum field theory (QFT) leads to ultraviolet (UV) divergences, and that a QFT has to be regularized by modifying its short-distance or large-energy–momentum structure. Since such a modification is somewhat arbitrary, it is necessary to prove that the resulting large scale predictions are, at least to a large extent, short-distance insensitive.

Such a proof relies on the renormalization theory, as initiated in Refs. [24, 25, 54], and the corresponding RG [55] (for early works, see Refs. [56]). Therefore, we describe here the essential steps of a proof of the *perturbative renormalizability* of the scalar ϕ^4 QFT in dimension 4. All the basic difficulties of renormalization theory are already present in this simple example, and it will eventually become apparent how to extend the method, based on power counting, to other theories (however, the preservation of symmetries requires a specific discussion).

We have followed the elegant presentation of Callan [57, 58], which makes it possible to prove renormalizability, and RG equations (in Callan–Symanzik’s form [60]) simultaneously. This presentation is especially suited to our general purpose, since a large part of this work is devoted to applications of RG. Moreover, it already emphasizes, at the technical level, the direct relation between renormalizability and the existence of an RG. Although this may not always be obvious, the background of the discussion is effective field theory (EFT), and emergent renormalizable theory (see Sections 8.8.1 and 8.9).

As a technical tool, we define the initial unrenormalized theory by *momentum cut-off regularization*, although this leads to a derivation slightly more complicated than dimensional regularization. However, dimensional regularization has no direct physical interpretation and, moreover, already performs a partial renormalization, since it cancels what in momentum regularization are power-law divergences. This may lead to erroneous conclusions from the physics viewpoint. By contrast, the inverse of the momentum cut-off is an artificial replacement for a true initial short-distance scale below which the *local EFT* is no longer meaningful.

Renormalization involves a renormalization of the field necessary to generate a limiting distribution, something expected when one sums over an infinite number of random variables (*cf.* the central limit theorem of probabilities). However, we have already pointed out that it also requires *a tuning of the parameters of the initial action as a function of the regularization parameter*, something non-physical in the framework of particle physics and which, even in statistical physics, is peculiar. *The interpretation of QFTs as effective low energy or large distance theories* (EFTs), and the RG formulated in terms of initial (bare) parameters, provide a framework for discussing this issue.

The proof of renormalizability given here applies only to massive theories and, thus, in Section 9.9, we discuss the existence of a massless theory. A different, homogeneous form of RG equations follows. In Section 9.10.1, we discuss the covariance of RG functions.

Another, non-perturbative RG approach, which is outlined in Appendix A9.1, deals with a general EFT. It has been employed to give an alternative proof of renormalizability [59]. It relies on a partial integration of large momentum modes in a form proposed by Wegner and Wilson [61–63], and reduces to the QFT bare RG in the perturbative regime.

It is described in the spirit of this work in Chapter 16 of Ref. [64].

Appendix A9.2 contains a few remarks about divergences in super-renormalizable theories and the technique of normal-ordering.

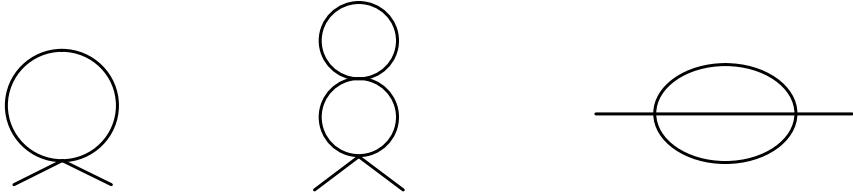


Fig. 9.1 Contributions to the ϕ -field two-point vertex function: $l = 0, n = 2, \delta = 2$

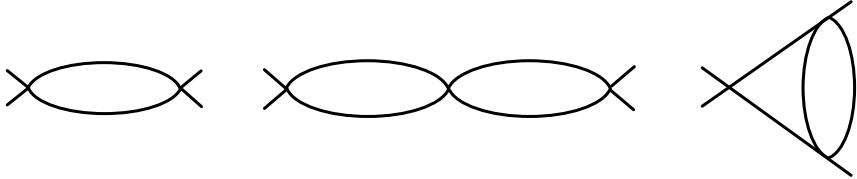


Fig. 9.2 Contributions to the ϕ -field four-point vertex function: $l = 0, n = 4, \delta = 0$

9.1 Power counting. Dimensional analysis

We discuss the divergences of the correlation functions of a relativistic QFT corresponding to the classical, non-regularized, Euclidean action for a scalar field $\phi(x)$ in four dimensions (the dimension relevant to particle physics),

$$\mathcal{S}(\phi) = \int d^4x \left[\frac{1}{2} (\nabla\phi(x))^2 + \frac{1}{2} r\phi^2(x) + \frac{1}{4!} g\phi^4(x) \right], \quad (9.1)$$

where $\nabla \equiv \{\partial/\partial x_1, \dots, \partial/\partial x_4\}$, and r and $g \geq 0$ are two constant parameters.

When the coupling constant g vanishes, the coefficient r of ϕ^2 is the physical mass squared but, for $g > 0$, the existence of a necessary phase transition requires $r < 0$.

Renormalization theory is mainly formulated in terms of *vertex (one-particle irreducible or 1PI) functions*,

$$\Gamma^{(n)}(x_1, \dots, x_n) = \langle \phi(x_1) \cdots \phi(x_n) \rangle_{\text{1PI}},$$

(in the notations of equation (8.1)) which have simpler properties than correlation functions. These are generated by the functional $\Gamma(\varphi)$, Legendre transform of the generating functional of connected correlation functions (Section 7.7),

$$\Gamma(\varphi) = \sum_{n=0} \frac{1}{n!} \int \Gamma^{(n)}(x_1, \dots, x_n) \prod_{i=1}^n d^4x_i \varphi(x_i). \quad (9.2)$$

Power counting in four dimensions. In Section 8.3, we have shown that the ϕ^4 vertex has UV dimension 0 (which in a scalar QFT is the same as mass dimension) and the action is renormalizable in the sense of power counting: the superficial degree of divergence of vertex functions is independent of the order in perturbation theory.

The degree of divergence δ of the vertex function $\Gamma^{(n)}$ is $\delta = 4 - n$ (equations (8.9) and (8.14)). In Figs. 9.1 and 9.2, the superficially divergent one- and two-loop Feynman diagrams corresponding to the two-point and four-point functions are displayed.

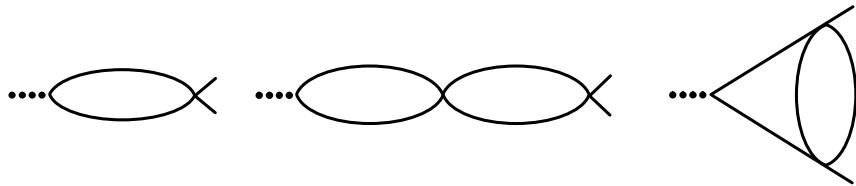


Fig. 9.3 Contributions to the $\langle\phi^2\phi\phi\rangle$ vertex function: $l = 1, n = 2, \delta = 0$

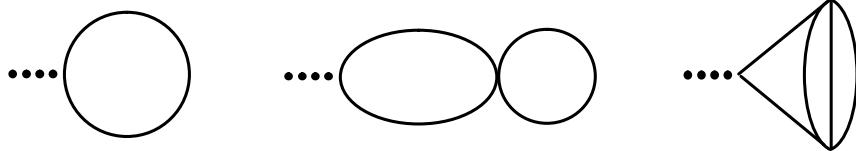


Fig. 9.4 Contributions to the constant $\langle\phi^2\rangle$ expectation value: $l = 1, n = 0, \delta = 2$

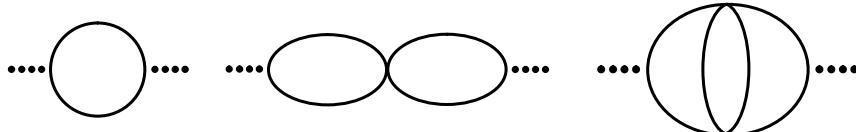


Fig. 9.5 Contributions to the $\langle\phi^2\phi^2\rangle$ correlation function: $l = 2, n = 0, \delta = 0$

$\phi^2(x)$ insertions. We also need vertex functions with $\phi^2(x)$ insertions. The degree δ of divergence of the vertex function with l insertions of $\frac{1}{2}\phi^2$,

$$\Gamma^{(l,n)}(y_1, \dots, y_l; x_1, \dots, x_n) = 2^{-l} \langle \phi^2(y_1) \cdots \phi^2(y_l) \phi(x_1) \cdots \phi(x_n) \rangle_{\text{1PI}},$$

is $\delta = 4 - n - 2l$ (equation (8.43)).

In Figs. 9.3–9.5, the first few superficially divergent diagrams corresponding to the new divergent functions are displayed. Diagrams with $n = 0$ never arise as subdiagrams, since they have no external lines, and can be discussed separately.

9.2 Regularization. Bare and renormalized QFT

To define a finite perturbation theory, we replace the Euclidean action $\mathcal{S}(\phi)$ by a regularized action $\mathcal{S}_\Lambda(\phi)$, using one of the regularization schemes described in Section 8.4.2. We refer explicitly to a cut-off, because it is in the spirit of this work. Some arguments would have to be slightly modified in the case of dimensional regularization.

Cut-off regularization. We introduce the regularized action (the subscript Λ emphasizes that we are using a momentum-cut-off scheme),

$$\mathcal{S}_\Lambda(\phi) = \int d^4x \left[\frac{1}{2} (\nabla_\Lambda \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \phi^4(x) \right], \quad (9.3)$$

where (a form of Pauli–Villars’s regularization),

$$\nabla_\Lambda \equiv \nabla \prod_{i=1}^s \left(1 - \nabla^2 / M_i^2 \right)^{1/2}, \quad (9.4)$$

and the masses $M_i > 0$ are of the order of the cut-off Λ .

To generate finite ϕ and ϕ^2 vertex functions in four dimensions, $s = 2$ suffices (the fields contributing to the field integral then become continuous). The regularization yields propagators, in the Fourier representation, of the form

$$[\tilde{\Delta}(p)]^{-1} = p^2 \prod_{i=1}^2 (1 + p_i^2/M_i^2) + r.$$

Dimensional analysis. In what follows, mass dimensional analysis, which in a scalar QFT is equivalent to UV dimension, is useful: the action is dimensionless, space has dimension -1 , by convention. Then, the field ϕ has mass dimension 1, as well as the cut-off Λ . The parameter r has dimension 2 and the coupling g is dimensionless.

Connected correlation functions satisfy

$$W^{(n)}(x_1/\lambda, \dots, x_n/\lambda; \lambda^2 r, g, \lambda\Lambda) = \lambda^n W^{(n)}(x_1, \dots, x_n; r, g, \Lambda). \quad (9.5)$$

After Fourier transformation and factorization of the δ -function of momentum conservation, one obtains

$$\widetilde{W}^{(n)}(\lambda p_1, \dots, \lambda p_n; \lambda^2 r, g, \lambda\Lambda) = \lambda^{4-3n} \widetilde{W}^{(n)}(p_1, \dots, p_n; r, g, \Lambda). \quad (9.6)$$

From the expansion (9.2), one infers the scaling of the vertex function $\Gamma^{(n)}(x_1, \dots, x_n)$, and its Fourier transform $\tilde{\Gamma}^{(n)}(p_1, \dots, p_n)$ (after factorization of a δ -function):

$$\begin{aligned} \Gamma^{(n)}(x_1/\lambda, \dots, x_n/\lambda; \lambda^2 r, g, \lambda\Lambda) &= \lambda^{3n} \Gamma^{(n)}(x_1, \dots, x_n; r, g, \Lambda), \\ \tilde{\Gamma}^{(n)}(\lambda p_1, \dots, \lambda p_n; \lambda^2 r, g, \lambda\Lambda) &= \lambda^{4-n} \tilde{\Gamma}^{(n)}(p_1, \dots, p_n; r, g, \Lambda). \end{aligned}$$

The function $\tilde{\Gamma}^{(n)}$ has a mass dimension $(4 - n)$, which coincides with its UV dimension in the sense of power counting.

The argument generalizes to

$$\tilde{\Gamma}^{(l,n)}(\lambda q_i, \lambda p_j; \lambda^2 r, g, \lambda\Lambda) = \lambda^{4-n-2l} \tilde{\Gamma}^{(l,n)}(q_i, p_j; r, g, \Lambda). \quad (9.7)$$

9.2.1 Bare and renormalized action: Counter-terms

Before we discuss renormalization, a few definitions are required. The action (9.3) is often called the *bare action*. The parameters r and g that appear in the action are then called *bare parameters*; the correlation functions of the *bare* field ϕ are *bare correlation functions*. From the physics viewpoint, the bare action is obtained from the *effective action at the cut-off scale* by performing a Gaussian renormalization (equation (8.52)) and neglecting all non-renormalizable interactions.

To the bare parameters r and g correspond two *renormalized parameters* (or effective parameters at the physical scale) m_r and g_r : m_r characterizes the imaginary-time decay of correlation functions and is proportional to the physical mass of the theory; g_r is dimensionless and becomes the expansion parameter of the renormalized theory.

We define a renormalized field ϕ_r obtained from the bare field by a rescaling,

$$\phi(x) = Z^{1/2} \phi_r(x). \quad (9.8)$$

The correlation functions of the field ϕ_r are the *renormalized correlation functions*.

We also define the *renormalized action* \mathcal{S}_r , which is the initial action expressed in terms of the renormalized field ϕ_r and renormalized parameters, which we write as

$$\mathcal{S}_\Lambda(\phi) \equiv \mathcal{S}_r(\phi_r) = \mathcal{S}_{r,0}(\phi_r) + \mathcal{S}_{C.T.}(\phi_r), \quad (9.9)$$

where $\mathcal{S}_{r,0}(\phi_r)$ is the *tree order action* (which, for $\Lambda \rightarrow \infty$, reduces to the classical action):

$$\mathcal{S}_{r,0}(\phi_r) = \int d^4x \left[\frac{1}{2} (\nabla_\Lambda \phi_r(x))^2 + \frac{1}{2} m_r^2 \phi_r^2(x) + \frac{1}{4!} g_r \phi_r^4(x) \right], \quad (9.10)$$

$\mathcal{S}_{C.T.}(\phi_r)$ is the sum of *counter-terms*:

$$\mathcal{S}_{C.T.}(\phi_r) = \int d^4x \left[\frac{1}{2} (Z - 1) (\nabla \phi_r(x))^2 + \frac{1}{2} Z \delta r \phi_r^2(x) + \frac{1}{4!} g_r (Z_g - 1) \phi_r^4(x) \right], \quad (9.11)$$

and δr and Z_g are two additional renormalization constants. The identity between the renormalized action (9.9) and the regularized action (9.3) is expressed by relation (9.8) and the relations between renormalized and bare quantities:

$$g = g_r Z_g / Z^2, \quad r = m_r^2 / Z + \delta r, \quad (9.12)$$

where Z is the field renormalization constant, Z_g/Z^2 is the coupling constant renormalization constant and δr characterizes the mass renormalization.

Perturbative renormalization theorem. In this chapter, we derive (non-rigorously) the following theorem: the renormalization constants δr , Z_g , and Z , calculated as power series in g_r of the form,

$$\begin{cases} \delta r = \Lambda^2 [a_1(\Lambda) g_r + a_2(\Lambda) g_r^2 + O(g_r^3)] \\ Z_g = 1 + b_1(\Lambda) g_r + b_2(\Lambda) g_r^2 + O(g_r^3) \\ Z = 1 + c_1(\Lambda) g_r + c_2(\Lambda) g_r^2 + O(g_r^3), \end{cases} \quad (9.13)$$

can be chosen as functions of g_r , in such a way that the renormalized vertex functions, order by order in g_r , have a finite limit for $\Lambda \rightarrow \infty$ (a mathematical limit that is not necessarily physical).

Perturbative and loop expansions. The action (9.1) can be rewritten as

$$\mathcal{S}(\phi) = \frac{1}{g} \int d^4x \left[\frac{1}{2} (\sqrt{g} \nabla \phi(x))^2 + \frac{1}{2} r (\sqrt{g} \phi(x))^2 + \frac{1}{4!} (\sqrt{g} \phi(x))^4 \right]. \quad (9.14)$$

The loop expansion in the ϕ^4 QFT is thus an expansion in powers of g at $\phi \sqrt{g}$ fixed.

Because the renormalization constants are series in g_r , the expansion in powers of g_r at $(\sqrt{g_r} \phi_r)$ fixed is no longer a loop expansion in the diagrammatic sense, in contrast with the expansion in powers of g (equation (9.14)). At order g_r^{L-1} , contribute L loop diagrams and diagrams with less than L loops multiplied by contributions coming from renormalization constants. Nevertheless, in the next section, when no confusion is possible, we call the expansion in powers of g_r at $(\sqrt{g_r} \phi_r)$ fixed, loop expansion.

9.2.2 Bare and renormalized correlation functions

The relation between bare and renormalized fields $\phi = Z^{1/2} \phi_r$ directly implies the relations between connected renormalized and bare correlation functions,

$$W_r^{(n)}(x_1, \dots, x_n) = Z^{-n/2} W^{(n)}(x_1, \dots, x_n),$$

which can be summarized by

$$\mathcal{W}(J/\sqrt{Z}) = \mathcal{W}_r(J), \quad (9.15)$$

in which $\mathcal{W}(J)$ and $\mathcal{W}_r(J)$ are, respectively, the generating functionals of bare and renormalized connected correlation functions.

After Legendre transformation, one verifies that the corresponding generating functionals of vertex functions are related by (J and φ are dual)

$$\Gamma_r(\varphi) = \Gamma(\varphi\sqrt{Z}), \quad (9.16)$$

a relation that, for the renormalized and bare vertex functions, translates into

$$\Gamma_r^{(n)}(x_1, \dots, x_n) = Z^{n/2} \Gamma^{(n)}(x_1, \dots, x_n). \quad (9.17)$$

Operator ϕ^2 insertions. We also need the bare and renormalized ϕ^2 insertions. Therefore, we introduce a source $J_2(x)$ for $\phi^2(x)$ and add to the bare action (9.1) the corresponding source term:

$$\mathcal{S}(\phi, J_2) = \mathcal{S}(\phi) + \frac{1}{2} \int d^4x J_2(x) \phi^2(x). \quad (9.18)$$

Functional differentiations with respect to $J_2(x)$ of the field integral

$$\mathcal{Z}(J, J_2) = \int [d\phi] \exp \left[-\mathcal{S}(\phi, J_2) + \int d^4x J(x) \phi(x) \right], \quad (9.19)$$

then generate insertions of the operator $-\frac{1}{2}\phi^2(x)$ (Section 8.6). Similarly, the field integral

$$\mathcal{Z}_r(J, J_2) = \int [d\phi_r] \exp \left[-\mathcal{S}_r(\phi_r, J_2) + \int d^4x J(x) \phi_r(x) \right], \quad (9.20)$$

where

$$\mathcal{S}_r(\phi_r, J_2) = \mathcal{S}_r(\phi_r) + \frac{1}{2} Z_2 \int J_2(x) \phi_r^2(x) d^4x, \quad (9.21)$$

in which Z_2 is a new renormalization constant, generates the renormalized correlation functions with $-\frac{1}{2}\phi^2$ insertions (we temporarily normalize $\mathcal{Z}(J, J_2)$ to $\mathcal{Z}(0, J_2) = 1$ in order to eliminate the pure $\phi^2(x)$ correlation functions).

The relation between renormalized and bare functionals is then

$$\mathcal{W}_r(J, J_2) = \mathcal{W}(J/\sqrt{Z}, J_2 Z_2/Z). \quad (9.22)$$

After Legendre transformation, the relation implies

$$\Gamma_r(\varphi, J_2) = \Gamma(\varphi\sqrt{Z}, J_2 Z_2/Z), \quad (9.23)$$

or, in terms of vertex functions,

$$\Gamma_r^{(l,n)}(y_1, \dots, y_l; x_1, \dots, x_n) = Z^{(n/2)-l} Z_2^l \Gamma^{(l,n)}(y_1, \dots, y_l; x_1, \dots, x_n). \quad (9.24)$$

If the source J_2 is a constant, then it simply generates a shift of the bare parameter r , and the action (9.18) can be rewritten as

$$\mathcal{S}(\phi, J_2) = \int d^4x \left[\frac{1}{2} (\nabla\phi(x))^2 + \frac{1}{2} (r + J_2) \phi^2(x) + \frac{1}{4!} g \phi^4(x) \right]. \quad (9.25)$$

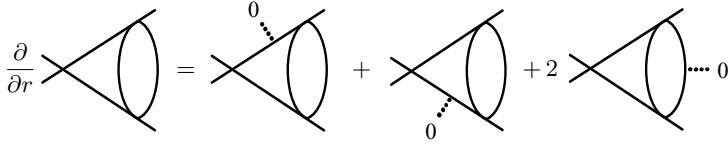


Fig. 9.6 A contribution to $\partial\Gamma^{(4)}/\partial r$ and the corresponding contributions to $\Gamma^{(1,4)}$

In this limit, differentiation with respect to J_2 is equivalent to differentiation with respect to r . Moreover, if $J_2(x)$ is a constant, its Fourier transform is proportional to $\delta^{(d)}(p)$, which means that it generates vertex functions with insertions of the Fourier transform of $\phi^2(x)$ at zero momentum,

$$\frac{\partial}{\partial r} \Big|_g \tilde{\Gamma}^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n) = \tilde{\Gamma}^{(l+1,n)}(0, q_1, \dots, q_l; p_1, \dots, p_n). \quad (9.26)$$

Equation (9.26) has a diagrammatic interpretation: the diagrams contributing to the right-hand side are obtained from the diagrams contributing to $\Gamma^{(l,n)}$ by doubling a propagator in all possible ways (up to a sign). In Fig. 9.6, we illustrate the relation between $\Gamma^{(4)}$ and $\Gamma^{(1,4)}$ with a two-loop order diagram.

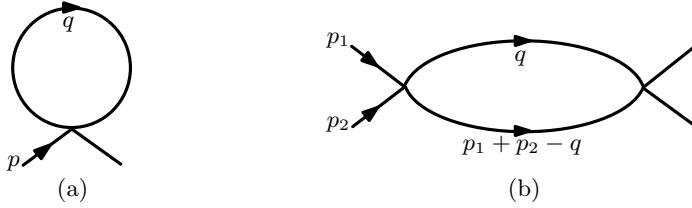


Fig. 9.7 One-loop divergent diagrams in the ϕ^4 QFT

9.3 One-loop divergences

In Section 8.5, we have calculated the one-loop divergences in the ϕ_6^3 (the subscript refers to the space dimension) QFT. We examine here the one-loop divergences of the ϕ_4^4 QFT. We expand the generating functional of renormalized vertex functions $\Gamma_r(\varphi)$ at one-loop order (really in powers of g_r at $\varphi\sqrt{g_r}$ fixed).

At tree order, the counter-terms do not contribute and thus (equation (9.10)),

$$\begin{aligned} \Gamma_r(\varphi) &= \Gamma_{r,0}(\varphi) = \lim_{\Lambda \rightarrow \infty} S_{r,0}(\varphi) \\ &= \int d^4x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} m_r^2 \varphi^2(x) + \frac{1}{4!} g_r \varphi^4(x) \right]. \end{aligned} \quad (9.27)$$

It follows that,

$$\tilde{\Gamma}_r^{(2)}(p) = p^2 + m_r^2, \quad \tilde{\Gamma}_r^{(4)}(p_1, \dots, p_4) = g_r, \quad \tilde{\Gamma}_r^{(n)} = 0, \text{ for } n > 4. \quad (9.28)$$

At one-loop order contribute the one-loop terms generated by the tree order action and the counter-terms at leading order. The former is given by equation (7.93):

$$\begin{aligned} \Gamma_1(\varphi) &= \frac{1}{2} \text{tr} \ln \left[1 + (m_r^2 - \nabla_\Lambda^2)^{-1} g_r \varphi^2 / 2 \right] \\ &= \frac{1}{4} \text{tr} (m_r^2 - \nabla_\Lambda^2)^{-1} g_r \varphi^2 - \frac{1}{16} \text{tr} \left[(m_r^2 - \nabla_\Lambda^2)^{-1} g_r \varphi^2 \right]^2 + O(\varphi^6). \end{aligned}$$

The first two terms in the expansion in powers of φ^2 correspond to the two divergent diagrams displayed in Fig. 9.7.

Using a specific regularization of the form (9.4) with $s = 2$, $M_1 = M_2 = \Lambda$ (to simplify the propagator, to m_r^2 have been added terms of order m_r^4/Λ^2), one finds:

(i) for $n = 2$, the coefficient (a) of $\frac{1}{2}g_r\varphi^2$,

$$\frac{1}{2}(\text{a}) = \frac{1}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 + m_r^2)(1 + p^2/\Lambda^2)^2} = \frac{1}{16\pi^2} \left(\frac{\Lambda^2}{2} - m_r^2 \ln \frac{\Lambda}{m_r} \right) + O(\Lambda^0);$$

(ii) for $n = 4$, the coefficient (b) of $g_r^2\varphi^4/4!$,

$$-\frac{3}{2}(\text{b}) = -\frac{3}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 + m_r^2)_\Lambda [(q - p_1 - p_2)^2 + m_r^2]_\Lambda} \sim -\frac{3}{16\pi^2} \ln \frac{\Lambda}{m_r}.$$

The divergent part $\Gamma_1^{\text{div.}}(\varphi)$ of $\Gamma_1(\varphi)$ generated by the tree action (9.10) is thus

$$\Gamma_1^{\text{div.}} = \frac{1}{16\pi^2} \int d^4x \left[\frac{1}{2} \left(\frac{\Lambda^2}{2} - m_r^2 \ln \frac{\Lambda}{m_r} \right) g_r \varphi^2(x) - \frac{3}{4!} \ln \frac{\Lambda}{m_r} g_r^2 \varphi^4(x) \right]. \quad (9.29)$$

Note the absence of a term proportional to $\int d^4x (\nabla \varphi_r)^2$ in the ϕ^4 QFT at one-loop order, as a consequence at one-loop order of the $\phi \rightarrow -\phi$ symmetry of the action.

We now consider the modified action

$$\mathcal{S}_{r,1}(\phi_r) = \mathcal{S}_{r,0}(\phi_r) - \Gamma_1^{\text{div.}}(\phi_r).$$

At this order, $\Gamma_1^{\text{div.}}(\phi_r)$ contributes additively to $\Gamma_r(\varphi)$. Its addition cancels the divergences in the $\Lambda \rightarrow \infty$ limit, and the ϕ^4 QFT is renormalized at one-loop order,

$$\Gamma_r(\varphi) = \Gamma_{r,0}(\varphi) + \Gamma_{r,1}(\varphi), \quad \text{with} \quad \Gamma_{r,1}(\varphi) \underset{\Lambda \rightarrow \infty}{=} \Gamma_1(\varphi) - \Gamma_1^{\text{div.}}(\varphi).$$

Identifying $\mathcal{S}_{r,1}(\phi_r)$ with the action (9.9), we infer the divergent part of the counter-terms expanded at one-loop order and, in the parametrization (9.13), a_1 , b_1 , and c_1 .

The condition of finiteness of vertex functions determines these coefficients only up to arbitrary finite constants. The difference between the vertex functions corresponding to two different choices is of the form of the tree order functions, that is, a constant for the four-point function in Fourier space and a first degree polynomial in p^2 , p being the momentum, for the two-point function. In this chapter, it is convenient to impose a set of *renormalization conditions* to the renormalized vertex functions $\tilde{\Gamma}_r^{(n)}$. In the Fourier representation, we impose

$$\begin{aligned} \tilde{\Gamma}_r^{(2)}(p=0) &= m_r^2, \\ \frac{\partial}{\partial p^2} \tilde{\Gamma}_r^{(2)}(p)|_{p=0} &= 1, \\ \tilde{\Gamma}_r^{(4)}(0,0,0,0) &= g_r, \end{aligned} \quad (9.30)$$

conditions consistent with the tree approximation (9.28). They completely determine the three renormalization constants. At one-loop order, one obtains

$$\begin{cases} a_1(\Lambda) = -\frac{1}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 + m_r^2)_\Lambda} = -\frac{1}{16\pi^2} \left(\frac{1}{2}\Lambda^2 - m_r^2 \ln(\Lambda/m_r) - \frac{1}{2}m_r^2 \right), \\ b_1(\Lambda) = \frac{3}{2} \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 + m_r^2)_\Lambda^2} = \frac{1}{16\pi^2} \left(3 \ln(\Lambda/m_r) - \frac{17}{4} \right), \\ c_1(\Lambda) = 0. \end{cases} \quad (9.31)$$

Operator ϕ^2 insertions. One obtains the one-loop contribution to the superficially divergent $\langle\phi^2\phi\phi\rangle$ vertex (1PI) function by adding $J_2(x)$ to $\frac{1}{2}g_r\varphi_r^2$ in the tr ln , and identifying the $J_2\varphi^2$ term:

$$\tilde{\Gamma}_r^{(1,2)}(q; p_1, p_2) = 1 - \frac{1}{2} \frac{g_r}{(2\pi)^4} \int \frac{d^4 k}{(k^2 + m_r^2)_\Lambda [(k+q)^2 + m_r^2]_\Lambda} + O(g_r^2). \quad (9.32)$$

The additional renormalization condition,

$$\tilde{\Gamma}_r^{(1,2)}(q=0; p_1=p_2=0) = 1, \quad (9.33)$$

which again is consistent with the tree approximation, determines the new renormalization constant Z_2 (equation (9.24)). At one-loop order,

$$Z_2 - 1 = \frac{1}{2} \int \frac{d^4 q}{(2\pi)^4} \frac{1}{(q^2 + m_r^2)_\Lambda^2} g_r + O(g_r^2) = \frac{1}{16\pi^2} (\ln(\Lambda/m_r) - \frac{17}{12}) g_r + O(g_r^2). \quad (9.34)$$

The fine-tuning problem. In the conventional interpretation of the renormalization procedure, one adjusts the parameters of the initial (or bare) action as functions of the cut-off and physical parameters, and then takes the infinite cut-off limit. This tuning procedure is quite strange, since the bare parameters are singular in this limit and become meaningless. Although we will follow this formal procedure, we will give it a different interpretation. First, we will implicitly assume that the cut-off is large but not infinite (we simply neglect the contributions that formally vanish for infinite cut-off in the perturbative expansion). Second, for the fine tuning of the ϕ^2 coefficient, we have a physics interpretation in the framework of statistical physics: this parameter plays the role of the temperature and we know that the correlation length can become large only near the critical temperature. However, for particle physics, the origin of such tuning, which demands fixing values of parameters with unreasonable precision, is unknown. Finally, the tuning of the coupling constant renormalization is examined in Section 9.11. Its interpretation involves ‘bare’ (and perturbative) RG equations.

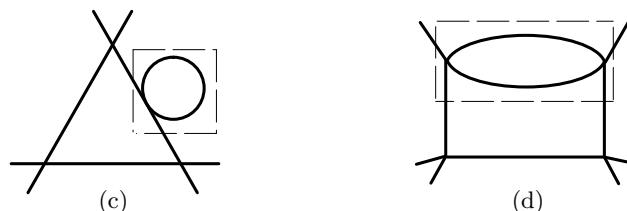


Fig. 9.8 The six-point function: divergent subdiagrams



Fig. 9.9 Two-loop contributions from one-loop counter-terms

9.4 Divergences beyond one-loop: Skeleton diagrams

Power counting shows that, to all orders in the loop expansion, the two-point function has superficial quadratic divergences and the four-point function logarithmic divergences. However, at higher orders a new difficulty arises: superficially convergent diagrams may have divergent subdiagrams. For example, the six-point function at one-loop order is finite but, at two-loop order, the diagrams displayed in Fig. 9.8 contribute.

Inside the dashed boxes, one recognizes divergent subdiagrams. However, they can be identified with one-loop divergences of the two-point function (c) and the four-point function (d), for which counter-terms have already been provided. Indeed, at this order a diagram (c') appears, in which the one-loop counter-term for the two-point function is inserted on a propagator, and another one (d'), in which the vertex of the tree order action is replaced by the one-loop counter-term of the four-point function (Fig. 9.9).

This property is generally true: counter-terms that render the divergent functions finite, at higher orders also cancel the divergence of subdiagrams of superficially convergent functions.

The proof of this property is based on the introduction of *skeleton diagrams*: a skeleton diagram is a superficially convergent without divergent subdiagrams.

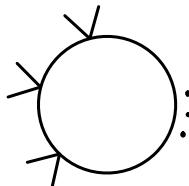


Fig. 9.10 One-loop skeleton diagrams

For example, the one-loop diagrams of the $2n$ -point functions, for $n > 2$, are all skeleton diagrams (see Fig. 9.10).

An arbitrary superficially convergent diagram can then be obtained from a skeleton diagram by replacing all vertices by $\Gamma^{(4)}$ and all propagators by $(\Gamma^{(2)})^{-1}$ and expanding in powers of the coupling constant g_r .

For example, the diagrams (c) and (d) are generated by the expansion of the *dressed skeleton diagram* of Fig. 9.11.

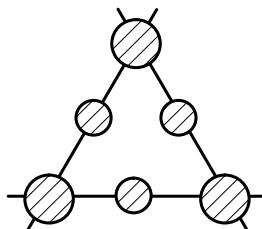


Fig. 9.11 Dressed skeleton diagram

An important property is the following: if in a dressed skeleton diagram, $\Gamma^{(4)}$ and $\Gamma^{(2)}$ are replaced by the renormalized functions $\Gamma_r^{(4)}$ and $\Gamma_r^{(2)}$, the dressed skeleton diagram is finite.

This is a direct consequence of the bounds on the large momentum behaviour [65]:

$$\left| \tilde{\Gamma}_r^{(2)}(\lambda p) \right| \leq \lambda^2 \times \text{power of } \ln \lambda, \quad \left| \tilde{\Gamma}_r^{(4)}(\lambda p_i) \right| \leq \text{power of } \ln \lambda, \quad \left| \tilde{\Gamma}_r^{(1,2)}(\lambda q; \lambda p_1, \lambda p_2) \right| \leq \text{power of } \ln \lambda, \quad \left. \right\} \text{at any finite order for } \lambda \rightarrow \infty. \quad (9.35)$$

A few comments about them can be found at the end of Section 9.6. These bounds for the large momentum behaviour of the various renormalized functions, which are valid for arbitrary momenta, differ from the tree order behaviour only by powers of logarithms (at any finite order in the loop expansion). Therefore, power counting arguments still apply and superficially convergent diagrams are thus convergent.

Similar estimates exist for superficially convergent functions, but are then valid only for generic momenta (see Sections 11.3 and 11.4).

The bounds (9.35) together with the skeleton expansion completely reduce the problem of renormalization of superficially convergent vertex functions to the renormalization of the divergent vertex functions.

The argument also applies to the vertex functions $\Gamma^{(l,n)}$ with l ϕ^2 insertion. For example, the diagram of Fig. 9.12, which contributes to the superficially convergent function $\Gamma^{(1,4)}$, has divergent subdiagrams and is generated from a skeleton diagram by replacing propagators by $(\Gamma^{(2)})^{-1}$, vertices by $\Gamma^{(4)}$, and the $\phi^2\phi\phi$ vertex by $\Gamma^{(1,2)}$, as shown in Fig. 9.13.

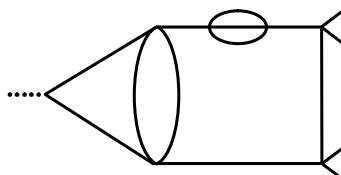


Fig. 9.12 Divergent contribution to $\Gamma^{(1,4)}$

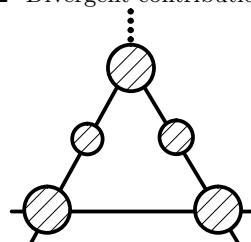


Fig. 9.13 Dressed skeleton diagram with ϕ^2 insertion

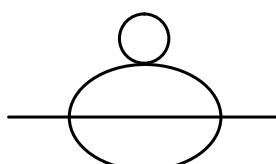


Fig. 9.14 Three-loop contribution to the two-point function

Superficially divergent functions. In the next section, we thus examine the renormalization of $\Gamma^{(2)}$, $\Gamma^{(4)}$ and $\Gamma^{(1,2)}$. The diagrams contributing to these functions are superficially divergent, but also have divergent subdiagrams corresponding to the divergence of the same functions at lower orders. Fig. 9.14 provides an example.

Overlapping divergences. However, a new problem arises, the problem of *overlapping divergences*. For example, let us examine the diagram of Fig. 9.15.

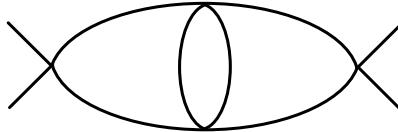


Fig. 9.15 Three-loop contribution to the four-point function

Figure 9.16 displays the three divergent subdiagrams. These subdiagrams have a common part. Therefore, the concept of insertion of divergent diagrams of lower order is no longer well-defined. This is the problem of the so-called *overlapping divergences*.

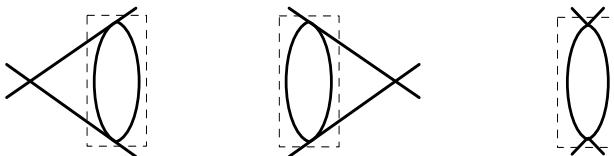


Fig. 9.16 Overlapping divergent subdiagrams in the diagram 9.15

In Section 9.5, we develop a specific technique, based on differentiating diagrams with respect to the mass, to deal with this problem.

9.5 Callan–Symanzik equations

The starting point of the analysis is equation (9.26), which shows that a differentiation with respect to the bare mass improves the large momentum behaviour of Feynman diagrams. This provides a method to relate superficially divergent vertex functions to functions which have a skeleton expansion. Furthermore, at a given number of loops, a function which has a skeleton expansion is only expressed in terms of divergent functions which have at least *one loop less*. We see here a mechanism to prove renormalizability by induction. However, we want to insert into the skeleton expansion renormalized vertex functions. This introduces some additional difficulties that will become apparent once we have transformed equation (9.26) into an equation for the renormalized vertex functions.

We first introduce a notation and apply the chain rule to transform differentiation with respect to m_r at g, Λ fixed, into differentiation at g_r, Λ fixed:

$$D_r \equiv m_r \frac{\partial}{\partial m_r} \Big|_{g, \Lambda} = m_r \frac{\partial}{\partial m_r} \Big|_{g_r, \Lambda} + \beta(g_r, m_r/\Lambda) \frac{\partial}{\partial g_r} \Big|_{m_r, \Lambda} \quad (9.36)$$

with the definition (we take dimensional analysis immediately into account),

$$D_r g_r = \beta(g_r, m_r/\Lambda). \quad (9.37)$$

Similarly, we define the functions η , η_2 and σ as

$$D_r \ln Z(g_r, m_r/\Lambda) = \eta(g_r, m_r/\Lambda), \quad (9.38)$$

$$D_r \ln (Z_2/Z) = \eta_2(g_r, m_r/\Lambda), \quad (9.39)$$

$$ZZ_2^{-1} (D_r m^2) = m_r^2 \sigma(g_r, m_r/\Lambda). \quad (9.40)$$

In terms of the differential operator

$$D_{CS} = D_r - \frac{1}{2} n \eta(g_r, m_r/\Lambda) - l \eta_2(g_r, m_r/\Lambda),$$

one finds the equation,

$$\begin{aligned} D_{CS} \tilde{\Gamma}_r^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n) \\ = m_r^2 \sigma(g_r, m_r/\Lambda) \tilde{\Gamma}_r^{(l+1,n)}(0, q_1, \dots, q_l; p_1, \dots, p_n). \end{aligned} \quad (9.41)$$

In the infinite cut-off limit, equation (9.41) becomes an equation for the renormalized vertex functions called Callan–Symanzik (CS) equation [60], which has also played an important role in the calculations of universal quantities in continuous phase transitions (Chapters 15–32).

To prove renormalizability, we prove inductively on the number of loops both the existence of the CS equation and the finiteness of vertex functions.

Renormalization conditions. The CS equation in the form (9.41) expresses only that we have rescaled the vertex functions and made an arbitrary change of parametrization. To be able to prove that the renormalized vertex functions have a finite $\Lambda \rightarrow \infty$ limit, it is necessary to determine the renormalization constants and, therefore, to impose on equation (9.41) the consequences of the renormalization conditions (9.30) and (9.33).

(i) $n = 2, l = 0$:

at zero momentum, equation (9.41) yields

$$\left(m_r \frac{\partial}{\partial m_r} - \eta(g_r, \Lambda/m_r) \right) m_r^2 = m_r^2 \sigma(g_r, \Lambda/m_r),$$

and thus, the relation

$$\sigma = 2 - \eta. \quad (9.42)$$

If one then differentiates the same equation with respect to a momentum squared, one finds

$$-\eta = m_r^2 (2 - \eta) \frac{\partial}{\partial p^2} \tilde{\Gamma}_r^{(1,2)}(0; p, -p) \Big|_{p^2=0}. \quad (9.43)$$

(ii) $n = 4, l = 0$:

at zero momentum, equation (9.41) yields

$$\beta - 2g_r \eta = m_r^2 (2 - \eta) \tilde{\Gamma}_r^{(1,4)}(0; 0, 0, 0, 0). \quad (9.44)$$

(iii) $n = 2, l = 1$:

again, the zero momentum limit yields

$$-\eta - \eta_2 = m_r^2 (2 - \eta) \tilde{\Gamma}_r^{(2,2)}(0, 0; 0, 0). \quad (9.45)$$

We have related all the coefficients of the partial differential equation (9.41) to values of vertex functions at zero momentum. From these relations it follows that, if we can show that the renormalized vertex functions have an infinite cut-off limit, the functions β , η , and η_2 will also have a limit.

Also note that, if we know the coefficients of the CS equations, we can calculate the renormalization constants from the set of equations (9.37–9.40).

Leading order contributions.

(i) $\tilde{\Gamma}_r^{(1,2)}(0; p, -p)$ at order g does not depend on p (equation (9.32)). The equation (9.34) then implies

$$\tilde{\Gamma}_r^{(1,2)}(0; p, -p) = 1 + O(g_r^2).$$

Therefore, the expansion of $\eta(g_r)$, which can be calculated from equation (9.43), begins at order g_r^2 .

(ii) The first diagram contributing to $\tilde{\Gamma}^{(1,4)}$ is of order g_r^2 (see Fig. 9.17 (a)). It then follows from equation (9.44) and the preceding remark that the function $\beta(g_r)$ has also an expansion that begins at order g_r^2 . Thus, the operator $\beta_2 \partial/\partial g_r$, which appears in the CS equation is of order g_r .

(iii) The function $\tilde{\Gamma}^{(2,2)}$ has a first contribution of order g_r (see Fig. 9.17 (b)). Equation (9.45) then shows that η_2 also begins at order g_r .

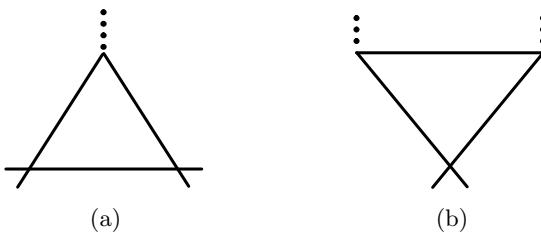


Fig. 9.17 Contributions (a) to $\tilde{\Gamma}^{(1,4)}$, and (b) to $\tilde{\Gamma}^{(2,2)}$

Cluster properties and analyticity at low momentum. In Section 8.5, we have used the regularity of the one-loop diagrams near zero momentum to show that the divergent contributions are polynomials in the momentum variables. This is more generally true: in a massive theory, connected and vertex (or 1PI) functions are analytic functions around zero momentum, as can be seen on the expression of regularized Feynman diagrams. This property, which will again be needed in the inductive proof, implies cluster properties: connected correlation functions decrease exponentially in space for large separations of the arguments (for details, see Section A7.3).

9.6 Inductive proof of renormalizability

In Section 9.3, we have constructed a theory finite at one-loop order. We now assume that the vertex functions defined by equations (9.24) and renormalization conditions (9.30, 9.33) have a finite limit up to loop order L , at m_r and g_r fixed, for $\Lambda \rightarrow \infty$.

This means that $\tilde{\Gamma}_r^{(2)}, \tilde{\Gamma}_r^{(4)}, \tilde{\Gamma}_r^{(1,2)}$ have a limit for $\Lambda \rightarrow \infty$ up to order g_r^L, g_r^{L+1} and g_r^L , respectively.

As we have shown in Section 9.5, from equations (9.24), (9.30), and (9.33) follow the CS equations (9.41) and the relations (9.42–9.45).

We now use the CS equation (9.41) in the form

$$m_r \frac{\partial}{\partial m_r} \tilde{\Gamma}_r^{(l,n)} = \left(-\beta_2 \frac{\partial}{\partial g_r} + \frac{n}{2} \eta + l \eta_2 \right) \tilde{\Gamma}_r^{(l,n)} + m_r^2 (2 - \eta) \tilde{\Gamma}_r^{(l+1,n)}, \quad (9.46)$$

and show that the right-hand side is finite at loop order $(L + 1)$.

We note that $\Gamma^{(l,n)}$ in the right-hand side is only needed at loop order L , because its coefficient is of order g_r . For $\Gamma^{(l+1,n)}$ two cases arise: either it has a skeleton expansion and is, therefore, finite at loop order $(L + 1)$, or the CS equation has to be iterated. However, before discussing vertex functions, we examine the coefficient functions.

9.6.1 Coefficients of the CS equation

- (i) Because $\partial \tilde{\Gamma}^{(1,2)} / \partial p^2$ is of order g_r^2 , equation (9.43) then implies that η is finite up to order g_r^L .
- (ii) The function $\tilde{\Gamma}^{(1,4)}$ is superficially convergent. Therefore, it has a skeleton expansion. The first dressed skeleton diagram contributing to $\tilde{\Gamma}^{(1,4)}$ is shown in Fig. 9.18.

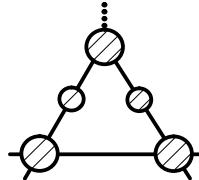


Fig. 9.18 Skeleton diagram contributing to $\Gamma^{(1,4)}$

If the functions $\tilde{\Gamma}_r^{(2)}$, $\tilde{\Gamma}_r^{(4)}$ and $\tilde{\Gamma}_r^{(1,2)}$ are finite up to L loops, $\tilde{\Gamma}_r^{(1,4)}$ is finite up to loop order $(L + 1)$, which means up to order g_r^{L+2} . Equation (9.44) then shows that the combination $\beta - 2g_r\eta$ is finite up to order g_r^{L+2} . Since η is finite up to order g_r^L , this implies that β is finite up to order g_r^{L+1} .

- (iii) The function $\tilde{\Gamma}^{(2,2)}$ is also superficially convergent. It has a skeleton expansion. The first dressed skeleton is shown in Fig. 9.19.

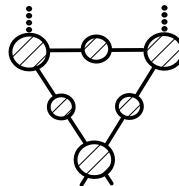


Fig. 9.19 Skeleton diagram contributing to $\Gamma^{(2,2)}$

Therefore, the function $\tilde{\Gamma}_r^{(2,2)}$ is also finite up to loop order $(L + 1)$, which means up to order g_r^{L+1} . Since $\tilde{\Gamma}^{(2,2)}$ is of order g_r , at order g_r^{L+1} the sum $\eta + \eta_2$ calculated from equation (9.45) involves only η at order g_r^L and $\tilde{\Gamma}_r^{(2,2)}$ at order g_r^{L+1} , and is thus finite. The function η_2 is then finite up to order g_r^L .

We now prove that the functions $\tilde{\Gamma}_r^{(2)}$, $\tilde{\Gamma}_r^{(4)}$, and $\tilde{\Gamma}_r^{(1,2)}$ are, with the induction assumptions, finite up to loop order $(L + 1)$.

9.6.2 The $\langle\phi\phi\phi\rangle$ vertex function ($l = 0, n = 4$)

We consider the coefficient of order g_r^{L+2} in equation (9.46):

$$\left[m_r \frac{\partial}{\partial m_r} \tilde{\Gamma}_r^{(4)} \right]_{L+2} = \left[\left(-\beta_2 \frac{\partial}{\partial g_r} + 2\eta \right) \tilde{\Gamma}_r^{(4)} \right]_{L+2} + m_r^2 \left[(2-\eta) \tilde{\Gamma}_r^{(1,4)} \right]_{L+2}.$$

Since $\beta_2 \partial / \partial g_r$ is of order g_r and η of order g_r^2 for g_r small, in the first term of the right-hand side we need $\tilde{\Gamma}_r^{(4)}$ only up to order g_r^{L+1} , which is finite by assumption. We now separate in $\tilde{\Gamma}_r^{(4)}$ the leading term g_r and a remainder of order g_r^2 . For the terms of order g_r^2 and higher, we need η only up to order g_r^L and β up to order g_r^{L+1} , which are finite. The leading term in $\tilde{\Gamma}_r^{(4)}$ then involves the combination

$$\left[\left(-\beta_2 \frac{\partial}{\partial g_r} + 2\eta \right) g_r \right]_{L+2} = [-\beta_2 + 2\eta g_r]_{L+2},$$

and we have shown above that $\beta - 2g_r\eta$ is finite up to order g_r^{L+2} .

Finally, $\tilde{\Gamma}_r^{(1,4)}$ is finite up to order g_r^{L+2} . In addition, its expansion in powers of g_r begins only at order g_r^2 . Therefore, the factor $(2-\eta)$ is only needed up to order g_r^L . The conclusion is that the left-hand side is finite at loop order $(L+1)$.

Perturbative integration of CS equations. In what follows, we denote by \mathbf{p} the set of all four momenta (p_1, p_2, p_3, p_4) .

By integrating equation (9.46), we now want to show that $\tilde{\Gamma}_r^{(4)}$ itself is finite at $(L+1)$ loop order. The function $\tilde{\Gamma}_r^{(4)}$ is dimensionless, that is, invariant in a dilatation

$$(\mathbf{p}, m_r, \Lambda) \mapsto (\rho \mathbf{p}, \rho m_r, \rho \Lambda).$$

Therefore, we set

$$m_r \frac{\partial}{\partial m_r} \tilde{\Gamma}_r^{(4)} = f^{(4)}(\mathbf{p}/m_r, m_r/\Lambda, g_r), \quad (9.47)$$

with

$$\lim_{\Lambda \rightarrow \infty} f^{(4)}(\mathbf{p}/m_r, m_r/\Lambda, g_r) < \infty, \text{ and } f^{(4)}(0, m_r/\Lambda, g_r) = m_r \frac{\partial}{\partial m_r} g_r = 0.$$

We integrate equation (9.47) between m_r and Λ :

$$\tilde{\Gamma}_r^{(4)}(\mathbf{p}/m_r, m_r/\Lambda, g_r) = \tilde{\Gamma}_r^{(4)}(\mathbf{p}/\Lambda, 1, g_r) - \int_{m_r}^{\Lambda} \frac{d\rho}{\rho} f^{(4)}(\mathbf{p}/\rho, \rho/\Lambda, g_r).$$

Renormalization conditions (9.30) together with the regularity at low momentum imply

$$\lim_{\Lambda \rightarrow \infty} \tilde{\Gamma}_r^{(4)}(\mathbf{p}/\Lambda, 1, g_r) = \tilde{\Gamma}_r^{(4)}(0, 1, g_r) = g_r.$$

The integral has a large cut-off limit if we restrict the domain of integration to $[m_r, \mu]$, $m_r \ll \mu \ll \Lambda$:

$$\lim_{\Lambda \rightarrow \infty} \int_{m_r}^{\mu} \frac{d\rho}{\rho} f^{(4)}(\mathbf{p}/\rho, \rho/\Lambda, g_r) < \infty.$$

We still have to examine the values of ρ of order Λ . Then Λ/ρ is of order 1, while \mathbf{p}/ρ is small. The integral depends on the small momentum behaviour of the four-point vertex function. Again, we use the property that, in a massive theory (the mass is $\rho \sim \Lambda$), the vertex functions are analytic around $\mathbf{p} = 0$. Since, in addition, $f^{(4)}$ vanishes at $\mathbf{p} = 0$,

$$\left| f^{(4)}(\mathbf{p}/\rho, \rho/\Lambda, g_r) \right| < \frac{C(\mathbf{p}, g_r)}{\rho^2}, \quad \text{for } \rho \gg |\mathbf{p}|,$$

the remaining integral is then bounded by

$$\int_\mu^\Lambda \frac{d\rho}{\rho} \left| f^{(4)}(\mathbf{p}/\rho, \rho/\Lambda, g_r) \right| < \frac{1}{2} C(\mathbf{p}, g_r) \left(\frac{1}{\mu^2} - \frac{1}{\Lambda^2} \right).$$

We conclude that $\tilde{\Gamma}_r^{(4)}$ has a large cut-off limit at $(L+1)$ loop order given by

$$\tilde{\Gamma}_r^{(4)}(\mathbf{p}, m_r, g_r) = g_r - \int_1^\infty \frac{d\rho}{\rho} f^{(4)}(\mathbf{p}/\rho m_r, 0, g_r), \quad (9.48)$$

both sides being expanded up to order g_r^{L+2} .

9.6.3 The $\langle \phi^2 \phi \phi \rangle$ vertex function ($l=1, n=2$)

We now repeat the argument for $\tilde{\Gamma}_r^{(1,2)}$. We consider the term of order g_r^{L+1} in equation (9.46),

$$m_r \frac{\partial}{\partial m_r} \left[\tilde{\Gamma}_r^{(1,2)} \right]_{L+1} = - \left[\beta \frac{\partial}{\partial g_r} \tilde{\Gamma}_r^{(1,2)} \right]_{L+1} + \left[(\eta + \eta_2) \tilde{\Gamma}_r^{(1,2)} \right]_{L+1} + m_r^2 \left[(2 - \eta) \tilde{\Gamma}_r^{(2,2)} \right]_{L+1}.$$

The first term in the right-hand side involves $\tilde{\Gamma}_r^{(1,2)}$ up to order L , since $\beta \partial / \partial g_r$ is of order g_r , and β up to order $(L+1)$. Both are finite. The second term again involves $\tilde{\Gamma}_r^{(1,2)}$ up to order L , since $(\eta + \eta_2)$ is of order g_r and $(\eta + \eta_2)$ at order $(L+1)$, which is finite (although η and η_2 separately are not). The last term involves $\tilde{\Gamma}_r^{(2,2)}$ up to order $(L+1)$ and η up to order L , since $\tilde{\Gamma}_r^{(2,2)}$ is of order g_r . We conclude that $m_r \partial \tilde{\Gamma}_r^{(1,2)} / \partial m_r$ is finite at $(L+1)$ loop order.

The function $\tilde{\Gamma}_r^{(1,2)}$ is also dimensionless. Its value at zero momentum is fixed by the renormalization condition (9.33), therefore,

$$m_r \frac{\partial}{\partial m_r} \tilde{\Gamma}_r^{(1,2)}(0; 0, 0) = 0.$$

The analysis is then the same as for $\tilde{\Gamma}_r^{(4)}$, and we conclude that $\tilde{\Gamma}_r^{(1,2)}$ has a finite limit at infinite cut-off at loop order $(L+1)$. We can now use equation (9.43) and argument (i): Since $\tilde{\Gamma}^{(1,2)}$ is finite up to order g_r^{L+1} , η is finite up to the same order g_r^{L+1} .

9.6.4 The $\langle\phi\phi\rangle$ vertex function ($l = 0, n = 2$)

The term of order g_r^{L+1} in equation (9.46) is

$$m_r \frac{\partial}{\partial m_r} \left[\tilde{\Gamma}_r^{(2)} \right]_{L+1} = \left[\left(-\beta \frac{\partial}{\partial g_r} + \eta \right) \tilde{\Gamma}_r^{(2)} \right]_{L+1} + m_r^2 \left[(2 - \eta) \tilde{\Gamma}_r^{(1,2)} \right]_{L+1}. \quad (9.49)$$

In the first term of the right-hand side, we need $\tilde{\Gamma}_r^{(2)}$ only up to order g_r^L , since it is multiplied by terms of order g_r . We have also shown previously that η is finite up to order g_r^{L+1} and β is finite at this order by argument (ii). For the second term, we have just shown that the two factors are finite up to order g_r^{L+1} . We now consider the quantity $\tilde{\Gamma}_r^{(2)}(p) - m_r^2 - p^2$. Renormalization conditions imply that it vanishes as $(p^2)^2$ for $|\mathbf{p}|$ small. It has mass dimension 2. We set

$$m_r \frac{\partial}{\partial m_r} \left[\tilde{\Gamma}_r^{(2)}(p) - m_r^2 - p^2 \right] = m_r^2 f^{(2)}(p/m_r, m_r/\Lambda, g_r), \quad (9.50)$$

with

$$\begin{cases} \lim_{\Lambda \rightarrow \infty} f^{(2)}(p/m_r, m_r/\Lambda, g_r) < \infty, \\ f^{(2)}(p/m_r, m_r/\Lambda, g_r) = O(p^4) \text{ for } |p| \rightarrow 0. \end{cases}$$

The integration of equation (9.50) then yields

$$\begin{aligned} \tilde{\Gamma}_r^{(2)}(p, m_r, g_r, \Lambda) - m_r^2 - p^2 &= \Lambda^2 \left[\tilde{\Gamma}_r^{(2)}(p/\Lambda, 1, g_r, 1) - 1 - p^2/\Lambda^2 \right] \\ &\quad - \int_{m_r}^{\Lambda} \rho d\rho f^{(2)}(p/\rho, \rho/\Lambda, g_r). \end{aligned} \quad (9.51)$$

The integrated term in the right-hand side decreases like $(p^2)^2/\Lambda^4$ for Λ large and, therefore, goes to zero.

For the same reason, $f^{(2)}(p/\rho, \rho/\Lambda, g_r)$ is of order $1/\rho^4$, for $\rho \sim \Lambda$, and, therefore, the infinite cut-off limit can be taken in the integral.

This concludes the induction. The advantage of the method is that, simultaneously, we have proved renormalizability and derived the CS equations,

$$\begin{aligned} &\left(m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{\eta}{2} \eta(g_r) - l\eta_2(g_r) \right) \tilde{\Gamma}_r^{(l,n)}(q_j; p_i) \\ &= (2 - \eta(g_r)) m_r^2 \tilde{\Gamma}_r^{(l+1,n)}(0, q_j; p_i). \end{aligned} \quad (9.52)$$

9.6.5 The large momentum behaviour of superficially divergent vertex functions

We infer from this derivation that we can also use induction to estimate the large momentum behaviour of vertex functions. If we assume the bounds (9.35) at loop order L , then $\tilde{\Gamma}^{(1,4)}$ and $\tilde{\Gamma}^{(2,2)}$ are given by convergent integrals at loop order $(L+1)$. It is not too difficult to bound their large momentum behaviour by powers of logarithms.

Let us now again consider the example of the four-point function. Once we have established the representation (9.48), the induction hypothesis tells us that the function $f^{(4)}$ in the right-hand side is bounded by (we omit now the cut-off dependence)

$$|f^{(4)}(\lambda \mathbf{p}/m_r, g_r)| < \text{const. } (\ln \lambda)^{k(L)} \quad \text{for } \lambda \rightarrow \infty.$$

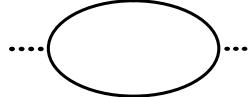


Fig. 9.20 One-loop contribution to $\Gamma^{(2,2)}$

We then decompose the integral over ρ into the sum of two terms,

$$\int_1^\infty \frac{d\rho}{\rho} f^{(4)}(\lambda \mathbf{p}/\rho m_r, g_r) = \int_\lambda^\infty \frac{d\rho}{\rho} f^{(4)}(\lambda \mathbf{p}/\rho m_r, g_r) + \int_1^\lambda \frac{d\rho}{\rho} f^{(4)}(\lambda \mathbf{p}/\rho m_r, g_r).$$

The first integral can be bounded by a constant. The second integral can be bounded using the large-momentum behaviour of $f^{(4)}$,

$$\int_1^\lambda \frac{d\rho}{\rho} \left| f^{(4)}(\lambda \mathbf{p}/\rho m_r, g_r) \right| < \text{const.} \quad \int_1^\lambda \frac{d\rho}{\rho} [\ln(\lambda/\rho)]^{k(L)} \sim (\ln \lambda)^{k(L)+1}.$$

The argument is the same for $\tilde{\Gamma}^{(1,2)}$. This last bound can then be used to bound $\tilde{\Gamma}^{(2)}$.

9.7 The $\langle \phi^2 \phi^2 \rangle$ vertex function

We have seen that the superficial degree of divergence of the vertex function $\tilde{\Gamma}^{(2,0)}$ is 0. Actually, even in free QFT ($g_r = 0$), $\tilde{\Gamma}^{(2,0)}$ is divergent (see Fig. 9.20). An additional renormalization is needed. We impose the renormalization condition

$$\tilde{\Gamma}_r^{(2,0)}(p=0, m_r, g_r) = 0, \quad (9.53)$$

which is obviously consistent with the tree level approximation. We expect the relation between bare and renormalized vertex function to be

$$\tilde{\Gamma}_r^{(2,0)}(p) = (Z_2/Z)^2 \left[\tilde{\Gamma}^{(2,0)}(p) - \tilde{\Gamma}^{(2,0)}(0) \right]. \quad (9.54)$$

Differentiating with respect to m_r at g and Λ fixed, we then obtain

$$\begin{aligned} & \left[m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - 2\eta_2(g_r) \right] \tilde{\Gamma}_r^{(2,0)}(p; m_r, g_r) \\ &= m_r^2 [2 - \eta(g_r)] \left[\tilde{\Gamma}_r^{(3,0)}(p, -p, 0; m_r, g_r) - \tilde{\Gamma}_r^{(3,0)}(0, 0, 0; m_r, g_r) \right]. \end{aligned} \quad (9.55)$$

The derivation of the equation follows the same lines as in previous cases. It uses the properties that $\tilde{\Gamma}^{(3,0)}$ has a skeleton expansion and that the L loop order of $\tilde{\Gamma}^{(2,0)}$ is of order g_r^{L-1} . Finally, in the integration with respect to m_r , one takes into account the renormalization condition (9.53) and notes that $\tilde{\Gamma}^{(2,0)}$ is dimensionless in mass units.

It is possible then to summarize all CS equations by

$$\begin{aligned} & \left[m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{\eta}{2} \eta(g_r) - l\eta_2(g_r) \right] \tilde{\Gamma}_r^{(l,n)}(q_j; p_i; m_r, g_r) \\ &= m_r^2 (2 - \eta(g_r)) \tilde{\Gamma}_r^{(l+1,n)}(0, q_j; p_i; m_r, g_r) + \delta_{n0}\delta_{l2}B(g_r), \end{aligned} \quad (9.56)$$

with the notation

$$-m_r^2 (2 - \eta(g_r)) \tilde{\Gamma}_r^{(3,0)}(0, 0, 0; m_r, g_r) = B(g_r), \quad (9.57)$$

since the left-hand side is dimensionless.

9.8 The renormalized action: General construction

The proof of the renormalizability, *in the sense of power counting*, of arbitrary massive local QFTs is a rather simple generalization of the proof we have given previously for the ϕ^4 QFT in four dimensions. Therefore, we summarize only the main steps of the construction of the renormalized theory.

We consider an arbitrary QFT renormalizable by power counting. We perform a loop expansion of the generating functional of vertex functions,

$$\Gamma(\varphi) = \mathcal{S}_{r,0}(\varphi) + \sum_{l=1}^{\infty} \Gamma_l(\varphi).$$

After addition to the tree order action $\mathcal{S}_{r,0}(\phi)$ of all the counter-terms needed to render the theory finite up to loop-order L , for $\Lambda \rightarrow \infty$, the functionals $\Gamma_l(\varphi)$ have a finite limit for $l \leq L$, and the diagrams contributing to $\Gamma_{L+1}(\varphi)$ have no divergent subdiagrams. Therefore, the divergent part $\Gamma_{L+1}^{\text{div}}(\varphi)$ of $\Gamma_{L+1}(\varphi)$ is a general local functional linear combination of all vertices of non-positive UV dimensions (except if symmetries forbid some terms) [54]. By adding the counter-terms $-\Gamma_{L+1}^{\text{div}}(\varphi)$ to the renormalized action, one renders the theory finite at loop order $(L+1)$. The counter-terms are defined only up to an arbitrary finite part, a linear combination of the same vertices that appear in $\Gamma_{L+1}^{\text{div}}$. *The resulting renormalized action is a general local functional of the fields, a linear combination of all vertices of non-positive UV dimension.*

This statement summarizes the result derived in Section 9.6 in the case of the ϕ^4 QFT in four dimensions. The divergent part of $\Gamma(\varphi)$ at loop order L , after renormalization up to order $(L-1)$, has the form

$$\Gamma_L^{\text{div.}}(\varphi) = - \int d^4x \left[\frac{1}{2} \delta m_L^2 \varphi^2(x) + \frac{1}{2} \delta Z_L (\nabla \varphi(x))^2 + \frac{1}{4!} g_r \delta Z_{g,L} \varphi^4(x) \right], \quad (9.58)$$

because the vertex ϕ^2 has dimension -2 , and the vertices $(\nabla \phi)^2$ and ϕ^4 dimension 0 . No odd power of ϕ appears because the tree order action is symmetric in $\phi \mapsto -\phi$.

Note that, quite generally, if some monomials allowed by power counting and symmetries are absent from the initial (bare) action, they will be generated as counter-terms. Note also that algorithms have been developed to generate directly renormalized Feynman diagrams [66, 67, 68].

9.9 The massless theory

We have derived the renormalizability of the ϕ^4 QFT by a method that applies only to massive QFTs, because the mass insertion operation plays an essential role in decreasing the degree of divergence of Feynman diagrams. To derive a renormalized massless ϕ_4^4 QFT, we rescale the mass, at fixed cut-off and momenta [69],

$$m_r \mapsto m_r / \rho, \quad \rho \rightarrow \infty.$$

At fixed cut-off, vertex functions have a limit, as will be discussed extensively in Chapters 14–17, in dimensions larger than or equal to the dimension in which the theory is exactly renormalizable, provided this dimension is larger than 2 (because the propagator is $1/p^2$). In addition, the set of arguments of the vertex functions in the momentum representation must be non-exceptional, that is, all non-trivial subsets of momenta should have a non-vanishing sum.

These conditions are met in the ϕ_4^4 QFT (at non-exceptional momenta) and we now examine what happens when the cut-off is removed. We combine CS equations with Weinberg's theorem [65].

9.9.1 Large momentum behaviour and massless theory

We start from the CS equations

$$\left[m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \eta(g_r) \right] \tilde{\Gamma}_r^{(n)}(p_i) = m_r^2 (2 - \eta) \tilde{\Gamma}_r^{(1,n)}(0; p_i).$$

Weinberg's theorem states that if we scale all momenta $p_i \mapsto \rho p_i$, the large ρ behaviour at *non-exceptional momenta*, at any finite order of perturbation theory, is given by the mass dimension up to powers of logarithms. Thus,

$$\begin{aligned} \tilde{\Gamma}_r^{(l,n)}(\rho p_i) &\underset{\rho \rightarrow \infty}{\sim} \rho^{4-n-2l} \times \text{power of } \ln \rho, \\ \tilde{\Gamma}_r^{(l+1,n)}(0; \rho p_i) &\underset{\rho \rightarrow \infty}{\sim} \frac{1}{\rho^2} \rho^{4-n-2l} \times \text{power of } \ln \rho. \end{aligned}$$

Therefore, if in the asymptotic expansion of $\tilde{\Gamma}_r^{(n)}(\rho p_i)$ for ρ large, we neglect all terms subleading by a power ρ^{-2} up to powers of $\ln \rho$, we obtain a vertex function $\tilde{\Gamma}_{r,\text{as.}}^{(n)}$ that satisfies the homogeneous RG equation,

$$\left(m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \eta(g_r) \right) \tilde{\Gamma}_{r,\text{as.}}^{(n)}(p_i) = 0. \quad (9.59)$$

However, we know from dimensional analysis that

$$\tilde{\Gamma}_r^{(n)}(p_i, m_r) = m_r^{4-n} \tilde{\Gamma}_r^{(n)}(p_i/m_r, 1). \quad (9.60)$$

Therefore, scaling all momenta by a factor ρ is equivalent, up to a global factor, to scaling the mass by a factor ρ^{-1} . The solutions of equation (9.59) are thus the vertex functions of the massless ϕ^4 QFT.

The critical bare mass. When the physical mass vanishes, the parameter r in expression (9.1) is fixed to a critical value $r = r_c$ and begins with a term of order g_r . The first of equations (9.31) then shows that r is negative for g_r small. This is the sign one expects from the point of view of phase transitions: the existence of a phase transition requires a double-well potential. This justifies our notation r for the coefficient of ϕ^2 in the bare action, instead of the traditional m^2 .

Perturbative solution of the homogeneous CS equations. It is actually interesting to study the structure of vertex functions implied by equation (9.59). For illustration purpose, we consider here only the two-point function $\tilde{\Gamma}_{r,\text{as.}}^{(2)}(p)$. We introduce the function $\zeta(g_r)$ defined by

$$\ln \zeta(g_r) = \int_0^{g_r} \frac{dg' \eta(g')}{\beta(g')}.$$

Since both η and β are of order g_r^2 , the function has a perturbative expansion.

We then set

$$\tilde{\Gamma}_{r,\text{as.}}^{(2)}(p) = p^2 \zeta(g_r) A(g_r, p/m_r).$$

The function A satisfies

$$\left(m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} \right) A(g_r, p/m_r) = 0. \quad (9.61)$$

We expand A and $\beta(g_r)$ in powers of g_r , setting

$$A(g_r, p/m_r) = 1 + \sum_1^\infty g_r^n a_n(p/m_r), \quad \beta(g_r) = \sum_2^\infty \beta_n g_r^n. \quad (9.62)$$

Introducing these expansions into equation (9.61), we find for $n \geq 2$,

$$-za'_n(z) + \sum_{m=1}^{n-1} m a_m(z) \beta_{n-m+1} = 0. \quad (9.63)$$

The equation for $n = 1$ is special (C_1 is a constant),

$$-za'_1(z) = 0 \Rightarrow a_1(z) = C_1.$$

The solution of the equation for $n = 2$ is given by (C_2 is another constant)

$$-za'_2(z) + C_1 \beta_2 = 0 \Rightarrow a_2(z) = C_1 \beta_2 \ln z + C_2.$$

From this example, the general structure of $a_n(z)$ can be guessed: it is a polynomial of degree $(n - 1)$ in $\ln z$ of the form

$$a_n(z) = P_{n-1}(\ln z), \quad \text{with} \quad P'_{n-1}(x) = \sum_{m=1}^{n-1} m P_{m-1}(x) \beta_{n-m+1}. \quad (9.64)$$

The new information specific to the order n is characterized by two constants, β_n , which enters in the coefficient of $\ln z$, and C_n , which is the integration constant (to which one should add the coefficients of $\eta(g_r)$, which appear in the function $\zeta(g_r)$). Moreover, the term of highest degree in P_n is entirely determined by one-loop results, the next term by one- and two-loop results and so on. Finally, $\tilde{\Gamma}_{r,\text{as.}}^{(2)}(p)$ is entirely determined by the functions $\beta(g_r)$ and $\eta(g_r)$ and, for example, $\tilde{\Gamma}_{r,\text{as.}}^{(2)}(1, g_r)/m_r^2$ which is a third function of g_r .

It also follows from these equations that $\tilde{\Gamma}_{r,\text{as.}}^{(2)}(p)$ has a limit for $p = 0$:

$$\tilde{\Gamma}_{r,\text{as.}}^{(2)}(p^2 = 0) = 0, \quad (9.65)$$

confirming, as expected, that the theory is massless. However, its derivative $\partial \tilde{\Gamma}_{r,\text{as.}}^{(2)} / \partial p^2$ has no zero-momentum limit. It is easy to verify that no other vertex function has a zero-momentum limit either.

We have constructed a massless theory by scaling a massive theory and shown that the corresponding vertex functions satisfy a homogeneous CS equation, also called an RG equation. We now show that such an equation can be derived more directly if one assumes the existence of a renormalized massless QFT.

9.9.2 RG equations in a massless QFT

Renormalization conditions. We have shown that the renormalized massless ϕ^4 QFT exists in four dimensions. To determine renormalization constants by renormalization conditions, we have to impose them at non-exceptional momenta; in particular, we cannot use zero momentum except for the two-point function as we have indicated previously. Therefore, we introduce an arbitrary mass scale μ (the physical mass scale) and impose

$$\tilde{\Gamma}_r^{(2)}(p^2 = 0) = 0, \quad (9.66)$$

$$\frac{\partial}{\partial p^2} \tilde{\Gamma}_r^{(2)}(p^2 = \mu^2) = 1, \quad (9.67)$$

$$\tilde{\Gamma}_r^{(4)}(p_i = \mu \theta_i) = g_r, \quad (9.68)$$

in which the θ_i form a set of arbitrary non-exceptional numerical vectors.

The bare vertex functions in a massless theory depend only on the cut-off and momenta, since the bare mass parameter is fixed by imposing that the renormalized mass m_r vanishes. The renormalized vertex functions depend on the arbitrary scale μ and momenta. They are related for Λ large by

$$\tilde{\Gamma}_r^{(n)}(p_i; \mu, g_r) = Z^{n/2}(\Lambda/\mu, g_r) \tilde{\Gamma}^{(n)}(p_i; \Lambda, g). \quad (9.69)$$

RG equations. The bare theory does not depend on the parameter μ , which has just been introduced to fix the normalization scale. Therefore,

$$\mu \frac{\partial}{\partial \mu} \tilde{\Gamma}^{(n)}(p_i; \Lambda, g) \Big|_{g, \Lambda} = 0.$$

Then if one differentiates equation (9.69) with respect to μ at Λ and g fixed, using chain rule, one finds [70–72]

$$\left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \tilde{\eta}(g_r) \right] \tilde{\Gamma}_r^{(n)}(p_i; \mu, g_r) = 0, \quad (9.70)$$

with the definitions

$$\tilde{\beta}(g_r) = \mu \frac{\partial}{\partial \mu} g_r \Big|_{g, \Lambda}, \quad \tilde{\eta}(g_r) = \mu \frac{\partial}{\partial \mu} \ln Z \Big|_{g, \Lambda}. \quad (9.71)$$

A priori, the functions $\tilde{\beta}$ and $\tilde{\eta}$ could also depend of the dimensionless ratio Λ/μ but, since they can also be expressed directly in terms of the renormalized vertex functions, they must have a large cut-off limit.

Equation (9.70) is analogous to equation (9.59). It differs only in the definition of g_r and a finite field renormalization. Both sets of equations will be essential tools for the analysis of the large momentum behaviour of vertex functions. Moreover, equation (9.70) is used in Chapter 15 to discuss the small momentum behaviour in massless theories.

ϕ^2 insertions. In the massless theory, one can also define renormalized vertex functions with ϕ^2 insertions (see Section 9.7 and equations (9.54) and (11.6)), given by

$$\begin{aligned} & \tilde{\Gamma}_r^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n; \mu, g_r) \\ &= Z^{(n/2)-l} Z_2^l \left[\tilde{\Gamma}^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n; \Lambda, g) - \delta_{l2} \delta_{n0} C(g) \right]. \end{aligned} \quad (9.72)$$

The ϕ^2 renormalization constant Z_2 can be fixed by a renormalization condition of the form

$$\tilde{\Gamma}_r^{(1,2)}(q; p_1, p_2) \Big|_{p_1^2 = p_2^2 = q^2 = \mu^2} = 1. \quad (9.73)$$

The additional renormalization constant $C(g)$ needed for the $\langle \phi^2 \phi^2 \rangle$ correlation function can be fixed by the renormalization condition (see equation (9.53)),

$$\tilde{\Gamma}_r^{(2,0)}(p^2 = \mu^2) = 0. \quad (9.74)$$

From equation (9.72), RG equations can be derived by differentiating with respect to μ . Introducing the RG function

$$\tilde{\eta}_2(g_r) = \mu \frac{\partial}{\partial \mu} \ln Z_2 \Big|_{g, \Lambda}, \quad (9.75)$$

one obtains (see equation (9.56)),

$$\begin{aligned} & \left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \tilde{\eta}(g_r) - l \tilde{\eta}_2(g_r) \right] \tilde{\Gamma}_r^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n; \mu, g_r) \\ &= \delta_{n0} \delta_{l2} C_r(g_r). \end{aligned} \quad (9.76)$$

9.10 Homogeneous RG equations: Massive QFT

The renormalized vertex functions $\tilde{\Gamma}_r^{(l,n)}$ of the massless theory have been defined and shown to satisfy RG equations. Since a constant source term for ϕ^2 at zero momentum generates a mass shift (equation (9.25)), one could think about expressing the vertex functions of a massive theory in terms of the vertex functions with ϕ^2 insertions of the massless theory [70–72]. However, an immediate problem arises: one verifies that ϕ^2 insertions at zero momentum in a massless theory are infrared (IR) divergent.

The problem can be solved by the following method: one introduces a space-dependent source $J_2(x)$ for $\phi^2(x)$ insertions [74]. One performs a partial summation of the two-point function, and then takes the constant source limit. After summation, the propagator becomes massive, and the limit is no longer IR divergent.

We consider the renormalized action with a source $J_2(x)$ for renormalized ϕ^2 insertions,

$$\mathcal{S}_r(\phi_r, J_2) = \int d^4x \left[\frac{1}{2} Z (\nabla \phi_r(x))^2 + \frac{1}{2} (\delta m^2 + Z_2 J_2(x)) \phi_r^2(x) + \frac{1}{4!} g_r Z_g \phi_r^4(x) \right], \quad (9.77)$$

where the renormalization constants are those of the massless theory. In the Fourier representation, a vertex function in the presence of the source $J_2(x)$, expressed in terms of its Fourier transform $\tilde{J}_2(q)$, has the expansion,

$$\begin{aligned} \tilde{G}_r^{(n)}(p_1, p_2, \dots, p_n; J_2) &= \sum_{l=0}^n \frac{1}{2^l} \frac{1}{l!} \int d^4q_1 d^4q_2 \cdots d^4q_l \tilde{J}_2(q_1) \tilde{J}_2(q_2) \cdots \tilde{J}_2(q_l) \\ &\times (2\pi)^4 \delta^{(4)}(q_1 + \cdots + q_l + p_1 + \cdots + p_n) \tilde{\Gamma}_r^{(l,n)}(q_1, q_2, \dots, q_l; p_1, p_2, \dots, p_n), \end{aligned} \quad (9.78)$$

in which the vertex function $\tilde{G}_r^{(n)}$ differs from $\tilde{\Gamma}_r^{(n)}$ because the δ -function of momentum conservation cannot be factorized when J_2 is space dependent.

Applying the differential operator

$$D \equiv \mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_r) \frac{\partial}{\partial g_r}, \quad (9.79)$$

to equation (9.78), using the RG equations (9.76), and noting that

$$\int d^4q \tilde{J}_2(q) \frac{\delta}{\delta \tilde{J}_2(q)} = l, \quad (9.80)$$

one obtains

$$\begin{aligned} &\left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \tilde{\eta}(g_r) - \tilde{\eta}_2(g_r) \int d^4q \tilde{J}_2(q) \frac{\delta}{\delta \tilde{J}_2(q)} \right] \\ &\times \tilde{G}_r^{(n)}(J_2; p_1, \dots, p_n) = 0. \end{aligned} \quad (9.81)$$

After some partial summation, one can set $\tilde{J}_2(q) = \tau_r$ constant and the equation becomes [70–74]

$$\left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \tilde{\eta}(g_r) - \tilde{\eta}_2(g_r) \tau_r \frac{\partial}{\partial \tau_r} \right] \tilde{\Gamma}_r^{(n)}(p_1, \dots, p_n; \mu, g_r, \tau_r) = 0. \quad (9.82)$$

Therefore, we have derived a new RG equation for a massive theory, which differs from the original CS equation in two respects:

- (i) Vertex functions now depend on two mass parameters while we know that only one is required. However, in this parametrization the massless limit of correlation functions is directly obtained by setting $\tau_r = 0$ and no asymptotic expansion at large momenta is needed.
- (ii) In contrast to the CS equation, the RG equation (9.82) is homogeneous. It is thus easier to solve.

9.10.1 Covariance of RG functions

In the example of the massless ϕ^4 theory, we have introduced two different renormalization schemes. It is interesting to understand how RG functions of different schemes are related. Theories may differ by a redefinition of the coupling constant and a finite field renormalization. We denote by g_r and \tilde{g}_r the coupling constants in two schemes. Comparing two RG equations, one infers

$$\beta(g_r) \frac{\partial}{\partial g_r} = \tilde{\beta}(\tilde{g}_r) \frac{\partial}{\partial \tilde{g}_r} \Rightarrow \beta(g_r) \frac{\partial \tilde{g}_r}{\partial g_r} = \tilde{\beta}(\tilde{g}_r). \quad (9.83)$$

In a ϕ^4 -like QFT, the function $\beta(g_r)$ has the expansion,

$$\beta(g_r) = \beta_2 g_r^2 + \beta_3 g_r^3 + O(g_r^4). \quad (9.84)$$

Expanding \tilde{g}_r in terms of g_r ,

$$\tilde{g}_r = g_r + \gamma_2 g_r^2 + O(g_r^3), \quad (9.85)$$

and using equation (9.84), after a short calculation one finds

$$\tilde{\beta}(\tilde{g}_r) = \beta_2 \tilde{g}_r^2 + \beta_3 \tilde{g}_r^3 + O(\tilde{g}_r^4). \quad (9.86)$$

The first two terms in the expansion of the β -function are universal and all others are formally arbitrary.

One should not conclude from this observation that the physical consequences inferred from the RG β -function for finite coupling are arbitrary. Only regular mappings $g_r \mapsto \tilde{g}_r$ are allowed. In particular, this implies that $\partial \tilde{g}_r / \partial g_r$ in equation (9.83) must be strictly positive. Therefore, the *sign and the existence of zeros* of the β -function are properties of the QFT. Equation (9.83) also shows that the slope of the β -function at a zero is scheme independent.

Of course, we do not know in general which renormalization scheme leads to regular functions of the coupling constant. Our intuition is that ‘natural’ definitions as induced by momentum or minimal subtraction have the most chance to satisfy this criterion.

In the same way, if we call $\zeta(\tilde{g}_r)$ the additional finite field renormalization, we find

$$\tilde{\eta}(\tilde{g}_r) = \eta(g_r) + \tilde{\beta}(\tilde{g}_r) \frac{\partial}{\partial \tilde{g}_r} \ln \zeta(\tilde{g}_r). \quad (9.87)$$

Since the field renormalization appears at order g_r^2 , $\ln \zeta(\tilde{g}_r)$ is of order \tilde{g}_r^2 and, therefore, the modification to $\tilde{\eta}$ of order \tilde{g}_r^3 . The coefficient of order g_r^2 is universal. The value of η at a zero of $\beta(g_r)$ is also universal. Similar arguments apply to η_2 : its first coefficient is universal as well as its value at a zero of $\beta(g_r)$.

Finally, these arguments generalize to the relation between bare (Section 9.11) and renormalized RG functions.

Several coupling constants. We shall meet actions depending on several fields and coupling constants g_i . In a change of parametrization of the coupling space, the RG β -functions transform like

$$\tilde{\beta}_i(\tilde{g}) = \sum_j \frac{\partial \tilde{g}_i}{\partial g_j} \beta_j(g), \quad (9.88)$$

where the mapping should satisfy $\det(T_{ij} \equiv \partial \tilde{g}_i / \partial g_j) > 0$. Thus the existence of zeros of the β -functions is universal and, since at a zero g_i^* ,

$$\frac{\partial \tilde{\beta}_i}{\partial \tilde{g}_j} = \sum_{k,l} T_{ik} \frac{\partial \beta_k}{\partial g_l} T_{lj}^{-1},$$

the eigenvalues of $\partial \beta_i / \partial g_j$ are scheme independent.

9.11 EFT and RG

In Sections 8.8.1 and 8.9, we have explained that the initial bare QFT is an *effective* microscopic theory (EFT) in which non-renormalizable interactions have been neglected. It is an approximation to a more complete *finite physical theory*, and designed to describe only its large distance scale, low energy properties. The cut-off Λ , introduced to render the QFT finite, represents the mass scale at which new physics has to be taken into account and at which the EFT is no longer relevant. In the framework of EFTs, the restriction to renormalizable QFTs is not a new law of nature but only a consequence of the weakness of the possible non-renormalizable interactions due to the large ratio between microscopic and physical scales.

The renormalizability of a QFT implies that the large distance, small physical mass behaviour of correlation functions is *universal*, that is, independent from the arbitrary cut-off implementation, order by order in a perturbative expansion, because the renormalized and bare correlation functions are asymptotically proportional.

The renormalized functions, when they exist beyond perturbation theory, contain the whole information about the asymptotic large distance properties.

One may then wonder why one should consider bare correlation functions. The main reason is that renormalization theory implies *tuning the parameters of the Lagrangian as a function of the cut-off*, something that is obviously non-physical in particle physics. The tuning of the ϕ^2 coefficient in the ϕ^4 QFT (it corresponds to adjusting the temperature close to the critical temperature in critical phenomena, see Chapter 15) seems unavoidable, and requires an explanation outside of the ϕ^4 theory. However, one wants to avoid at least the tuning of the ϕ^4 interaction, which, even for critical phenomena, corresponds to a non-generic situation.

A further reason is that the *renormalized theory may not exist beyond perturbation theory*, equation (9.90) being then valid only for Λ large but finite (see Section 9.12).

Finally, one may want to study corrections to the leading large distance behaviour.

Super-renormalizable interactions. Renormalization theory implicitly assumes that the scale in super-renormalizable interactions is the physical scale, for example, the physical mass, instead of the cut-off scale. This is a highly non-generic situation.

9.11.1 Bare (or microscopic) and renormalized vertex functions

Rescaling the parameter r , we rewrite the bare (effective microscopic) action (9.3) as

$$\mathcal{S}_\Lambda(\phi) = \int d^4x \left[\frac{1}{2} (\nabla_\Lambda \phi(x))^2 + \frac{1}{2} \Lambda^2 r \phi^2(x) + \frac{1}{4!} g \phi^4(x) \right], \quad (9.89)$$

where parameters g and r are dimensionless and the cut-off Λ is the only mass parameter.

For generic values of r , the physical mass m is of order Λ and all correlations vanish at large distance. For a particular negative value $r = r_c$ (the critical temperature in the sense of critical phenomena, see Section 15.3), which must exist beyond perturbation theory, the physical mass m_r vanishes and for $r < r_c$ the \mathbb{Z}_2 reflection symmetry is spontaneously broken (see Section 16.5.4). A small physical mass in the unbroken phase implies $1 \gg r - r_c \geq 0$, a condition that implies a *fine tuning* of the parameter r .

Studying the low energy, low mass situation is equivalent to studying the large cut-off limit. Thus, we can use the results of the renormalization theory. For simplicity, specializing to $r = r_c$, the value at which the physical mass m_r vanishes, we introduce rescaled vertex functions, defined by the renormalization conditions (9.67) and (9.68), adapted to a massless theory at a physical scale $\mu \ll \Lambda$, and functions of a renormalized coupling constant g_r .

These vertex functions are related to the microscopic (bare) ones by equations of the form (9.69), which we write here slightly differently as

$$\tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu, \Lambda) = Z^{n/2}(g, \Lambda/\mu) \tilde{\Gamma}^{(n)}(p_i; g, \Lambda). \quad (9.90)$$

Renormalization theory tells us that the functions $\tilde{\Gamma}_r^{(n)}$ in equation (9.90) have, at p_i , g_r , and μ fixed, and *order by order in the perturbative expansion*, finite limits for $\Lambda \rightarrow \infty$. Moreover, the renormalized functions $\tilde{\Gamma}_r^{(n)}$ do not depend on the specific regularization procedure and, given the normalization conditions (9.68), are *universal*. More precisely, the corrections at loop order L are of the order $(\ln \Lambda)^L / \Lambda^2$. We assume here that the sum of all corrections still vanishes for $\Lambda \rightarrow \infty$, something that has not been proved here.

9.11.2 Bare or asymptotic microscopic RG equations

Differentiating equation (9.90) with respect to Λ at g_r and μ fixed, one derives the new identity

$$\Lambda \frac{\partial}{\partial \Lambda} \Big|_{g_r, \mu \text{ fixed}} Z^{n/2}(g, \Lambda/\mu) \tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = o(\Lambda^{-1}). \quad (9.91)$$

Using chain rule, one infers from equation (9.91),

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g, \Lambda/\mu) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g, \Lambda/\mu) \right] \tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = 0, \quad (9.92)$$

where corrections vanishing for Λ large have been neglected (see Chapter 17). The functions β and η are given by

$$\beta(g, \Lambda/\mu) = \Lambda \frac{\partial}{\partial \Lambda} \Big|_{g_r, \mu} g, \quad \eta(g, \Lambda/\mu) = -\Lambda \frac{\partial}{\partial \Lambda} \Big|_{g_r, \mu} \ln Z(g, \Lambda/\mu). \quad (9.93)$$

Being dimensionless, they may depend only on the dimensionless quantities g and Λ/μ . However, the functions β and η can also be directly calculated from equation (9.92) in terms of functions $\tilde{\Gamma}^{(n)}$ which do not depend on μ . Therefore, the functions β and η cannot depend on the ratio Λ/μ . This simplifies equation (9.92), which becomes [72]

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right) \tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = 0. \quad (9.94)$$

Equation (9.94) is satisfied, when the *cut-off is large with respect to the physical scale, but not infinite*, by the bare vertex functions of the ϕ^4 QFT. It is a direct consequence of the existence of a perturbative renormalized theory. Its solution also implies the existence of a renormalized theory, but in a slightly different form and with a new interpretation.

Massive theory. RG equations can also be proved for $0 < (r - r_c) \ll 1$. They take a form analogous to equations (9.82), τ_r being replaced by $\tau = r - r_c$ [72].

Leading corrections for Λ large. Finally, we characterize the terms neglected in equation (9.94) more precisely. In a series expansion in powers of g ,

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = \sum_{\nu} \tilde{\Gamma}_{\nu}^{(n)}(p_i, \Lambda) g^{\nu}, \quad (9.95)$$

the coefficients $\tilde{\Gamma}_{\nu}^{(n)}$ have an asymptotic expansion for all $|p_i|/\Lambda$ small of the form

$$\tilde{\Gamma}_{\nu}^{(n)}(p_i, \Lambda) = \sum_{\ell=0}^{L(n, \nu)} \left[(\ln \Lambda)^{\ell} A_{\ell \nu}^{(1)}(p_i) + \frac{1}{\Lambda^2} (\ln \Lambda)^{\ell} A_{\ell \nu}^{(2)}(p_i) + \dots \right], \quad (9.96)$$

with $L(n, \nu) = \nu - 1$ for $n = 4$ and $L(n, \nu) = \nu - n/2$ for $n \neq 4$.

The RG equation (9.94) is exact for the sum of the perturbative contributions which do not vanish for Λ large, as can be verified by expanding equation (9.94) in powers of g .

Bare and renormalized RG equations. Formally, the bare and renormalized RG equations differ only by their parametrization. Even the first terms of the small coupling expansion are the same (Section 9.10.1). However, the bare RG equations combined with the Gaussian renormalization, describing the RG flow from the microscopic scale to the physical scale, determine some properties of the renormalized parameters. By contrast, the renormalized equations describe only the flow between different physical scales. Still, the two approaches are equivalent to study the leading IR behaviour of a massless (or critical) or near massless QFT, but the latter is more practical.

9.12 Solution of bare RG equations: The triviality issue

The RG equations (9.94) (as well as equations (9.70)) can be solved by the method of characteristics. Introducing a dilatation parameter λ , one looks for functions $g(\lambda)$ and $Z(\lambda)$ such that

$$\lambda \frac{d}{d\lambda} \left[Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \Lambda\lambda) \right] = 0. \quad (9.97)$$

Differentiating explicitly with respect to λ , one finds that equations (9.97) and (9.94) are compatible provided that

$$\lambda \frac{d}{d\lambda} g(\lambda) = \beta(g(\lambda)), \quad (9.98)$$

$$\lambda \frac{d}{d\lambda} \ln Z(\lambda) = \eta(g(\lambda)), \quad (9.99)$$

or after integration with the boundary conditions $g(1) = g$ and $Z(1) = 1$,

$$\ln \lambda = \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} \quad (9.100)$$

$$\ln Z(\lambda) = \int_1^\lambda \frac{d\sigma}{\sigma} \eta(g(\sigma)) = \int_g^{g(\lambda)} dg' \frac{\eta(g')}{\beta(g')}. \quad (9.101)$$

Equation (9.97) implies

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \Lambda\lambda),$$

and after the rescaling $\Lambda\lambda \mapsto \Lambda$, since $\tilde{\Gamma}^{(n)}$ has mass dimension $(4 - n)$,

$$\tilde{\Gamma}^{(n)}(\lambda p_i; g, \Lambda) = \lambda^{4-n} Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \Lambda). \quad (9.102)$$

The equation shows that the function $g(\lambda)$ is the effective coupling at the scale $\mu = \lambda\Lambda$ and we are interested in the limit $\lambda \rightarrow 0$. Equation (9.94) is the RG equation in differential form. Equations (9.102), (9.100), and (9.101) are the integrated RG equations.

9.12.1 RG functions at leading order

After a short calculation, one obtains for the massless propagator

$$\Delta(x) = \frac{1}{4\pi^2} \frac{1}{x^2}. \quad (9.103)$$

The one-loop contribution to the two-point function is a constant which is cancelled by a shift of r_c and thus $\eta(g) = O(g^2)$. The one-loop diagram contributing to the four-point function (first diagram in Fig. 9.2) can be written as

$$B_4(p) = \int d^4x e^{ipx} \Delta^2(x).$$

The integral diverges at $x = 0$ and, at this leading order, we simply cut the integral at $|x| = 1/\Lambda$ and, after having also cut the integral in the IR at $|x| = 1/\mu$ for $|x|$ large, we set $p = 0$. The divergent part (consistent with equation (9.29)) is then

$$B_4(p) \sim \frac{1}{8\pi^2} \ln(\Lambda/\mu) \Rightarrow \tilde{\Gamma}^{(4)}(p_i; g, \Lambda) = g - \frac{3}{16\pi^2} [\ln(\Lambda/\mu) + O(1)] g^2 + O(g^3). \quad (9.104)$$

Substituting the expansion into equation (9.94) with $n = 4$, one infers (as expected from the discussion of Section 9.10.1)

$$\beta(g) = \frac{3}{16\pi^2} g^2 + O(g^3). \quad (9.105)$$

The two-loop diagram contributing to the two-point function (third diagram in Fig. 9.1) can be written as

$$\Omega_4(p) = \int d^4x e^{ipx} \Delta^3(x).$$

Again we introduce an IR cut-off μ and a UV cut-off Λ . We can then expand in powers of the momentum p . The relevant contribution is of order p^2 and given by

$$[\Omega_4(p)]_{\text{div.}} \sim -\frac{1}{8} p^2 \int d^4x x^2 \Delta^3(x) \sim -\frac{1}{4} \frac{1}{(8\pi^2)^2} p^2 \ln(\Lambda/\mu). \quad (9.106)$$

Inserting the result into equation (9.94) with $n = 2$, one finds

$$\eta(g) = \frac{1}{(8\pi^2)^2} \frac{g^2}{24} + O(g^3). \quad (9.107)$$

Since the β -function is of order g^2 , for $g(\lambda) \rightarrow 0$, $Z(\lambda)$ has a constant limit (for details see Section 17.2).

9.12.2 The triviality issue

For g small, the β -function is positive. If it remains positive for all $g > 0$ (in agreement with empirical evidence), then equation (9.98) implies that $g(\lambda)$ is an increasing function of λ . For physical scales μ such that $\mu/\Lambda = \lambda \rightarrow 0$, in equation (9.100) $\ln \lambda \rightarrow -\infty$ and the integral in the right-hand side must diverge. Therefore, the effective coupling $g(\lambda)$ must go to zero. The integral is then dominated by the g^2 term of $\beta(g)$ (equation (9.105)). One concludes

$$g(\lambda = \mu/\Lambda) \sim \frac{16\pi^2}{3} \frac{1}{\ln(\Lambda/\mu)}. \quad (9.108)$$

The effective coupling $g(\mu/\Lambda)$ can be identified with the renormalized coupling g_r . This leads to the *triviality* issue: in the infinite cut-off limit, the *renormalized coupling vanishes*: it is impossible to define a *renormalized* ϕ^4 QFT in four dimensions for non-zero coupling [73]. However, in the logic of *EFTs*, the problem is reformulated differently. The QFT has only a limited energy or momentum range of validity, where it is consistent (in the sense of satisfying all usual physical requirements). This range is small compared to the cut-off, which represents the scale of some new physics and can no longer be assumed to be infinite.

RG arguments then imply that the consistent physical range decreases when the renormalized or effective charge increases, as indicated by equation (9.108). Note that if g is generic (not too small) and Λ/μ very large, then the renormalized coupling constant g_r becomes essentially independent of the initial coupling constant g .

Finally, in the $O(N)$ symmetric $(\phi^2)^2$ field theory, for N large, one finds that the renormalized theory has a non-physical pole, dubbed the ‘Landau ghost’ (Section 18.5).

A9 Functional RG equations. Super-renormalizable QFTs. Normal order

A9.1 Large-momentum mode integration and functional RG equations

Here, we briefly describe a direct construction of a general functional RG, closer to Wilson's ideas [62] (in a form also developed by Wegner [61]), which was proposed by Polchinski [59] as an alternative method to prove renormalizability. The physical motivation for such a method will become more apparent when we discuss RG and critical phenomena starting in Chapter 15. The method is based on a recursive integration over the large momentum modes of the field. The procedure leads to an exact RG in the space of all local interactions: one expresses the equivalence between a variation of the cut-off and a modification of the coefficients of the interaction terms. The equivalence takes the form of functional RG equations for an effective Hamiltonian, valid for large distance or low momentum (in the cut-off scale). For a simple and much more detailed introduction see, for example, Chapter 16 of Ref. [64].

A9.1.1 A basic equivalence

From the arguments of Section 8.4.3, based on a Gaussian integration, it follows that the two actions (translation invariance is assumed)

$$\begin{aligned}\mathcal{S}(\phi) &= \frac{1}{2} \int d^d x d^d y \phi(x) \Delta^{-1}(x-y) \phi(y) + V(\phi), \quad \text{and} \\ \mathcal{S}(\phi_1, \phi_2) &= \frac{1}{2} \int d^d x d^d y [\phi_1(x) \Delta_1^{-1}(x-y) \phi_1(y) + \phi_2(x) \Delta_2^{-1}(x-y) \phi_2(y)] \\ &\quad + V(\phi_1 + \phi_2),\end{aligned}\tag{A9.1}$$

with $\Delta = \Delta_1 + \Delta_2$, generate the same perturbation theory.

Here, Δ^{-1} means inverse of Δ in the sense of operators (as well as $\Delta_1^{-1}, \Delta_2^{-1}$):

$$\int d^d z \Delta^{-1}(x-z) \Delta(z-y) = \delta^{(d)}(x-y).$$

We take for Δ, Δ_1 and Δ_2 massless propagators of the general form

$$\Delta(x-y) = \frac{1}{(2\pi)^d} \int d^d k e^{i(x-y)} \tilde{\Delta}(k), \quad \text{with } \tilde{\Delta}(k) = \frac{C(k^2/\Lambda^2)}{k^2},\tag{A9.2}$$

in which the function $C(t)$ is smooth, go to 1 for t small and decreases faster than any power for t large (see Section A8.3), and Λ is a large momentum cut-off.

We now use this equivalence in the limit in which the propagator Δ_2 goes to 0. Then only small values of ϕ_2 contribute to the partition function. Expanding the interaction for ϕ_2 small,

$$V(\phi_1 + \phi_2) = V(\phi_1) + \int d^d x \frac{\delta V(\phi_1)}{\delta \phi(x)} \phi_2(x) + \frac{1}{2} \int d^d x d^d y \frac{\delta^2 V}{\delta \phi(x) \delta \phi(y)} \phi_2(x) \phi_2(y) + \dots,$$

we integrate over ϕ_2 to obtain the leading order correction

$$\begin{aligned}&\int [d\phi_2] \exp \left\{ - \left[\frac{1}{2} \int d^d x d^d y \phi_2(x) \Delta_2^{-1}(x-y) \phi_2(y) + V(\phi + \phi_2) - V(\phi) \right] \right\} \\ &\sim 1 + \frac{1}{2} \int d^d x d^d y \Delta_2(x-y) \left[\frac{\delta V}{\delta \phi(x)} \frac{\delta V}{\delta \phi(y)} - \frac{\delta^2 V}{\delta \phi(x) \delta \phi(y)} \right] + \dots\end{aligned}\tag{A9.3}$$

Taking the logarithm of expression (A9.3), we rewrite the partition function as a field integral over a field ϕ (the index 1 is no longer useful) with the effective action

$$\begin{aligned} \mathcal{S}'(\phi) &= \frac{1}{2} \int d^d x d^d y \phi(x) \Delta_1^{-1}(x-y) \phi(y) + V(\phi) \\ &\quad + \frac{1}{2} \int d^d x d^d y \Delta_2(x-y) \left[\frac{\delta^2 V}{\delta \phi(x) \delta \phi(y)} - \frac{\delta V}{\delta \phi(x)} \frac{\delta V}{\delta \phi(y)} \right] + \dots . \end{aligned} \quad (\text{A9.4})$$

We have thus established that the actions (A9.1) and (A9.4) lead to the same partition function, up to a trivial change of normalization. Note that the same identity can be proven by playing with integrations by parts as in the case of the quantum equation of motion (7.33).

We will now show that this equivalence can be used to partially integrate out the large momenta in a field theory with a cut-off Λ .

A9.1.2 Large-momentum mode partial integration and RG equations

We now choose

$$\tilde{\Delta}(k) = \frac{C(k^2/\Lambda^2)}{k^2}, \quad \tilde{\Delta}_1(k) = \frac{C(k^2/\Lambda^2(1+\sigma)^2)}{k^2}, \quad (\text{A9.5})$$

where the propagator Δ_1 is obtained from Δ by a rescaling of the cut-off Λ .

For $\sigma \rightarrow 0$, at leading order Δ_2 is then given by

$$\tilde{\Delta}_2(k) \sim \frac{2\sigma}{\Lambda^2} C'(k^2/\Lambda^2) \equiv \sigma \tilde{D}(k). \quad (\text{A9.6})$$

We note that the propagator $\tilde{\Delta}_2$ has no pole at $k = 0$. Moreover, if we choose a function $C(t)$ which is very close to 1 for t small

$$|C(t) - 1| t^{-p} \rightarrow 0 \quad \forall p > 0,$$

then $\tilde{\Delta}_2$ is only large for k of order Λ . The integration over ϕ_2 thus corresponds to an integration over the large momentum modes of the field ϕ .

The equivalence between actions (A9.1) and (A9.4), which is the starting point of an RG, can be written as ($D(x)$ is the Fourier transform of $\tilde{D}(k)$)

$$\Lambda \frac{d}{d^d \Lambda} V(\phi, \Lambda) = \frac{1}{2} \int d^d x d^d y D(x-y) \left[\frac{\delta^2 V}{\delta \phi(x) \delta \phi(y)} - \frac{\delta V}{\delta \phi(x)} \frac{\delta V}{\delta \phi(y)} \right], \quad (\text{A9.7})$$

or, after Fourier transformation ($\tilde{\phi}(p)$ is the Fourier transform $\phi(x)$),

$$\Lambda \frac{d}{d \Lambda} V(\phi, \Lambda) = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \tilde{D}(k) \left[\frac{\delta^2 V}{\delta \tilde{\phi}(k) \delta \tilde{\phi}(-k)} - \frac{\delta V}{\delta \tilde{\phi}(k)} \frac{\delta V}{\delta \tilde{\phi}(-k)} \right]. \quad (\text{A9.8})$$

The equation can also be derived from the quantum field equations and, therefore, partial integration does not imply a loss of information, by contrast with the lattice equation.

To study the existence of fixed points, we start with a given interaction $V_0(\phi)$ at a scale Λ_0 and use equation (A9.7) to calculate the effective interaction $V(\phi, \Lambda)$ at a scale $\Lambda \ll \Lambda_0$. A fixed point is defined by the property that $V(\phi, \Lambda)$, after a suitable rescaling of ϕ , goes to a limit.

We denote by $\tilde{V}^{(n)}(p_1, p_2, \dots, p_n)$ the coefficients of $V(\phi, \Lambda)$ in an expansion in powers of $\tilde{\phi}(p)$. Equation (A9.8) can then be written in component form as

$$\begin{aligned} \Lambda \frac{d}{d\Lambda} \tilde{V}^{(n)}(p_1, p_2, \dots, p_n) &= \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \tilde{D}(k) \tilde{V}^{(n+2)}(p_1, p_2, \dots, p_n, k, -k) \\ &\quad - \frac{1}{2} \sum_I D(p_0) \tilde{V}^{(l+1)}(p_{i_1}, \dots, p_{i_l}, p_0) \tilde{V}^{(n-l+1)}(p_{i_{l+1}}, \dots, p_{i_n}, -p_0), \end{aligned} \quad (\text{A9.9})$$

in which the momentum p_0 is determined by momentum conservation, and the set $I \equiv \{i_1, i_2, \dots, i_l\}$ runs over all distinct subsets of $\{1, 2, \dots, n\}$.

Equation (A9.9) shows that, even if we start with a pure $g\phi^4$ interaction, at scale Λ we obtain a general local interaction because all functions $\tilde{V}^{(n)}$ are coupled. However, in the spirit of the perturbative methods used so far, it is possible to solve equation (A9.7) as an expansion in the coupling constant g with the ansatz that the terms of $V(\phi, \Lambda)$ quadratic and quartic in ϕ are of order g and the general term of degree $2n$ is of order g^{n-1} .

Correlation functions. To generate correlation functions, one has to add a source to the interaction $V(\phi)$:

$$V(\phi) \mapsto V(\phi) - \int d^d x J(x) \phi(x).$$

However, equation (A9.4) then shows that $\mathcal{S}'(\phi)$ becomes in general a complicated functional of the source $J(x)$. A solution to this problem is the following: one takes a source whose Fourier transform $\tilde{J}(k)$ vanishes for $k^2 \geq \Lambda^2$, together with a propagator Δ_2 which propagates only momenta such that $k^2 \geq \Lambda^2$. This implies that $C'(t)$ vanishes identically for $t \leq 1$ (unfortunately, such cut-off functions are inconvenient for practical calculations). Then $\int d^d x J(x) \phi_2(x)$ does not contribute in integral (A9.3) and $\mathcal{S}'(\phi) - \mathcal{S}(\phi)$ does not depend on $J(x)$.

However, we note that then the RG transformation is such that the correlation functions corresponding to the action $\mathcal{S}(\phi)$ and $\mathcal{S}'(\phi)$ are only identical when all momenta are smaller than the cut-off. The differences between correlation functions are smooth functions of momenta and thus decay at large distances in space faster than any power.

A9.2 The ϕ^4 QFT in three dimensions: Divergences

Although much of this work is focused on strictly renormalizable QFTs, we make a few remarks here about UV divergences in super-renormalizable scalar QFTs in three and two dimensions. Note that, below four dimensions, the perturbative expansion of the *massless theory is IR divergent* (see Section 10.5.6).

We first take the example of the ϕ^4 QFT in three dimensions. As explained in Chapter 8, the ϕ^4 QFT then has only three superficially divergent diagrams, all contributing to the two-point function, which are displayed in Fig. 9.21.

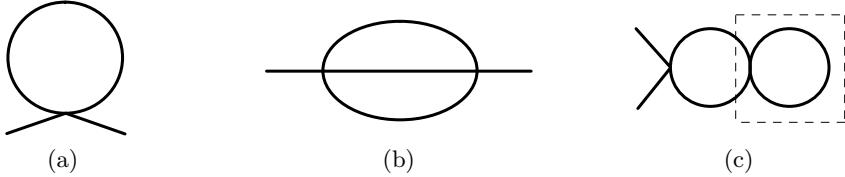


Fig. 9.21 The three divergent diagrams in the ϕ_3^4 QFT

The two first diagrams are given by

$$(a) = \frac{1}{(2\pi)^3} \int \frac{d^3 q}{(q^2 + m^2)_\Lambda}.$$

$$(b) = \frac{1}{(2\pi)^6} \int \frac{d^3 q_1 d^3 q_2}{(q_1^2 + m^2)_\Lambda (q_2^2 + m^2)_\Lambda \left[(p - q_1 - q_2)^2 + m^2 \right]_\Lambda}.$$

By adding counter-terms of the form

$$\frac{1}{2} \int d^3 x [a_1(\Lambda)g + a_2(\Lambda)g^2] \phi^2(x),$$

we can render the first two diagrams finite.

The diagram (c) is also superficially divergent, but it is clear that the counter-term which renders the diagram (a) finite, also renormalizes the diagram (c). Thus, the counter-terms that renormalize the diagrams (a) and (b) render the whole theory finite.

Actually the divergence (a), which corresponds to a self-contraction of the vertex ϕ^4 , can be eliminated *a priori* by replacing the vertex ϕ^4 by a *normal-ordered vertex* $: \phi^4 :_n$,

$$: \phi^4 :_n(x) = \phi^4(x) - 6\phi^2(x) \langle \phi^2(x) \rangle + 3 (\langle \phi^2(x) \rangle)^2, \quad (A9.10)$$

in which the expectation value $\langle \phi^2(x) \rangle$ is calculated in a free QFT with a mass μ which may or may not be equal to the renormalized mass m ,

$$\langle \phi^2(x) \rangle_\mu = \frac{1}{(2\pi)^3} \int \frac{d^3 q}{(q^2 + \mu^2)_\Lambda}. \quad (A9.11)$$

The denomination normal order comes from the operator terminology. The quantity $: \phi^4(x) :$ is such that

$$\begin{aligned} \langle : \phi^4(x) : \rangle &= 0 \\ \langle : \phi^4(x) : \phi(y_1) \phi(y_2) \rangle &= 0, \end{aligned} \quad (A9.12)$$

in which again the expectation values are calculated with the action $S_\mu(\phi)$,

$$S_\mu(\phi) = \frac{1}{2} \int d^3 x \left[(\nabla \phi(x))^2 + \mu^2 \phi^2(x) \right]. \quad (A9.13)$$

The expectation values calculated with another mass are then finite. Finally, we still have to add a counter-term for diagram (b).

A9.3 Super-renormalizable scalar QFTs in two dimensions: Normal order

An action $\mathcal{S}(\phi)$ of the form

$$\mathcal{S}(\phi) = \int d^2x \left[\frac{1}{2}(\nabla\phi(x))^2 + \frac{1}{2}m^2\phi^2(x) + V(\phi(x)) \right], \quad (A9.14)$$

in which $V(\phi)$ is an arbitrary regular function of ϕ , is super-renormalizable. If $V(\phi)$ is a polynomial, only a finite number of diagrams are superficially divergent. However, it is a peculiarity of dimension 2 that the field is dimensionless and, therefore, the interaction $V(\phi)$ may have an infinite series expansion in powers of ϕ . Although the theory is super-renormalizable, one finds an infinite number of superficially divergent diagrams. One then notes that all divergences come from self-contractions of the vertex (see Fig. 9.22). Therefore, the operation : $V(\phi)$: removes all divergences.



Fig. 9.22 Divergent contributions

To obtain an explicit expression for : $V(\phi)$:, we first consider the special interaction

$$V(\phi) = e^{\lambda\phi}. \quad (A9.15)$$

The expectation value of $V(\phi)$ in the presence of a source term can then be calculated explicitly:

$$\begin{aligned} & \int [d\phi] \exp \left\{ - \int d^2y \left[\frac{1}{2} \left((\nabla\phi(y))^2 + \mu^2\phi^2(y) \right) - \left(J(y) + \lambda\delta^{(2)}(x-y) \right) \phi(y) \right] \right\} \\ &= \exp \left[\frac{1}{2} \int d^2y d^2y' J(y) \Delta(y, y') J(y') + \lambda \int d^2y \Delta(x, y) J(y) + \frac{1}{2}\lambda^2 \Delta(x, x) \right], \end{aligned} \quad (A9.16)$$

in which $\Delta(x, y)$ is the free propagator with mass μ . The normal ordering operation has to suppress the term coming from self-contractions, which is proportional to $\Delta(x, x)$. It is thus clear that the normal ordered interaction is

$$: \exp [\lambda\phi(x)] : = \exp [\lambda\phi(x) - \lambda^2 \langle \phi^2(x) \rangle / 2], \quad (A9.17)$$

in which we have used

$$\langle \phi^2(x) \rangle = \Delta(x, x). \quad (A9.18)$$

We then express an arbitrary interaction term as a Laplace transform,

$$V(\phi(x)) = \int d\rho(\lambda) e^{\lambda\phi(x)}. \quad (A9.19)$$

The normal ordering is a linear operation. Thus,

$$: V(\phi) : = \int d\rho(\lambda) : e^{\lambda\phi(x)} : .$$

We then use the result (A9.17) for the exponential interaction and obtain

$$:V(\phi): = \left\{ \exp \left[-\frac{1}{2} \langle \phi^2 \rangle (\partial / \partial \phi)^2 \right] \right\} V(\phi), \quad (A9.20)$$

or more explicitly,

$$:V(\phi): = V(\phi) + \left[\sum_{n=1}^{\infty} \frac{(-1)^n}{2^n n!} \langle \phi^2 \rangle^n \left(\frac{\partial}{\partial \phi} \right)^{2n} \right] V(\phi).$$

The existence of the Laplace transform of a given interaction is irrelevant in this argument, since the final identity is purely algebraic.

10 Dimensional continuation, regularization, minimal subtraction (MS). Renormalization group (RG) functions

In this chapter, we introduce the notions of *dimensional continuation* and *dimensional regularization*, by defining a continuation of Feynman diagrams to analytic functions of the space dimension.

Dimensional continuation, which is essential for generating Wilson–Fisher’s famous ε -expansion in the theory of critical phenomena [75], and dimensional regularization seem to have no meaning outside the perturbative expansion of quantum field theory (QFT), and thus no direct physical interpretation.

Dimensional regularization is a powerful regularization technique, which is often used, when applicable, because it leads to simpler perturbative calculations. Dimensional regularization *performs a partial renormalization*, cancelling what would show up as power-law divergences in momentum or lattice regularization. In particular, it cancels the commutator of quantum operators in local QFTs. These cancellations may be convenient, but may also, occasionally remove divergences that have an important physical meaning.

It is not applicable when some essential property of the field theory is specific to the initial dimension. For example, in even space dimensions, the relation between γ_S (identical to γ_5 in four dimensions) and the other γ matrices involving the completely antisymmetric tensor $\epsilon_{\mu_1 \dots \mu_d}$ (see Section A12) may be needed in theories violating parity symmetry.

Its use requires some care in massless theories, because its rules may lead to unwanted cancellations between ultraviolet (UV) and infrared (IR) logarithmic divergences.

Within the framework of dimensional regularization, we also introduce the concept of renormalization by *MS*. We show that the MS scheme simplifies the expressions of renormalization constants and RG functions.

We perform explicit calculations at two-loop order, first in the simple one-component ϕ^4 QFT, and then in an N -component QFT with a general four-field interaction. The results of these calculations will be directly used for lower order estimates of critical exponents (see Sections 15.2–15.5).

10.1 Dimensional continuation and dimensional regularization

We first define dimensional continuation, and then dimensional regularization, which simplifies perturbative calculations in QFT.

10.1.1 Dimensional continuation

We express the Fourier representation of a regularized scalar propagator $\tilde{\Delta}(p)$ of a massive theory, in d dimensions, as a Laplace transform (see equation (A8.23)),

$$\tilde{\Delta}(p) = \int_0^\infty dt \rho(t\Lambda^2) e^{-(p^2+m^2)t}, \quad (10.1)$$

where Λ is the cut-off and

$$\lim_{t \rightarrow \infty} \rho(t) = 1, \quad \rho(t) = O(t^\sigma), \text{ with } \sigma > d, \text{ for } t \rightarrow 0.$$

For $\rho(t) \equiv 1$, one recovers the unregularized propagator $1/(p^2 + m^2)$.

Using the representation (10.1), and the property that the momentum associated with an internal line is a linear combination of loop and external momenta, one can then express a scalar Feynman diagram γ with constant vertices, and external momenta p_α , in the form

$$I_\gamma(\{p_\alpha\}) = (2\pi)^{-Ld} \int \prod_{i=1}^I dt_i \rho(\Lambda^2 t_i) \prod_{\ell=1}^L d^d q_\ell \\ \times \exp \left[- \sum_{\ell, \ell'=1}^L q_\ell \cdot q_{\ell'} M_{\ell \ell'}(t_i) - 2 \sum_{\ell=1}^L q_\ell \cdot k_\ell(\{p_\alpha\}, t_i) - S[\{(p_\alpha)\}, t_i] \right]. \quad (10.2)$$

The Gaussian integration over all loop momenta q_ℓ can then be performed. The resulting expression is

$$I_\gamma(p) = \frac{1}{(4\pi)^{Ld/2}} \int \prod_{i=1}^I dt_i \rho(\Lambda^2 t_i) (\det \mathbf{M})^{-d/2} \\ \times \exp \left[\sum_1^L k_\ell (M^{-1})_{\ell \ell'} k_{\ell'} - S(\{p_\alpha\}, t_i) \right]. \quad (10.3)$$

In the integrated expression, the dependence in the dimension d is explicit and, therefore, continuation in d is achieved. This continuation makes it possible to study properties of phase transitions near two (Chapter 19) and four dimensions (Chapter 15).

10.1.2 Dimensional regularization: Defining properties of d -dimensional integrals

After continuation, we can choose $\text{Re } d$, small enough, even negative if necessary, such that all Feynman diagrams become formally convergent in the sense of power counting and take the infinite cut-off limit. The result is a meromorphic function of d , the dimensional regularization of the Feynman diagram [76].

If $d > 2$, one can then also take the massless limit.

One can also define directly an integral in d dimensions by a set of conditions which, when d is an integer, lead to the usual integral. It should satisfy the three conditions,

- (i) $\int d^d p F(p+q) = \int d^d p F(p)$ translation
- (ii) $\int d^d p F(\lambda p) = |\lambda|^{-d} \int d^d p F(p)$ dilatation
- (iii) $\int d^d p d^{d'} q f(p)g(q) = \int d^d p f(p) \int d^{d'} q g(q)$ factorization.

These simple rules define a dimensional continuation of Feynman diagrams. It is equivalent to the previous method. Indeed, from property (iii), one derives

$$\int d^d p e^{-tp^2} = \left(\int ds e^{-ts^2} \right)^d = \left(\frac{\pi}{t} \right)^{d/2}. \quad (10.4)$$

and then one can calculate the Feynman diagram using the Laplace representation. However, one can also use the rules more directly with an unregularized propagator.

Examples.

(i) The massless propagator is given by

$$\Delta(x) = \frac{1}{(2\pi)^d} \int d^d p \frac{e^{ip \cdot x}}{p^2} = \frac{1}{(2\pi)^d} \int_0^\infty ds \int d^d p e^{-sp^2 + ip \cdot x}.$$

The Gaussian momentum integral can be performed, followed by the s integral, and the result is

$$\Delta(x) = \frac{\Gamma(d/2 - 1)}{4\pi^{d/2}} |x|^{2-d}. \quad (10.5)$$

The pole at $d = 2$ reflects the IR divergences of the two-dimensional massless theory.

(ii) The one-loop contribution to the two-point function in the massless ($m = 0$) ϕ^3 QFT as well as the contribution to the four-point function in the ϕ^4 QFT (first diagram of Fig. 9.2) involve the function

$$\begin{aligned} I_\gamma(p) &= \int d^d x e^{ip \cdot x} \Delta^2(x) = \frac{\Gamma^2(d/2 - 1)}{16\pi^d} \int d^d x e^{ip \cdot x} |x|^{4-2d} \\ &= \frac{\Gamma^2(d/2 - 1)}{16\pi^d \Gamma(d-2)} \int_0^\infty ds s^{d-3} \int d^d x e^{ip \cdot x - sx^2}. \end{aligned}$$

Finally, integrating over x and then s , one obtains,

$$I_\gamma(p) = \frac{1}{(4\pi)^{d/2}} \Gamma(2 - d/2) \frac{\Gamma^2((d/2) - 1)}{\Gamma(d-2)} (p^2)^{(d/2)-2}. \quad (10.6)$$

This expression has a pole at $d = 2$ corresponding to IR (low momentum) singularities because the theory is massless and has poles at $d = 4, 6$, and so on, which clearly are consequences of the UV (large momentum) divergences of the Feynman diagram.

It is interesting to explain the interplay between dimensional continuation and cut-off regularization in this example. If we regularize the propagator, for example by the method of Pauli–Villars, the Feynman diagram I_γ becomes a regular function of d for $d > 2$ up to some even integer larger than 4.

In the neighbourhood of $d = 4$, it has the form

$$I_\gamma \sim \frac{1}{8\pi^2 (4-d)} \left[(p^2)^{(d/2)-2} - \Lambda^{d-4} \right], \quad \text{for } d \rightarrow 4.$$

At $d < 4$ fixed, the limit of infinite cut-off yields the continuation of the initial diagram with a pole at $d = 4$. At cut-off fixed, the $d = 4$ limit yields a finite result in which $\ln \Lambda$ has replaced the pole at $d = 4$.

(iii) The one-loop contribution to the two-point function in a massive ϕ^4 QFT (Fig. 9.1) is proportional to

$$\begin{aligned} \Omega_d(m) &= \int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} = \int_0^\infty dt \frac{d^d q}{(2\pi)^d} e^{-t(q^2 + m^2)} \\ &= \frac{1}{(4\pi)^{d/2}} \int_0^\infty dt t^{-d/2} e^{-m^2 t} = \frac{1}{(4\pi)^{d/2}} \Gamma(1 - d/2) m^{d-2}. \end{aligned} \quad (10.7)$$

This expression has poles for $d = 2 + 2n$, $n \geq 0$, corresponding to expected UV divergences, but nothing equivalent to quadratic divergences. The divergence at $d = 4$ obtained after analytic continuation has the same nature as at $d = 2$.

Important remark. The expression (10.7) has the following property: for $\text{Re } d > 2$, the limit $m = 0$ vanishes. The result generalizes, leading to the peculiar result

$$\int \frac{d^d p}{p^2} = 0, \quad (10.8)$$

for this integral that exists for no value of d .

The argument generalizes, for $\text{Re } d > 2n$, to

$$\int \frac{d^d p}{p^{2n}} = 0. \quad (10.9)$$

This result confirms that dimensional regularization is a partial renormalization since all UV divergences that take the form of powers of the cut-off, in cut-off regularization, are cancelled.

Dangerous extrapolation. Let us also point out one dangerous consequence. The result (10.9) extrapolated naively to $d \rightarrow 2n$ would lead to a cancellation between a UV and IR divergence:

$$\int \frac{d^d p}{p^{2n}} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \left[\int_1^\infty p^{d-1-2n} dp + \int_0^1 p^{d-1-2n} dp \right].$$

For example, in a field theory involving massless fields having a propagator $1/p^2$, IR divergences appear in 2 dimensions. If this theory is renormalizable in 2 dimensions, it also has UV divergences. In such a case, UV and IR singularities get mixed. Therefore, to be able to identify poles coming from the large momentum region, it is necessary to introduce an IR cut-off, for example, by giving a mass to the field.

Operator commutation in local products. Dimensional regularization leads to the commutation of quantum operators in local field theories. Indeed commutators, for example, in the case of a scalar field take the form (6.7),

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\hbar \delta^{(d-1)}(x - y) = i\hbar (2\pi)^{1-d} \int d^{d-1}p e^{ip \cdot (x-y)},$$

where $\hat{\pi}(x)$ is the momentum conjugate to the field $\hat{\phi}(x)$. As we have already stressed, in a local theory all fields are taken at the same point and, therefore a commutation in the product $\hat{\phi}(x)\hat{\pi}(x)$ generates a divergent contribution (for $d > 1$), formally proportional to

$$\delta^{d-1}(0) = (2\pi)^{1-d} \int d^{d-1}p.$$

In dimensional regularization, taking the limit $n = 0$ of the equation (10.9), one concludes $\int d^d p = 0$, in contrast with momentum regularization, where it is proportional to Λ^d . Therefore, the order between quantum operators becomes irrelevant, because the commutator vanishes. Dimensional regularization is thus especially useful for perturbative calculations in geometric models where the problems of quantization occur, like non-linear σ -models whose Hamiltonians have the generic form (3.24) (see Chapters 19 and 29), or gauge theories. However, the cancellations have then to be justified by other methods, like lattice regularization.

Continuation of tensor structures. So far, we have only considered diagrams corresponding to scalar fields. Any diagram which is not a scalar can be expanded on a set of fixed tensors with scalar coefficients. For example,

$$\int d^d q q_\mu q_\nu f(q^2, p^2, p \cdot q) = A(p^2) p_\mu p_\nu + B(p^2) \delta_{\mu\nu}. \quad (10.10)$$

The scalar diagrams contributing to $A(p^2)$ and $B(p^2)$ can be obtained by taking the trace and the scalar product with p_μ :

$$\begin{aligned} A(p^2) &= \frac{1}{d-1} \frac{1}{(p^2)^2} \int d^d q [d(p \cdot q)^2 - p^2 q^2] f(q^2, p^2, p \cdot q), \\ B(p^2) &= \frac{1}{d-1} \frac{1}{p^2} \int d^d q [-(p \cdot q)^2 + p^2 q^2] f(q^2, p^2, p \cdot q). \end{aligned} \quad (10.11)$$

We have reduced the problem to the calculation of integrals of the form (10.2) with additional factors polynomial in momenta. The integration over momenta can then also be performed to yield the continuation in d .

10.1.3 Dimensional regularization and UV divergences

When the dimension d approaches the dimension where UV divergences are expected, Feynman diagrams have poles as a function of the dimension. The singular contributions can be isolated by performing a Laurent expansion of the diagram. For example, the expression (10.6) is the value of a Feynman diagram of the massless ϕ^4 QFT, which is renormalizable for $d = 4$. The Laurent expansion is

$$I_\gamma = N_d \left[\frac{1}{4-d} + \frac{1}{2} - \frac{1}{2} \ln p^2 + O(d-4) \right].$$

As we have implicitly done previously, in general, we include the loop factor

$$N_d = \frac{\text{area of the sphere } S_{d-1}}{(2\pi)^d} = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}, \quad (10.12)$$

in the definition of the loop expansion parameter, because it is generated naturally by each loop integration.

As we have already shown in an example, powers of $\ln \Lambda$ (Λ being the cut-off) which appear in a cut-off regularization in the large Λ limit are replaced by powers of $1/(d-4)$.

For example, at loop order L , in a renormalizable theory like the ϕ^4 QFT, one expects multiple poles like $1/(d-4)^L$.

10.2 RG functions

We again discuss the example of the ϕ^4 QFT, first in the massive formalism of the Callan–Symanzik (CS) equation (see Chapter 9), and then in the case of the massless theory.

10.2.1 The massive ϕ^4 field theory

We consider the bare action,

$$\mathcal{S}(\phi_0) = \int d^d x \left[\frac{1}{2} (\nabla \phi_0(x))^2 + \frac{1}{2} m_0^2 \phi_0^2(x) + \frac{1}{4!} g_0 \phi_0^4(x) \right], \quad (10.13)$$

within the framework of *dimensional regularization*. Following the discussion of Section 9.2, we introduce a renormalized mass m , and a dimensionless renormalized coupling constant g , in such a way that the renormalized action can be written as

$$\mathcal{S}_r(\phi) = \int d^d x \left[\frac{1}{2} Z(g) (\nabla \phi(x))^2 + \frac{1}{2} m^2 Z_m(g) \phi^2(x) + \frac{1}{4!} m^{4-d} g Z_g(g) \phi^4(x) \right]. \quad (10.14)$$

With this parametrization, the renormalization constants $Z(g)$, $Z_m(g)$, and $Z_g(g)$, are dimensionless and, thus, depend on the only dimensionless parameter available, the coupling constant g . In particular, the mass is multiplicatively renormalizable [70, 71, 77].

Since the renormalized action is obtained from the bare action by the field rescaling $\phi_0 = \phi \sqrt{Z}$, the bare and renormalized parameters are related by

$$m_0^2 = m^2 Z_m(g) / Z(g), \quad (10.15)$$

$$g_0 = g m^{4-d} Z_g(g) / Z^2(g). \quad (10.16)$$

Note that, in the dimensional regularization scheme, renormalization constants depend on an additional hidden parameter, the dimension d .

From the relation (10.15), we can now calculate the RG β -function, the coefficient of the CS equation. Setting

$$g Z_g / Z^2 = G(g), \quad (10.17)$$

and differentiating equation (10.16) with respect to m at g_0 fixed, we find (equation (9.37))

$$0 = (4 - d) G(g) + \beta(g) \frac{\partial}{\partial g} G(g) \Rightarrow \beta(g) = -(4 - d) \left(\frac{d \ln G(g)}{dg} \right)^{-1}. \quad (10.18)$$

Using the notation of Section 9.2, we call $Z_2(g)$ the renormalization constant associated with the renormalization of ϕ^2 . From equations (9.38) and (9.39), we then derive

$$\eta(g) = \beta(g) \frac{d}{dg} \ln Z(g), \quad (10.19)$$

$$\eta_2(g) = \beta(g) \frac{d}{dg} \ln [Z_2(g) / Z(g)] \quad (10.20)$$

and, finally, from equations (10.15) and (9.40),

$$\frac{Z_m}{Z_2} \left[2 + \beta(g) \frac{d}{dg} \ln (Z_m/Z) \right] = \sigma(g). \quad (10.21)$$

This equation shows that Z_m/Z_2 is a finite function of g and, therefore, Z_m is not a new renormalization constant.

10.2.2 The massless theory

Dimensional regularization can also be used to define the massless ϕ^4 field theory. However, unlike the massive theory, due to small momentum (IR) divergences, the regularized theory only exists in an infinitesimal neighbourhood of the dimension 4. This problem is discussed at length in Chapters 15 and 16, devoted to critical phenomena. Then, since the massless theory is renormalizable, the MS scheme is also applicable.

The bare action can be written as

$$\mathcal{S}(\phi_0) = \int d^d x \left[\frac{1}{2} (\nabla \phi_0(x))^2 + \frac{1}{4!} g_0 \phi_0^4(x) \right]. \quad (10.22)$$

Note the absence of a bare mass term. Indeed, if the propagator is massless, no mass can be generated, because there is no dimensional parameter, besides the coupling constant: all diagrams contributing to the two-point function have a power-law behaviour given by simple dimensional considerations ($\varepsilon = 4 - d$ is infinitesimal):

$$\tilde{\Gamma}^{(2)}(p) = p^2 + \sum_{n=2} C_n(\varepsilon) g_0^n p^{2-n\varepsilon}.$$

To define a renormalized theory, it is necessary to introduce a mass scale μ , which takes care of the dimension of the ϕ^4 coupling constant. The renormalized action then takes the form

$$\mathcal{S}_r(\phi) = \int d^d x \left[\frac{1}{2} Z(g) (\nabla \phi(x))^2 + \frac{1}{4!} \mu^{4-d} g Z_g \phi^4(x) \right]. \quad (10.23)$$

To the action, correspond the relations between bare and renormalized vertex functions,

$$Z^{n/2}(g)\tilde{\Gamma}_r^{(n)}(p_i; \mu, g) = \tilde{\Gamma}^{(n)}(p_i; g_0), \quad (10.24)$$

with

$$g_0 = \mu^{4-d} g Z_g / Z^2. \quad (10.25)$$

A differentiation of equation (10.24) with respect to μ at g_0 fixed, yields RG equations that express that vertex functions depend on μ and g only through the combination $g_0(\mu, g)$:

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right] \tilde{\Gamma}_r^{(n)}(p_i; \mu, g) = 0, \quad (10.26)$$

where β and η are given by expressions formally identical to equations (10.18) and (10.19):

$$\beta(g) = -(4-d) \left(\frac{d \ln G(g)}{dg} \right)^{-1}, \quad \eta(g) = \beta(g) \frac{d}{dg} \ln Z(g). \quad (10.27)$$

As we have discussed in Section 9.10, it is then possible to define a massive theory by adding to the renormalized action a mass term of the form

$$\mathcal{S}_r(\phi, m) = \mathcal{S}_r(\phi) + \frac{1}{2} m^2 \int d^d x Z_2(g) \phi^2(x).$$

10.3 The structure of renormalization constants

The renormalizability of the ϕ^4 field theory in four dimensions implies that the renormalized vertex functions and, therefore, also the RG functions $\beta(g)$, $\eta(g)$, $\eta_2(g)$, and $\sigma(g)$ have a finite limit when the deviation $\varepsilon = 4 - d$ from the dimension 4 goes to zero. Since

$$G(g) = g + O(g^2),$$

the function $\beta(g)$ can be written as

$$\beta(g) = -\varepsilon g + \beta_2(\varepsilon)g^2 + \beta_3(\varepsilon)g^3 + \dots, \quad (10.28)$$

in which all the functions $\beta_n(\varepsilon)$ have regular Taylor series expansion at $\varepsilon = 0$:

$$\beta_n(\varepsilon) = \beta_n(0) + \varepsilon \beta'_n(0) + \dots$$

Conversely, we now determine the form of $G(g)$ from the knowledge of $\beta(g)$:

$$g \frac{G'(g)}{G(g)} = -\frac{\varepsilon g}{\beta(g)} \equiv \left[1 - \frac{1}{\varepsilon} \beta_2(\varepsilon)g - \frac{1}{\varepsilon} \beta_3(\varepsilon)g^2 - \dots \right]^{-1}.$$

Expanding the right-hand side in powers of g , we observe that, at a fixed order in g , the most singular term in ε comes from the term of order g^2 in $\beta(g)$:

$$\frac{G'(g)}{G(g)} = \frac{1}{g} + \left(\frac{\beta_2(0)}{\varepsilon} + O(1) \right) + g \left(\frac{\beta_2^2(0)}{\varepsilon^2} + O\left(\frac{1}{\varepsilon}\right) \right) + \dots$$

Integrating the expansion term by term, we find

$$G(g) = g + \sum_{n=2} g^n \left[\left(\frac{\beta_2(0)}{\varepsilon} \right)^{n-1} + \text{less singular terms} \right].$$

The coefficient $\tilde{G}_n(\varepsilon)$ of the expansion of $G(g)$ in powers of g ,

$$G(g) = g + \sum_2^\infty g^n \tilde{G}_n(\varepsilon),$$

has thus a Laurent series expansion in ε , for ε small, of the form

$$\tilde{G}_n(\varepsilon) = \frac{\beta_2^{n-1}(0)}{\varepsilon^{n-1}} + \frac{G_{n,2-n}}{\varepsilon^{n-2}} + \cdots + G_{n,0} + G_{n,1}\varepsilon + \cdots.$$

The finiteness of $\eta(g)$, and $\eta_2(g)$ leads to similar conclusions for $Z(g)$ and $Z_2(g)$, which can be written as

$$\begin{aligned} Z(g) &= 1 + \sum_1^\infty \frac{\alpha^{(n)}(g)}{\varepsilon^n} + \text{regular terms in } \varepsilon, \\ Z_2(g) &= 1 + \sum_1^\infty \frac{\alpha_2^{(n)}(g)}{\varepsilon^n} + \text{regular terms in } \varepsilon, \end{aligned}$$

with $\alpha^{(n)}(g) = O(g^{n+1})$, $\alpha_2^{(n)} = O(g^n)$.

We conclude that, at order L in the loop expansion, the divergent part of $\Gamma(\varphi)$, the generating functional of vertex functions, is a polynomial of degree L in $1/\varepsilon$.

10.4 MS scheme

Although the MS idea can be used in any regularization scheme (see, for example, equation (8.34)), it is especially useful in dimensional regularization. Renormalization constants are determined in the following way: instead of imposing renormalization conditions to divergent vertex functions, one just subtracts, at each order in the loop expansion, the singular part of Laurent expansion in ε [78, 79]. We denote by $\Gamma_L^{\text{div.}}(\varphi)$ the divergent part of the generating functional of vertex functions, renormalized up to $(L-1)$ loops:

$$\Gamma_L^{\text{div.}}(\varphi) = \sum_{\ell=1}^L \frac{\gamma_{L,\ell}}{\varepsilon^\ell}. \quad (10.29)$$

We then add, as a counter-term, $-\Gamma_L^{\text{div.}}(\varphi)$ to the action.

The $\overline{\text{MS}}$ scheme. In the calculation of low order Feynman diagrams, the factor $(N_d)^L$, where L is the number of loops of the diagram and N_d the loop factor (equation (10.12)),

$$N_d = 2/(4\pi)^{d/2} \Gamma(d/2) \Rightarrow N_4 = 1/8\pi^2,$$

is generated naturally. Therefore, it is convenient to rescale the loop expansion parameter (*e.g.* by multiplying it by the factor N_4/N_d) to suppress the factor, avoiding in this way to expand it in ε [80]. Combined with the MS scheme, this is called the modified MS, or $\overline{\text{MS}}$ scheme.

Example. We again consider the ϕ^4 QFT in the scheme of Section 10.2.1. Since the field renormalization vanishes at this order, the one-loop divergent part of $\Gamma(\varphi)$ can be inferred from the functional density at constant field φ . In the Fourier representation (see equation (7.93)),

$$\gamma_1(\varphi) = \Gamma_1(\varphi)/\text{volume} = \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} \ln(p^2 + m^2 + \frac{1}{2}gm^\varepsilon\varphi^2). \quad (10.30)$$

Setting $m^2 + \frac{1}{2}gm^\varepsilon\varphi^2 = K$, we calculate (using the result (A10.5))

$$\gamma(K) = \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} \ln(p^2 + K) \Rightarrow \gamma'(K) = \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} \frac{1}{p^2 + K} = \frac{1}{4}N_d \frac{\pi}{\sin(\pi d/2)} K^{d/2-1}.$$

Integrating over K , we obtain (see also equation (A8.22)),

$$\gamma_1(\varphi) = \frac{1}{2d}N_d \frac{\pi}{\sin(\pi d/2)} (m^2 + \frac{1}{2}gm^\varepsilon\varphi^2)^{d/2}. \quad (10.31)$$

From an expansion in φ and ε one infers, in the $\overline{\text{MS}}$ scheme,

$$\Gamma_1^{\text{div.}}(\varphi) = -\frac{N_4}{4\varepsilon} \left[m^2 g \int \varphi^2(x) d^d x + \frac{1}{4}g^2 m^\varepsilon \int \varphi^4(x) d^d x \right]. \quad (10.32)$$

RG functions at one-loop order. The functions Z , Z_m , Z_g at one-loop order then are

$$Z = 1 + O(g^2), \quad Z_g = 1 + \frac{3}{2}N_d \frac{g}{\varepsilon} + O(g^2), \quad Z_m = 1 + \frac{1}{2}N_d \frac{g}{\varepsilon} + O(g^2). \quad (10.33)$$

Directly calculating the vertex function $\langle \phi^2 \phi \phi \rangle$ at one-loop order, one notes that $Z_2 = Z_m$. Since the MS scheme eliminates any possible finite renormalization and Z_2/Z_m is finite, the relation remains true to all orders.

The RG functions are then

$$\beta(g) = -\varepsilon g + \frac{3}{2}N_d g^2 + O(g^3), \quad (10.34)$$

$$\eta(g) = O(g^2), \quad (10.35)$$

$$\eta_2(g) = -\frac{1}{2}gN_d + O(g^2). \quad (10.36)$$

10.4.1 RG functions in the MS scheme

The RG β -function. Expanding, for instance, the function (equation (10.17))

$$G(g) = g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n}, \quad G_n(g) = O(g^{n+1}), \quad (10.37)$$

one can calculate the RG function

$$\beta(g) = -\varepsilon \left[g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n} \right] \left[1 + \sum_1^\infty \frac{G'_n(g)}{\varepsilon^n} \right]^{-1}.$$

Since $G'_n(g)$ is of order g^n , one can expand the denominator as

$$\beta(g) = -\varepsilon \left[g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n} \right] \left[1 - \frac{G'_1(g)}{\varepsilon} + \frac{[G'_1(g)]^2}{\varepsilon^2} - \frac{G'_2(g)}{\varepsilon^2} + \dots \right].$$

This expression can be rewritten as

$$\beta(g) = -\varepsilon g - G_1(g) + gG'_1(g) + \sum_1^\infty \frac{b_n(g)}{\varepsilon^n}.$$

The finiteness of $\beta(g)$ then implies $b_n(g) = 0$ for all $n \geq 1$, and leads to the simple form

$$\beta(g) = -\varepsilon g + gG'_1(g) - G_1(g). \quad (10.38)$$

The functions $\beta(g)$ and $G_n(g)$, $n \geq 2$, are uniquely determined by the function $G_1(g)$, that is, the coefficients of $1/\varepsilon$ in the divergences.

The RG functions η and η_2 . General arguments show that, in the expansion of equation (10.29), the whole new L -loop information about divergences is contained in $\gamma_{L,1}(\varphi)$. All other counter-terms are determined by the counter-terms of previous orders. This implies that the RG functions $\eta(g)$, and $\eta_2(g)$ also have a simple dependence on ε . In the MS scheme, the renormalization constants Z and Z_2 have the form

$$Z(g) = 1 + \sum_1^\infty \frac{\alpha^{(n)}(g)}{\varepsilon^n}, \quad \alpha^{(n)}(g) = O(g^{n+1}), \quad (10.39)$$

$$Z_2(g) = 1 + \sum_1^\infty \frac{\alpha_2^{(n)}(g)}{\varepsilon^n}, \quad \alpha_2^{(n)}(g) = O(g^n). \quad (10.40)$$

Using relation (10.19) and the form (10.38) of $\beta(g)$, one obtains

$$\eta(g) = [-\varepsilon g + gG'_1(g) - G_1(g)] \left[\frac{1}{\varepsilon} \frac{d}{dg} \alpha^{(1)}(g) + O\left(\frac{1}{\varepsilon^2}\right) \right].$$

Since $\eta(g)$ has a finite limit, it is given by

$$\eta(g) = -g \frac{d}{dg} \alpha^{(1)}(g). \quad (10.41)$$

Similarly, for $\eta_2(g)$ we find

$$\eta_2(g) = -g \frac{d}{dg} \alpha_2^{(1)}(g). \quad (10.42)$$

Finally, the explicit dependence of the renormalization constants on ε can be obtained by calculating them from β , η , and η_2 . For example,

$$G(g) = g \exp \left(-\varepsilon \int_0^g \left[\frac{1}{-\varepsilon g' + b(g')} + \frac{1}{\varepsilon g'} \right] dg' \right), \quad (10.43)$$

in which we have set

$$\beta(g, \varepsilon) = -\varepsilon g + b(g). \quad (10.44)$$

Massive and massless scheme. Since, in the MS scheme the renormalization constants are uniquely defined, one concludes that the renormalization constants in the massless scheme of Section 10.2.2 and the massive scheme of Section 10.2.1 are identical.

10.5 RG functions at two-loop order: The ϕ^4 QFT

As an exercise, we now calculate explicitly, at two-loop order, the renormalization constants and RG functions in the QFT corresponding to the renormalized action (10.23), to which a source $\tau(x)$ for the $\phi^2(x)/2$ monomial is added,

$$\mathcal{S}_r(\phi) = \int d^d x \left[\frac{1}{2} Z (\nabla \phi(x))^2 + \frac{1}{4!} \mu^\varepsilon g Z_g \phi^4(x) + \frac{1}{2} Z_2 \tau(x) \phi^2(x) \right]. \quad (10.45)$$

We recall that, in dimensional regularization, no mass counter-term is generated in the massless theory.

We define the renormalization constants in the $\overline{\text{MS}}$ scheme (Section 10.4) where the loop factor (10.12) is factorized in each loop integral and not expanded in ε .

For technical convenience, we work with the massless theory. Note that for higher order calculations more sophisticated methods are used. For example, one considers diagrams to which all divergent subdiagrams have been subtracted. The remaining global divergence is then independent of external momenta and internal masses. To calculate the divergent part, one sets as many masses and momenta to 0 as possible, consistently in the diagrams and the subtracted subdiagrams, as long as one does not encounter IR (zero momentum) divergences [81]. An example is given below.

10.5.1 The perturbative expansion

In Section 7.4, we have already listed the diagrams contributing to $\Gamma^{(2)}$ at this order. We display them again in Fig. 10.1. The two first diagrams vanish in dimensional regularization and the renormalized two-point vertex function reduces to

$$\tilde{\Gamma}_r^{(2)}(p) = Z p^2 - \frac{1}{6} g^2 \mu^{2\varepsilon} A(p) + O(g^3), \quad (10.46)$$

where $A(p) \propto p^{2-2\varepsilon}$.

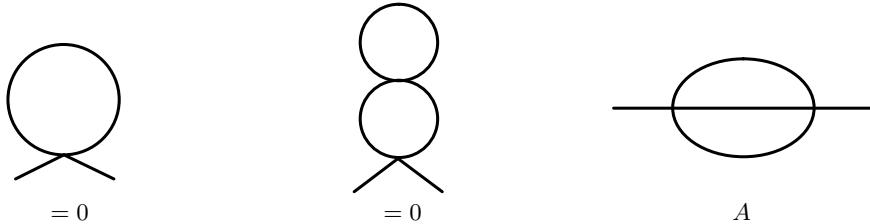


Fig. 10.1 Feynman diagrams contributing to $\Gamma^{(2)}$

The diagrams contributing to $\Gamma^{(4)}$ are displayed in Fig. 10.2. The renormalized four-point vertex function is given by (Z does not contribute at this order)

$$\begin{aligned} \tilde{\Gamma}_r^{(4)}(p_i) &= g Z_g \mu^\varepsilon - \frac{1}{2} Z_g^2 g^2 \mu^{2\varepsilon} [B(p_1 + p_2) + 2 \text{ terms}] + \frac{1}{4} g^3 \mu^{3\varepsilon} [B_d^2(p_1 + p_2) + 2 \text{ terms}] \\ &\quad + \frac{1}{2} g^3 \mu^{3\varepsilon} [C(p_1, p_2) + 5 \text{ terms}] + O(g^4), \end{aligned} \quad (10.47)$$

the additional terms restoring the permutation symmetry of the four-point function between the four momenta (p_1, p_2, p_3, p_4) .

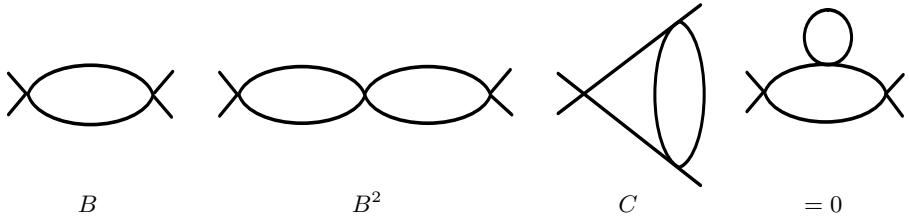


Fig. 10.2 Feynman diagrams contributing to $\Gamma^{(4)}$

10.5.2 Diagrams: Divergences

First, we list a few identities useful for the calculation of renormalization constants.

Feynman's parametrization. We use the special case of identity (A10.4) several times,

$$\frac{1}{a^\alpha b^\beta} = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 ds \frac{s^{\alpha-1}(1-s)^{\beta-1}}{(as + b(1-s))^{\alpha+\beta}}. \quad (10.48)$$

We also need the integral (A10.5),

$$\frac{1}{(2\pi)^d} \int \frac{d^d q}{(q^2 + 1)^\nu} = \frac{N_d}{2} \frac{\Gamma(d/2)\Gamma(\nu - d/2)}{\Gamma(\nu)}. \quad (10.49)$$

The massless propagator and diagram A. The massless propagator

$$\Delta(x) = \frac{1}{(2\pi)^d} \int d^d p \frac{e^{ipx}}{p^2},$$

is given by equation (10.5),

$$\Delta(x) = \frac{\Gamma(d/2 - 1)}{4\pi^{d/2}} |x|^{2-d}. \quad (10.50)$$

The diagram A (Fig. 10.1) can be expressed in terms of the propagator as

$$\begin{aligned} A(p) &= \int d^d x e^{ipx} \Delta^3(x) = \frac{\Gamma^3(d/2 - 1)}{64\pi^{3d/2}} \int d^d x e^{ipx} |x|^{6-3d} \\ &= \frac{N_d^2}{4} \Gamma^2(d/2) \Gamma^3(d/2 - 1) \frac{\Gamma(3-d)}{\Gamma(3d/2 - 3)} p^{2d-6} \underset{\varepsilon \rightarrow 0}{\sim} -\frac{N_d^2}{8\varepsilon} p^2. \end{aligned} \quad (10.51)$$

The bubble diagram B. The diagram B of Fig. 10.2 is given by

$$B(p) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(q^2(p+q)^2)}. \quad (10.52)$$

Using the identity (10.48), one can rewrite the diagram as

$$B(p) = \frac{1}{(2\pi)^d} \int_0^1 ds \int \frac{d^d q}{[(1-s)q^2 + s(p+q)^2]^2} = K_d p^{d-4},$$

with, for $\varepsilon = 4 - d \rightarrow 0$,

$$K_d = -\frac{N_d}{2} \frac{\pi}{\sin(\pi d/2)} \frac{\Gamma(d/2 - 1)\Gamma(d/2)}{\Gamma(d-2)} = \frac{N_d}{\varepsilon} (1 + \frac{1}{2}\varepsilon) + O(\varepsilon). \quad (10.53)$$

At the two-loop order, one needs,

$$B(p) = \frac{N_d}{\varepsilon} \left(1 + \left(\frac{1}{2} - \ln p \right) \varepsilon + O(\varepsilon^2) \right) = b_1 + B_r(p), \quad (10.54)$$

where $B_r(p)$ is the finite part and b_1 the divergent part of $B(p)$ in the $\overline{\text{MS}}$ scheme,

$$b_1 = \frac{N_d}{\varepsilon}. \quad (10.55)$$

The diagram C. More work is required for the evaluation of the divergent part of the third diagram of Fig. 10.2,

$$C(p_1, p_2) = \frac{1}{(2\pi)^{2d}} \int \frac{d^d q d^d k}{(q + p_1)^2 (q - p_2)^2 (q - k)^2 k^2}.$$

However, if we subtract to C the contribution of the divergent subdiagram, $C(p_1, p_2) - b_1 B_r(p_1 + p_2)$, the divergent part becomes momentum independent. Therefore, we can specialize the calculation of $C(p_1, p_2)$ to $C(p, 0)$, which is not IR divergent [81]. Then, after integration over k ,

$$C(p, 0) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(p + q)^2 q^2} B(q).$$

We set

$$C(p, 0) = K_d I(p),$$

where

$$I(p) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(p + q)^2 (q^2)^{3-d/2}}.$$

Using the identity (10.48) again, we can rewrite the integral as

$$\begin{aligned} I(p) &= \frac{1}{(2\pi)^d} \int_0^1 ds s^{2-d/2} \int \frac{d^d q}{[(1-s)q^2 + s(p+q)^2]^{4-d/2}} \\ &= \frac{N_d}{2} \Gamma(d/2) \frac{\Gamma(4-d)}{\Gamma(3-d/2)} \int_0^1 ds (1-s)^{d/2-4} s^{d-4} p^{2d-8} \\ &= \frac{N_d}{2} \Gamma(d/2) \frac{\Gamma(4-d)}{\Gamma(3-d/2)} \frac{\Gamma(d/2-1)\Gamma(d-3)}{\Gamma(3d/2-4)} p^{2d-8} \\ &= \frac{N_d}{2\varepsilon} (1 + \varepsilon - 2\varepsilon \ln p). \end{aligned}$$

Thus,

$$C(p, 0) = \frac{N_d^2}{2\varepsilon^2} (1 + \frac{3}{2}\varepsilon - 2\varepsilon \ln p). \quad (10.56)$$

Subtracting the divergent subdiagram, one obtains

$$C(p, 0) - b_1 B_r(p) = \frac{N_d^2}{2\varepsilon^2} (1 + \frac{1}{2}\varepsilon) + O(1), \quad (10.57)$$

which is momentum independent, as expected.

10.5.3 The renormalization constants

Field renormalization. At this order, only the diagram A contributes to $\tilde{\Gamma}^{(2)}$. Therefore, expressing that $\tilde{\Gamma}^{(2)}$ is finite and using equation (10.51), one finds

$$Z = 1 - \frac{N_d^2}{48\varepsilon} g^2 + O(g^3). \quad (10.58)$$

Coupling renormalization. After setting $\mu = 1$ and $B(p) = b_1 + B_r(p)$, where $B_r(p)$ is finite (equation (10.54)), the expansion (10.47) becomes

$$\begin{aligned} \tilde{\Gamma}_r^{(4)}(p_i) &= gZ_g - \frac{3}{2}b_1g^2Z_g^2 - \frac{1}{2}Z_g^2g^2 [B_r(p_1 + p_2) + 2 \text{ terms}] \\ &\quad + \frac{1}{4}g^3 [B_r^2(p_1 + p_2) + 2 \text{ terms}] + \frac{1}{2}b_1g^3 [B_r(p_1 + p_2) + 2 \text{ terms}] + \frac{3}{4}b_1^2g^3 \\ &\quad + \frac{1}{2}g^3 [C(p_1, p_2) + 5 \text{ terms}]. \end{aligned}$$

Then, using equation (10.55) and expressing that $\Gamma_r^{(4)}$ is finite at one-loop order, one finds

$$Z_g = 1 + \frac{3}{2}b_1g + O(g^2) = 1 + \frac{3N_d}{2\varepsilon}g + O(g^2).$$

The perturbative expansion becomes

$$\begin{aligned} \tilde{\Gamma}_r^{(4)}(p_i) &= gZ_g - \frac{3}{2}b_1g - \frac{15}{4}b_1^2g^3 - \frac{1}{2}g^2 [B_r(p_1 + p_2) + 2 \text{ terms}] \\ &\quad + \frac{1}{4}g^3 [B_r^2(p_1 + p_2) + 2 \text{ terms}] + \frac{1}{2}g^3 [C(p_1, p_2) - b_1B_r(p_1 + p_2) + 5 \text{ terms}]. \end{aligned}$$

The divergent part of $[C(p_1, p_2) - b_1B_r(p_1 + p_2)]$ is momentum independent and given by equation (10.57). Then, the divergent part of $\tilde{\Gamma}^{(4)}$ reduces to

$$\tilde{\Gamma}_{\text{div.}}^{(4)}(p_i) = g(Z_g - 1) - \frac{3}{2}b_1g^2 - \frac{15}{4}b_1^2g^3 + \frac{3N_d^2}{2\varepsilon^2}(1 + \frac{1}{2}\varepsilon)g^3.$$

Expressing that $\tilde{\Gamma}_{\text{div.}}^{(4)}$ vanishes at two-loop order, one finds

$$Z_g = 1 + N_d \frac{3g}{2\varepsilon} + N_d^2 \left(\frac{9}{4\varepsilon^2} - \frac{3}{4\varepsilon} \right) g^2 + O(g^3). \quad (10.59)$$

10.5.4 The ϕ^2 insertion

The two-loop expansion of the renormalized $\frac{1}{2}\phi^2\phi\phi$ vertex function is given by

$$\tilde{\Gamma}_r^{(2,1)}(p_1, p_2) = Z_2 - \frac{1}{2}gZ_gZ_2B(p_1 + p_2) + \frac{1}{4}g^2B^2(p_1 + p_2) + \frac{1}{2}g^2C(p_1, p_2) + O(g^3).$$

At this order, the field renormalization is not yet needed. We can specialize again to $p_1 = p$, $p_2 = 0$. Introducing the function B_r , we can rewrite the expressions as,

$$\begin{aligned} \tilde{\Gamma}_r^{(2,1)}(p, 0) &= Z_2 - \frac{1}{2}gZ_2B_r(p) - \frac{1}{2}b_1Z_2g - \frac{3}{4}b_1^2g^2 - \frac{3}{4}b_1B_r(p)g^2 \\ &\quad + \frac{1}{4}g^2(B_r^2(p) + 2b_1B_r(p) + b_1^2) + \frac{1}{2}g^2C(p, 0) + O(g^3). \end{aligned}$$

At one-loop order,

$$Z_2 = 1 + \frac{1}{2}b_1g = 1 + \frac{N_d}{2\varepsilon}g.$$

Then, the divergent part of $\tilde{\Gamma}^{(2,1)}$ is given by

$$\left[\tilde{\Gamma}^{(2,1)}(p, 0) \right]_{\text{div.}} = Z_2 - \frac{1}{2} b_1 g - \frac{3}{4} b_1^2 g^2 + \frac{1}{2} g^2 [C(p, 0) - b_1 B_r(p)] + O(g^3).$$

Expressing that $[\tilde{\Gamma}^{(2,1)}]_{\text{div.}}$ vanishes at two-loop order, one obtains

$$Z_2 - 1 - \frac{1}{2} b_1 g - \frac{3}{4} b_1^2 g^2 + g^2 \frac{N_d^2}{4\varepsilon^2} (1 + \frac{1}{2}\varepsilon) = 0,$$

and thus,

$$Z_2 = 1 + N_d \frac{g}{2\varepsilon} + N_d^2 \left(\frac{1}{2\varepsilon^2} - \frac{1}{8\varepsilon} \right) g^2 + O(g^3). \quad (10.60)$$

10.5.5 RG functions

Equations (10.18–10.20) then yield the three RG functions,

$$\tilde{\beta}(\tilde{g}) = N_d \beta(\tilde{g}) = -\varepsilon \tilde{g} + \frac{3}{2} \tilde{g}^2 - \frac{17}{12} \tilde{g}^3 + O(\tilde{g}^4), \quad (10.61)$$

$$\eta(\tilde{g}) = \frac{\tilde{g}^2}{24} + O(\tilde{g}^3), \quad (10.62)$$

$$\eta_2(\tilde{g}) = -\frac{\tilde{g}}{2} + \frac{5\tilde{g}^2}{24} + O(\tilde{g}^3), \quad (10.63)$$

with the notation

$$\tilde{g} = N_d g. \quad (10.64)$$

Field renormalization: Positivity of the RG function. Note that for g small, that is, in the perturbative domain, the field renormalization (10.58) satisfies $Z < 1$. This is a general property in unitary theories implied by the spectral representation of the two-point function (see Section 6.6). This implies $\eta(g) > 0$ for g small.

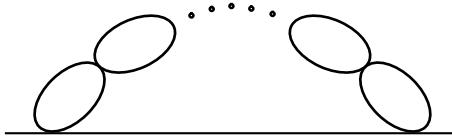


Fig. 10.3 Diagrams contributing to $\Gamma^{(2)}$

10.5.6 The massless theory at fixed dimension $d < 4$

We now show, by working out an example, that the perturbative expansion of the massless ϕ^4 QFT is IR divergent for all dimensions $d < 4$. We consider one contribution to the two-point vertex function of order g^{n+1} , proportional to the diagram (Fig. 10.3),

$$D_n(p) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(p+q)^2} B^n(q), \quad (10.65)$$

where $B(q)$ is the diagram (10.52),

$$B(q) = \frac{1}{(2\pi)^d} \int \frac{d^d k}{q^2(k+q)^2} = -\frac{N_d}{2} \frac{\pi}{\sin(\pi d/2)} \frac{\Gamma(d/2-1)\Gamma(d/2)}{\Gamma(d-2)} q^{d-4}.$$

Therefore, for

$$n(4-d) \geq d \Rightarrow n \geq d/(4-d),$$

the integral (10.65) is IR divergent. For $d = 3$, this happens for $n = 3$.

10.6 Generalization to N -component fields

We now generalize the preceding calculations to a field ϕ with N components ϕ_i and a ϕ^4 -like QFT, symmetric under a group, subgroup of $O(N)$, which admits only one quadratic invariant. In particular, this implies

$$\tilde{\Gamma}_{ij}^{(2)}(p) = \delta_{ij} \tilde{\Gamma}^{(2)}(p).$$

The renormalized action of the massless theory can be written as

$$\mathcal{S}_r(\phi) = \int d^d x \left[\frac{1}{2} Z \sum_i \nabla \phi_i(x) \nabla \phi_i(x) + \frac{1}{4!} \sum_{i,j,k,l} G_{ijkl} \phi_i(x) \phi_j(x) \phi_k(x) \phi_l(x) \right], \quad (10.66)$$

where G_{ijkl} is a tensor symmetric in its four indices. We denote by μ the renormalization scale. To determine the renormalization constants, we use again the $\overline{\text{MS}}$ scheme. We denote by g_{ijkl} the renormalized coupling constant, which is also symmetric in its four indices. At leading order,

$$G_{ijkl} = \mu^\varepsilon g_{ijkl} + O(g^2). \quad (10.67)$$

The symmetry implies a number of constraints on the tensors G_{ijkl} or g_{ijkl} . At two-loop order, one needs

$$\sum_k g_{ijkk} = \gamma_1 \delta_{ij}, \quad \gamma_1 = \frac{1}{N} \sum_{i,j} g_{iijj}, \quad (10.68)$$

$$\sum_{k,l,m} g_{iklm} g_{jklm} = \gamma_2 \delta_{ij}, \quad \gamma_2 = \frac{1}{N} \sum_{j,k,l,m} g_{jklm} g_{jklm}. \quad (10.69)$$

10.6.1 Renormalization constants

The calculation of renormalization constants reduces to the calculation of weight factors, the momentum dependence of the diagrams being the same as in Section 10.5.

The two-point function. At two-loop order, the renormalized two-point vertex function has the expansion

$$\tilde{\Gamma}_{ij}^{(2)}(p) = Z \delta_{ij} p^2 - \frac{1}{6} \sum_{k,l,m} g_{iklm} g_{jklm} \mu^{2\varepsilon} A(p) + O(g^3). \quad (10.70)$$

The field renormalization is then (equations (10.69)),

$$Z = 1 - \frac{N_d^2}{48\varepsilon} \gamma_2 + O(g^3). \quad (10.71)$$

The four-point function. The renormalized four-point vertex function $\langle \phi_i \phi_j \phi_k \phi_l \rangle_{\text{1PI}}$, at the same order is

$$\begin{aligned} \tilde{\Gamma}_{ijkl}^{(4)}(p_i) &= G_{ijkl} - \frac{1}{2} \sum_{m,n} (G_{ijmn} G_{mnkl} B(p_1 + p_2) + 2 \text{ terms}) \\ &\quad + \frac{1}{4} \sum_{m,n,p,q} (G_{ijmn} G_{mnpq} G_{pqkl} B^2(p_1 + p_2) + 2 \text{ terms}) \\ &\quad + \frac{1}{2} \sum_{m,n,p,q} (G_{klmn} G_{mpqi} G_{npqj} C(p_1, p_2) + 5 \text{ terms}) + O(G^4). \end{aligned} \quad (10.72)$$

The additional terms restore the permutation symmetry of the four-point function in its four arguments.

To evaluate renormalization constants, for notational convenience we now set $\mu = 1$. At one-loop order, the cancellation of divergence implies

$$G_{ijkl} = g_{ijkl} + \frac{1}{2} b_1 \sum_{m,n} (g_{ijmn} g_{mnkl} + 2 \text{ terms}) + O(g^2). \quad (10.73)$$

Then, expanding expression (10.72) in powers of g_{ijkl} , using the evaluations of the functions B (equation (10.54)) and C (equation (10.57)), and demanding the cancellation of the divergent part, one obtains

$$\begin{aligned} G_{ijkl} &= g_{ijkl} + \frac{N_d}{2\varepsilon} \sum_{m,n} (g_{ijmn} g_{mnkl} + 2 \text{ terms}) + \frac{N_d^2}{4\varepsilon^2} \sum_{m,n,p,q} (g_{ijmn} g_{mnpq} g_{pqkl} + 2 \text{ terms}) \\ &\quad + \frac{N_d^2}{4\varepsilon^2} \left(1 - \frac{\varepsilon}{2}\right) \sum_{m,n,p,q} (g_{ijmn} g_{mpqk} g_{npql} + 5 \text{ terms}) + O(g^4). \end{aligned}$$

The expansion of the bare coupling constant

$$g_{0;ijkl} = \mu^{-\varepsilon} Z^{-2} G_{ijkl}, \quad (10.74)$$

follows (equation (10.71)),

$$\begin{aligned} g_{0;ijkl} &= g_{ijkl} + \frac{N_d}{2\varepsilon} \sum_{m,n} (g_{ijmn} g_{mnkl} + 2 \text{ terms}) + \frac{N_d^2}{4\varepsilon^2} \sum_{m,n,p,q} (g_{ijmn} g_{mnpq} g_{pqkl} + 2 \text{ terms}) \\ &\quad + \frac{N_d^2}{4\varepsilon^2} \left(1 - \frac{\varepsilon}{2}\right) \sum_{m,n,p,q} (g_{ijmn} g_{mpqk} g_{npql} + 5 \text{ terms}) + \frac{N_d^2}{24\varepsilon} \gamma_2 + O(g^4). \quad (10.75) \end{aligned}$$

The $\frac{1}{2}\phi^2$ insertion. To determine the renormalization constant Z_2 of the local polynomial $\frac{1}{2}\phi^2(x)$, we calculate the renormalized vertex function $\frac{1}{2}\langle \tilde{\phi}^2(p_1+p_2) \tilde{\phi}_i(p_1) \tilde{\phi}_j(p_2) \rangle_{\text{1PI}}$, which has the form

$$\tilde{\Gamma}_{ij}^{(1,2)}(p_1, p_2) = \delta_{ij} \tilde{\Gamma}^{(1,2)}(p_1, p_2).$$

After introduction of the ϕ^2 renormalization constant Z_2 , the renormalized function $\tilde{\Gamma}_r^{(1,2)}$ has the two-loop expansion,

$$\begin{aligned} \tilde{\Gamma}_{r;ij}^{(1,2)}(p_1, p_2) &= Z_2 \delta_{ij} - \frac{1}{2} Z_2 \sum_k G_{kkij} B(p_1 + p_2) + \frac{1}{4} \sum_k G_{kkmn} G_{mni j} B^2(p_1 + p_2) \\ &\quad + \frac{1}{2} \sum_{k,p,q} G_{kpqi} G_{kpqj} C(p_1, p_2) + O(G^3). \end{aligned}$$

From equation (10.73), one infers,

$$\sum_k G_{kkij} = \gamma_1 + \frac{1}{2} b_1 (\gamma_1^2 + 2\gamma_2) + O(g^3).$$

It follows that

$$\begin{aligned}\tilde{\Gamma}_r^{(1,2)}(p_1, p_2) &= Z_2 - \frac{1}{2}Z_2 [\gamma_1 + \frac{1}{2}b_1(\gamma_1^2 + \gamma_2)] B(p_1 + p_2) + \frac{1}{4}\gamma_1^2 B^2(p_1 + p_2) \\ &\quad + \frac{1}{2}\gamma_2 C(p_1, p_2) + O(g^3).\end{aligned}\quad (10.76)$$

Then, setting $p_1 = p$, $p_2 = 0$, substituting $B = B_r + b_1$, one obtains

$$\begin{aligned}\tilde{\Gamma}_r^{(1,2)}(p, 0) &= Z_2 - \frac{1}{2}Z_2 [\gamma_1 + \frac{1}{2}b_1(\gamma_1^2 + 2\gamma_2)] b_1 - \frac{1}{2}Z_2 [\gamma_1 + \frac{1}{2}b_1(\gamma_1^2 + 2\gamma_2)] B_r(p) \\ &\quad + \frac{1}{4}b_1^2\gamma_1^2 + \frac{1}{2}b_1B_r(p)\gamma_1^2 + \frac{1}{4}\gamma_1^2B_r^2(p) + \frac{1}{2}\gamma_2 C(p, 0) + O(g^3).\end{aligned}$$

The renormalization constant at one-loop order is then

$$Z_2 = 1 + \frac{1}{2}b_1\gamma_1 + O(g^2).$$

The replacement, in the one-loop term, of Z_2 by the expansion yields

$$\begin{aligned}\tilde{\Gamma}_r^{(1,2)}(p, 0) &= Z_2 - \frac{1}{2}[\gamma_1 + b_1(\gamma_1^2 + \gamma_2)] b_1 - \frac{1}{2}[\gamma_1 + b_1(\gamma_1^2 + \gamma_2)] B_r(p) \\ &\quad + \frac{1}{4}b_1^2\gamma_1^2 + \frac{1}{2}b_1B_r(p)\gamma_1^2 + \frac{1}{4}\gamma_1^2B_r^2(p) + \frac{1}{2}\gamma_2 C(p, 0) + O(g^3).\end{aligned}$$

The divergent part of this expression takes the form,

$$\left[\tilde{\Gamma}_r^{(1,2)} \right]_{\text{div.}} = Z_2 - 1 - \frac{1}{2}b_1\gamma_1 - \frac{1}{4}(\gamma_1^2 + 2\gamma_2)b_1^2 + \frac{1}{2}\gamma_2 [C(p, 0) - b_1B_r(p)]_{\text{div.}}.$$

Using the result (10.57), one infers

$$Z_2 = 1 + \frac{N_d}{2\varepsilon}\gamma_1 + \frac{N_d^2}{4\varepsilon^2}\gamma_1^2 + \frac{N_d^2}{4\varepsilon^2}\left(1 - \frac{\varepsilon}{2}\right)\gamma_2 + O(g^3). \quad (10.77)$$

Actually, we need the renormalization constant which expresses the renormalized operator in terms of the bare fields,

$$\zeta_2 = Z_2/Z. \quad (10.78)$$

At two-loop order, one finds

$$\zeta_2 = Z_2 + \frac{N_d^2}{48\varepsilon}\gamma_2 + O(g^3). \quad (10.79)$$

10.6.2 RG equations

The relation between renormalized and bare vertex functions takes the form

$$\tilde{\Gamma}_{r;i_1 i_2 \dots i_n}^{(n)}(p, g, \mu) = Z^{n/2} \tilde{\Gamma}_{i_1 i_2 \dots i_n}^{(n)}(p, g_0, \Lambda), \quad (10.80)$$

in which g stands for g_{ijkl} and g_0 for $\mu^\varepsilon g_{0;ijkl}$.

We set

$$D \equiv \mu \frac{\partial}{\partial \mu} + \sum_{i,j,k,l} \beta_{ijkl}(g) \frac{\partial}{\partial g_{ijkl}},$$

with the definition

$$\sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial g_{0;ijkl}}{\partial g_{i'j'k'l'}} = -\varepsilon g_{0;ijkl}. \quad (10.81)$$

Differentiating equation (10.80) with respect to μ at g_0 and Λ fixed, one obtains the RG equation

$$(D - \frac{1}{2}n\eta(g)) \tilde{\Gamma}_{r;i_1 i_2 \dots i_n}^{(n)} = 0, \quad (10.82)$$

with the definition

$$\eta(g) = \sum_{i,j,k,l} \beta_{ijkl}(g) \frac{\partial \ln Z}{\partial g_{ijkl}}. \quad (10.83)$$

Similarly, to renormalize vertex functions with $\frac{1}{2}\phi^2(x)$ insertions, we must multiply each insertion by the matrix ζ_2 . This leads to the RG equation

$$[D - l\eta_2(g) - \frac{1}{2}n\eta(g)] \tilde{\Gamma}_{r;j_1 j_2 \dots j_l, i_1 \dots i_n}^{(l,n)} = 0, \quad (10.84)$$

with the definition

$$\eta_2(g) = \sum_{i,j,k,l} \beta_{ijkl}(g) \frac{\partial \ln \zeta_2}{\partial g_{ijkl}}. \quad (10.85)$$

RG functions. From the expressions (10.71–10.79), one derives the expansion of the RG functions at two-loop order:

$$\begin{aligned} \beta_{ijkl} &= -\varepsilon g_{ijkl} + \frac{N_d}{2} \sum_{m,n} (g_{ijmn}g_{mnkl} + 2 \text{ terms}) \\ &\quad - \frac{N_d^2}{4} \sum_{m,n,p,q} (g_{ijmn}g_{mpqk}g_{npql} + 5 \text{ terms}) \\ &\quad + \frac{N_d^2}{48} \sum_{m,n,p,q} (g_{ijkm}g_{mnpq}g_{npql} + 3 \text{ terms}) + O(g^4), \end{aligned} \quad (10.86)$$

and

$$\eta = \frac{N_d^2}{24} \gamma_2 + O(g^3), \quad \eta_2 = -\frac{N_d}{2} \gamma_1 + \frac{5N_d^2}{24} \gamma_2 + O(g^3). \quad (10.87)$$

10.6.3 $O(N)$ symmetry: Fixed point and exponents

These expressions will be used in Section 16.6 in a rather general form. Here, we specialize to the $(\phi^2)^2$ QFT with $O(N)$ symmetry. We then have to substitute

$$g_{ijkl} = \frac{g}{3} (\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \quad (10.88)$$

The two quantities γ_1 , γ_2 become

$$\gamma_1 = \frac{N+2}{3}g, \quad \gamma_2 = \frac{(N+2)}{3}g^2.$$

A short calculation in the notation of equation (10.64) leads to

$$\tilde{\beta}(\tilde{g}) = N_d \beta(\tilde{g}) = -\varepsilon \tilde{g} + \frac{1}{6}(N+8)\tilde{g}^2 - \frac{(3N+14)}{12}\tilde{g}^3 + O(\tilde{g}^4), \quad (10.89)$$

$$\eta(\tilde{g}) = \frac{(N+2)}{72}\tilde{g}^2 + O(\tilde{g}^3). \quad (10.90)$$

$$\eta_2(\tilde{g}) = -\frac{1}{6}(N+2)\tilde{g} \left(1 - \frac{5}{12}\tilde{g}\right) + O(\tilde{g}^3). \quad (10.91)$$

One notes that, for $\varepsilon = 4 - d$ positive and small, the β -function has a zero, $\tilde{g}^* \sim 6\varepsilon/(N+8)$, which corresponds in this framework to *Wilson–Fisher’s RG fixed point* in the theory of critical phenomena [75, 133] (see Chapter 15). From $\eta(\tilde{g}^*)$ and $\eta_2(\tilde{g}^*)$ one then infers two critical exponents of the N -vector model:

$$\eta = \eta(\tilde{g}^*), \quad \nu = 1/(2 + \eta_2(\tilde{g}^*)).$$

A10 Feynman parametrization

In explicit calculations of Feynman diagrams a simple identity is often useful. One starts from

$$\frac{1}{a^\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty dt t^{\alpha-1} e^{-at}. \quad (A10.1)$$

Therefore,

$$\prod_{i=1}^n (a_i)^{-\alpha_i} = \prod_{i=1}^n (\Gamma(\alpha_i))^{-1} \int_0^\infty \left(\prod_{i=1}^n dt t_i^{\alpha_i-1} \right) \exp \left(- \sum_{i=1}^n a_i t_i \right). \quad (A10.2)$$

Then, setting

$$t_i = su_i, \quad (A10.3)$$

with

$$u_i \geq 0, \quad \sum_{i=1}^n u_i = 1,$$

one can integrate over s to obtain

$$\prod_{i=1}^n \frac{1}{a_i^{\alpha_i}} = \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_n)}{\prod_{i=1}^n \Gamma(\alpha_i)} \int_{u_i \geq 0} \prod_{i=1}^n \frac{du_i u_i^{\alpha_i-1}}{\left(\sum_{i=j}^n a_j u_j \right)^{\alpha_i}} \delta \left(\sum_i u_i - 1 \right). \quad (A10.4)$$

If the quantities a_1, \dots, a_n correspond to propagators, $a_i \equiv p_i^2 + m_i^2$, the integral over momenta can then be explicitly performed.

At one-loop order, only one integral is needed:

$$\frac{1}{(2\pi)^d} \int \frac{d^d p}{(p^2 + 1)^\nu} = \frac{1}{(2\pi)^d \Gamma(\nu)} \int_0^\infty t^{\nu-1} dt \int d^d p e^{-tp^2 - t} = \frac{\Gamma(\nu - d/2)}{(4\pi)^{d/2} \Gamma(\nu)}. \quad (A10.5)$$

Note that the representation (10.3) leads to an expression similar to (A10.4). When $\rho(t)$ is of the form e^{-tm^2} , the argument of the exponential is linear in the variables t_i and, after the change of variables (A10.3), the integral over the homogeneous variable s can be performed.

In the calculations we have also used the integral

$$B(\alpha, \beta) \equiv \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \quad (A10.6)$$

11 Renormalization of local polynomials. Short-distance expansion (SDE)

In this chapter, we discuss two related topics: the renormalization of local polynomials (or local operators) of the field [82], and the SDE of the product of local operators, in space dimension 4 for simplicity. Both problems are related, since one can consider the insertion of a product of operators $A(x)B(y)$ as a regularization by point splitting of the operator $A[\frac{1}{2}(x+y)]B[\frac{1}{2}(x+y)]$. Therefore, in the limit $x \rightarrow y$, we expect the product to become dominated by a linear combination of the local operators which appear in the renormalization of the product AB , with singular coefficients, functions of $(x-y)$, replacing the usual cut-off dependent renormalization constants.

We first discuss the renormalization of local polynomials from the viewpoint of power counting. We use the relations between bare and renormalized operators to establish Callan–Symanzik (CS) equations for the insertion of operators of dimension 4 in the $\phi_{d=4}^4$ quantum field theory (QFT) [83]. We show that, in a QFT, there exist linear relations between operators, due to the equations of motion and relations derived in Section 7.5.

Then, starting with Section 11.3, we establish the existence of a SDE [84–87] for the product of two basic fields, and discuss the SDE at leading order in the ϕ^4 QFT.

We pass from short-distance behaviour to large-momentum behaviour, and derive a CS equation for the coefficient of the expansion at leading order [88]. Finally, we briefly discuss the generalization of this analysis to the SDE beyond leading order, to the SDE of arbitrary operators and to the light cone expansion (LCE) [89], which appears in the study of the large-momentum behaviour of real-time correlation functions (in contrast to Euclidean or imaginary time). In the whole chapter, we work in space-time dimension 4.

11.1 Renormalization of operator insertions

In Section 9.6, we showed how to renormalize insertions of the monomial $\phi^2(x)$. We could have considered other local monomials like $\phi^4(x)$, $(\nabla\phi(x))^2\dots$. They all generate new divergences which have to be cancelled by additional renormalizations.

In Section 8.6, we explained how to calculate the superficial degree of divergence of the insertion of a local monomial (or operator) $\mathcal{O}(\phi, x)$ by adding a source $g_{\mathcal{O}}(x)$ for the operator in the action:

$$\mathcal{S}(\phi) \mapsto \mathcal{S}(\phi) + \int d^4x \mathcal{O}(\phi, x) g_{\mathcal{O}}(x).$$

With this choice of sign, differentiation of $\mathcal{W}(J, g_{\mathcal{O}})$ with respect to $g_{\mathcal{O}}(x)$ generates insertions of $-\mathcal{O}(\phi, x)$ in connected correlation functions. However, since $\delta\Gamma/\delta g_{\mathcal{O}}(x) = -\delta\mathcal{W}/\delta g_{\mathcal{O}}(x)$, $\delta\Gamma/\delta g_{\mathcal{O}}(x)$ corresponds to the insertion of $\mathcal{O}(\phi, x)$ in vertex functions.

As a convention, we assign a canonical dimension $[g_{\mathcal{O}}]$ to the source $g_{\mathcal{O}}(x)$, opposite to the dimension of the vertex associated to $\mathcal{O}(\phi)$, and thus related to the dimension $[\mathcal{O}(\phi)]$ of the operator $\mathcal{O}(\phi, x)$ by

$$[g_{\mathcal{O}}] = 4 - [\mathcal{O}(\phi)].$$

Then, the implications of power counting and renormalization theory can be summarized as: the sum of counter-terms needed to render $\Gamma(\phi, g_{\mathcal{O}})$ finite is the most general local functional of $\phi(x)$ and $g_{\mathcal{O}}(x)$, allowed by power counting.

More precisely, it is *the most general linear combination of all vertices in $\phi(x)$ and $g_{\mathcal{O}}(x)$ of non-positive dimensions*.

In particular, this implies that an operator of a given dimension inserted once, in general, mixes with all operators of equal or lower dimension under renormalization. It is thus natural to study the renormalization of operators of increasing dimension.

We first verify this assertion in the case of the insertion of $\phi^2(x)$ in a ϕ^4 QFT in four dimensions, which we have already discussed in Chapter 9.

Regularization. For simplicity, in this chapter we assume an implicit momentum cut-off regularization, and denote the cut-off by Λ .

11.1.1 The $\phi^2(x)$ insertion

We use the conventions of Section 9.2 for the bare and renormalized actions:

$$\mathcal{S}(\phi) = \int d^4x \left[\frac{1}{2} (\nabla\phi(x))^2 + \frac{1}{2} r\phi^2(x) + \frac{1}{4!} g\phi^4(x) \right], \quad (11.1)$$

$$\mathcal{S}_r(\phi_r) = \int d^4x \left[\frac{1}{2} Z (\nabla\phi_r(x))^2 + \frac{1}{2} (m_r^2 + Z\delta r) \phi_r^2(x) + \frac{1}{4!} g_r Z_g \phi_r^4(x) \right]. \quad (11.2)$$

In the ϕ^4 field theory in four dimensions, the operator $\phi^2(x)$ has dimension $[\phi^2] = 2$.

Then, we denote by $t(x)$ the source for $\phi^2(x)$. Its dimension is $[t] = 2$.

Then, in addition to the vertices involving only the field ϕ , the following vertices arise as counter-terms:

$$\begin{aligned} & \int t(x)\phi^2(x)d^4x, \quad \text{which has dimension 0,} \\ & \int t^2(x)d^4x, \quad \text{which also has dimension 0,} \\ & \int t(x)d^4x, \quad \text{which has dimension -2.} \end{aligned}$$

The renormalized action then has the form (a and b are constants)

$$\mathcal{S}_r(\phi_r, t) = \mathcal{S}_r(\phi_r) + \frac{1}{2} Z_2 \int d^4x t(x)\phi_r^2(x) + \int d^4x \left[\frac{1}{2} at^2(x) + bt(x) \right]. \quad (11.3)$$

The last two terms only contribute to the vacuum amplitude. Expression (11.3) implies a set of relations between bare and renormalized generating functionals. First, for complete correlation functions,

$$\mathcal{Z}_r(J, t) = \mathcal{Z}(J/\sqrt{Z}, tZ_2/Z) \exp \left[- \int d^4x \left(\frac{1}{2} at(x)^2 + bt(x) \right) \right]. \quad (11.4)$$

For the connected functions, this implies

$$\mathcal{W}_r(J, t) = \mathcal{W}(J/\sqrt{Z}, tZ_2/Z) - \int d^4x \left(\frac{1}{2} at(x)^2 + bt(x) \right). \quad (11.5)$$

After Legendre transformation with respect to J , one obtains

$$\Gamma_r(\varphi, t) = \Gamma(\varphi\sqrt{Z}, tZ_2/Z) + \int d^4x \left(\frac{1}{2} at(x)^2 + bt(x) \right). \quad (11.6)$$

Expanding in powers of t and φ , one recovers the relations between bare and renormalized vertex functions described in Sections 9.7–9.10.

11.1.2 Operators of dimensions 3 and 4

The ϕ^3 insertion. To discuss the insertion of the $\phi^3(x)$ operator which has dimension 3, we introduce a source $t(x)$ that has dimension 1. The renormalized action $\mathcal{S}_r(\phi, t)$ then has the form (a_1, \dots, a_6 are constants)

$$\begin{aligned} \mathcal{S}_r(\phi, t) = & \mathcal{S}_r(\phi) + \int d^4x \left[t(x) \left(\frac{1}{3!} Z_3 \phi^3(x) + a_1 \phi(x) \right) + a_2 (\nabla t(x))^2 \right. \\ & \left. + t^2(x) (a_3 \phi^2(x) + a_4) + a_5 t^3(x) \phi(x) + a_6 t^4(x) \right]. \end{aligned} \quad (11.7)$$

In particular, the expression shows that the operator $\phi^3(x)$ mixes with $\phi(x)$ under renormalization, that the double insertion of ϕ^3 generates a counter-term proportional to ϕ^2 Therefore, to be able to write the equivalent of relations (11.4–11.6), we have to introduce explicitly a source for $\phi^2(x)$. We leave it to the reader to determine the renormalized action with sources for ϕ^3 and ϕ^2 , and the relations between renormalized and bare vertex functions. We postpone the discussion of the CS equations of correlation functions with ϕ^3 insertion until Section 11.2, in order to be able to incorporate the information provided by the field quantum equation of motion.

Operators of dimension 4 [83]. Quite generally, if an operator has a dimension strictly smaller than the space dimension d , its source has a strictly positive dimension, and the renormalized action is a polynomial in the source. If a source is coupled to operators of dimension d , corresponding to vertices of dimension 0 (here $\phi^4(x)$, $(\nabla\phi(x))^2$, an infinite series in the source is generated by the renormalization procedure, together with all operators of lower or equal dimensions.

For example, in $\phi_{d=4}^4$ QFT, if one inserts an operator of dimension 4 once, one has to consider the mixing of all linearly independent operators of dimensions 4 and 2 (parity in ϕ excludes odd dimensions). For example, the four operators

$$\begin{aligned} \mathcal{O}_1(\phi) &= \frac{1}{2} m_r^2 \phi^2(x), & \mathcal{O}_2(\phi) &= -\frac{1}{2} \nabla^2(\phi^2(x)), \\ \mathcal{O}_3(\phi) &= \frac{1}{2} [\nabla\phi(x)]^2, & \mathcal{O}_4(\phi) &= \frac{1}{4!} \phi^4(x), \end{aligned} \quad (11.8)$$

form a basis of linearly independent operators, which mix under renormalization. There exists another operator $\phi(x)\nabla^2\phi(x)$ of dimension 4, but it is a linear combination of \mathcal{O}_2 and \mathcal{O}_3 :

$$\frac{1}{2} \nabla^2(\phi^2(x)) = \phi(x) \nabla^2\phi(x) + [\nabla\phi(x)]^2.$$

The operator $\mathcal{O}_1(\phi)$ and, therefore, also the operator $\mathcal{O}_2(\phi)$, are multiplicatively renormalizable.

We can thus express the relation between bare and renormalized correlation functions $\Gamma_{\mathcal{O}_i}^{(n)}$ with \mathcal{O}_i insertion as

$$\left\{ \Gamma_{\mathcal{O}_i}^{(n)} \right\}_r = Z^{n/2} \sum_j Z_{ij} \Gamma_{\mathcal{O}_j}^{(n)}. \quad (11.9)$$

The renormalization matrix Z_{ij} has the form

$$\begin{pmatrix} (Z_2/Z) \mathbf{1}_2 & 0 \\ \mathbf{B} & \mathbf{A} \end{pmatrix},$$

in which \mathbf{A} and \mathbf{B} are 2×2 matrices. We have used for the renormalization of ϕ^2 the notation of Section 9.2.



Fig. 11.1 Divergent contribution to $\phi^2(x)\phi^4(y)$ insertion (dotted lines)

CS equations. From equation (11.9), we can derive CS equation for $\{\Gamma_{\mathcal{O}_i}^{(n)}\}_r$. However, here some care is required. The CS operation involves ϕ^2 insertions and the product, for example, $\{\phi^4 x\|_r \{\phi^2(y)\}_r$ inserted in a correlation function is not finite: since the source for ϕ^4 has dimension 0, and the source for ϕ^2 dimension 2, the product of ϕ^2 by both sources has dimension 4. Fig. 11.1 displays the first divergent diagram.

This implies

$$\frac{1}{4!} \left\{ \left\{ \phi^4(x) \right\}_r \left\{ \phi^2(y) \right\}_r \right\}_r = \frac{1}{4!} \left\{ \phi^4(x) \right\}_r \left\{ \phi^2(y) \right\}_r + C_4 \delta^{(4)}(x-y) \left\{ \phi^2(x) \right\}_r , \quad (11.10)$$

in which C_4 is a new renormalization constant. Identity (11.10) is only true as an insertion in an n -point correlation function for $n \neq 0$.

After Fourier transformation, and for an insertion of ϕ^2 at zero momentum, equation (11.10) becomes

$$\left\{ \left\{ \tilde{\mathcal{O}}_4(p) \right\}_r \left\{ \tilde{\mathcal{O}}_1(0) \right\}_r \right\}_r = \left\{ \tilde{\mathcal{O}}_4(p) \right\}_r \left\{ \tilde{\mathcal{O}}_1(0) \right\}_r + C_4 \left\{ \tilde{\mathcal{O}}_1(p) \right\}_r . \quad (11.11)$$

A similar equation holds for the operator \mathcal{O}_3 .

We now apply the CS operator, $m_r \partial / \partial m_r$, at g and Λ fixed, on equation (11.9).

A new set of RG functions is generated involving the matrix

$$\tilde{\eta}_{ij}(g_r, \Lambda/m_r) = \sum_k \left(m_r \frac{\partial}{\partial m_r} Z_{ik} \right) [Z^{-1}]_{kj} . \quad (11.12)$$

As a consequence of relation (11.11), two elements, $\tilde{\eta}_{31}$ and $\tilde{\eta}_{41}$, of the matrix $\tilde{\eta}_{ij}$ are not finite when the cut-off becomes infinite. Their divergent part cancels the divergences coming from the insertion of ϕ^2 as represented by equation (11.11). Defining then,

$$\eta_{i1}(g_r) = \tilde{\eta}_{i1} - m_r^2 \sigma(g_r) C_i , \quad (11.13)$$

we now obtain two finite RG functions.

For the other matrix elements, we just set

$$\eta_{ij} = \tilde{\eta}_{ij} . \quad (11.14)$$

The CS equations then read

$$\begin{aligned} & \left\{ \left[m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \eta(g_r) \right] \delta_{ij} - \eta_{ij}(g_r) \right\} \left\{ \tilde{\Gamma}_{\mathcal{O}_j}^{(n)} \right\}_r (p) \\ &= m_r^2 \sigma(g_r) \left\{ \tilde{\Gamma}_{\mathcal{O}_i}^{(1,n)} \right\}_r (0; p) . \end{aligned} \quad (11.15)$$

The matrix η_{ij} has the form

$$\eta_{ij} = \begin{bmatrix} \mathbf{c} & 0 \\ \mathbf{b} & \mathbf{a} \end{bmatrix} ,$$

in which \mathbf{a} , \mathbf{b} and \mathbf{c} are 2×2 matrices, and \mathbf{c} is diagonal:

$$\mathbf{c} = \begin{bmatrix} 2 + \eta_2 & 0 \\ 0 & \eta_2 \end{bmatrix} .$$

This completes the discussion of one insertion of the operators of dimension 4 in ϕ_4^4 field theory. It reveals the general features of the insertion of any other operator of higher dimension.

Double insertion of operators of dimension 4. Let us now briefly discuss the double $\phi^4(x)$ or $(\nabla\phi(x))^2$ insertion. It is similar to the $\phi^4\phi^2$ insertion. The relation between product of renormalized operators and renormalized product is

$$\begin{aligned} \{\{\phi^4(x)\}_r \{\phi^4(y)\}_r\}_r &= \{\phi^4(x)\}_r \{\phi^4(y)\}_r + \delta^{(4)}(x-y) \sum_{i=1}^4 D_{4i} \{\mathcal{O}_i(\phi(x))\}_r \\ &\quad + \nabla\delta^{(4)}(x-y) E_4 \nabla \mathcal{O}_1 \{\phi(x)\}_r + \nabla^2 \delta^{(4)}(x-y) F_4 \{\mathcal{O}_1(\phi(x))\}_r, \end{aligned} \quad (11.16)$$

in which D_{4i} , E_4 and F_4 are new renormalization constants. A similar equation is valid for $[\nabla\phi(x)]^2$. Again, equation (11.16) is valid only as an insertion.

11.1.3 Operator insertion: General case

Power counting arguments, based on the dimension of operators and sources, imply quite generally that if $\mathcal{O}(\phi, x)$ is an operator of canonical dimension $[\mathcal{O}(\phi)] = D$, then it renormalizes as

$$\{\mathcal{O}(\phi, x)\}_r = \sum_{\alpha: D_\alpha = [\mathcal{O}_\alpha] \leq D} Z_\alpha \mathcal{O}_\alpha(\phi, x), \quad (11.17)$$

where Z_α are renormalization constants.

If we now consider the product of two operators $\mathcal{O}(\phi)$ and $\mathcal{O}'(\phi)$ of dimensions D and D' at different points x and y , then,

$$\begin{aligned} \{\{\mathcal{O}(\phi, x)\}_r \{\mathcal{O}'(\phi, y)\}_r\}_r &= \{\mathcal{O}(\phi, x)\}_r \{\mathcal{O}'(\phi, y)\}_r \\ &\quad + \sum_{\alpha: [\mathcal{O}_\alpha] + [P_\alpha] \leq D + D' - d} C_\alpha \{\mathcal{O}_\alpha(\phi, x)\}_r P_\alpha(\nabla) \delta^{(4)}(x-y), \end{aligned} \quad (11.18)$$

in which $P_\alpha(\nabla)$ is a polynomial in ∇ . For example, in $\phi_{d=4}^4$ field theory, the product $\{\phi^6(x)\}_r \{\phi^8(y)\}_r$ involves all operators of dimension lower than or equal to 10.

11.1.4 Operator insertion and effective field theory

In Section 8.9, we have argued that to describe large scale physics, the effective action could be reduced to its renormalizable part. The argument was based on an analysis at leading order in perturbation theory; for dimensional reasons, the non-renormalizable interactions are strongly suppressed at large distance. However, beyond leading order, the insertion of these interactions generate strong divergences, which seem to cancel the suppression factor. These divergences were used in the traditional presentation of renormalization theory as an argument to exclude them.

However, we are now in a position to examine more precisely the effect of such insertions. We can invert the relations (11.17) in the form

$$\mathcal{O}(\phi, x) = \sum_{\alpha: D_\alpha \leq D} Z'_\alpha \{\mathcal{O}_\alpha(\phi, x)\}_r, \quad (11.19)$$

where $D_\alpha = [\mathcal{O}_\alpha]$ and Z'_α is another combination of renormalization constants. From power counting, we know that the behaviour of Z'_α for a large cut-off Λ is proportional to Λ^{D-D_α} , up to powers of $\ln\Lambda$.

The operator $\mathcal{O}(\phi, x)$ is multiplied by a factor Λ^{4-D} . Therefore, the contribution of the operator $\{\mathcal{O}_\alpha(\phi, x)\}_r$ is multiplied by a factor Λ^{4-D_α} up to logarithms. One infers that:

(i) the operators corresponding to non-negative powers of Λ , $D_\alpha \leq 4$, are those that are already present in the renormalizable action;

(ii) all other operators are suppressed by powers of Λ .

The conclusion is that the higher order contributions of non-renormalizable interactions renormalize the coefficients of the renormalizable action, that is, the field normalization and the ϕ^2 and ϕ^4 coefficients, and add subleading contributions, only modified by powers of logarithms compared the leading order contributions.

For example, in the $\phi_{d=4}^4$ QFT, the leading contributions, for Λ large, correspond to the operators of dimension 6, like ϕ^6 , and are suppressed by a factor $1/\Lambda^2$ up to powers of $\ln \Lambda$ to any finite order in perturbation theory. For example,

$$\frac{1}{\Lambda^2} \phi^6(x) = C_1(\Lambda) \{\phi^2(x)\}_r + C_2(\Lambda) \{\phi^4(x)\}_r + C_3(\Lambda) \left\{ (\nabla \phi(x))^2 \right\}_r + \sum_\alpha D_\alpha \{\mathcal{O}_{6,\alpha}(\phi, x)\}_r,$$

where C_1 grows like Λ^2 up to powers of logarithms, C_2, C_3 grow like powers of logarithms, the $\{\mathcal{O}_{6,\alpha}(\phi, x)\}_r$ are all operators of dimension 6, and the coefficients D_α decrease like $1/\Lambda^2$ up to logarithms, to any finite order in perturbation theory.

A more precise analysis, useful for numerical simulations of QFTs, also shows that the coefficients of the non-renormalizable interactions can be adjusted in order to cancel the leading corrections for Λ large to the renormalized functions, facilitating in this way the approach to the continuum limit, and leading to the concept of *improved action* [90].

11.2 Quantum field equations

We have discussed the renormalization of local polynomials. However, not all renormalizations are independent in a given field theory, because quantum field equations and other identities discussed in Section 7.5 imply relations between operators.

11.2.1 Insertion of the ϕ^3 operator

We return to the example of the $\phi^3(x)$ operator in the framework of the ϕ^4 field theory.

We consider the action

$$\begin{aligned} \mathcal{S}(\phi, t, u) \\ = \int d^4x \left[\frac{1}{2} (\nabla_\Lambda \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \phi^4(x) + \frac{1}{2} t(x) \phi^2(x) + \frac{1}{3!} u(x) \phi^3(x) \right]. \end{aligned} \quad (11.20)$$

(∇_Λ is a momentum regularized form of ∇ .) A differentiation of the partition function, corresponding to the action, with respect to the external fields $u(x)$ and $t(x)$ generates $-\frac{1}{3!} \phi^3(x)$ and $-\frac{1}{2} \phi^2(x)$ insertions, respectively.

The simplest relations are derived from the identity (see Section 7.5),

$$\int [d\phi] \left[\frac{\delta \mathcal{S}(\phi, t, u)}{\delta \phi(x)} - J(x) \right] \exp \left(-\mathcal{S}(\phi, t, u) + \int d^4y J(y) \phi(y) \right) = 0.$$

Since we have introduced sources for ϕ^2 and ϕ^3 , we can use them to express the terms ϕ^2 and ϕ^3 in $\delta \mathcal{S} / \delta \phi$ as functional derivatives with respect to $t(x)$ and $u(x)$. The partition function then satisfies

$$[-\nabla_\Lambda^2 + r + t(x)] \frac{\delta \mathcal{Z}}{\delta J(x)} - g \frac{\delta \mathcal{Z}}{\delta u(x)} - u(x) \frac{\delta \mathcal{Z}}{\delta t(x)} = J(x) \mathcal{Z}(J, t, u). \quad (11.21)$$

The functional $\mathcal{W} = \ln \mathcal{Z}$ then satisfies the equation

$$[-\nabla_\Lambda^2 + r + t(x)] \frac{\delta \mathcal{W}}{\delta J(x)} - g \frac{\delta \mathcal{W}}{\delta u(x)} - u(x) \frac{\delta \mathcal{W}}{\delta t(x)} = J(x). \quad (11.22)$$

The Legendre transformation is straightforward. The property (7.65) implies

$$\frac{\delta \mathcal{W}}{\delta u(x)} = -\frac{\delta \Gamma}{\delta u(x)}, \quad \frac{\delta \mathcal{W}}{\delta t(x)} = -\frac{\delta \Gamma}{\delta t(x)},$$

and one finds

$$[-\nabla_\Lambda^2 + r + t(x)] \varphi(x) + g \frac{\delta \Gamma}{\delta u(x)} + u(x) \frac{\delta \Gamma}{\delta t(x)} - \frac{\delta \Gamma}{\delta \varphi(x)} = 0. \quad (11.23)$$

Setting $t(x) = u(x) = 0$, we derive a first consequence of the relation. Since

$$\left. \frac{\delta \Gamma}{\delta u(x)} \right|_{u=t=0} = \Gamma_{\phi^3}(\varphi, x), \quad (11.24)$$

where $\Gamma_{\phi^3}(\varphi, x)$ is the generating functional of ϕ -field vertex functions with one $\frac{1}{3!}\phi^3(x)$ insertion, then

$$g \Gamma_{\phi^3}(\varphi, x) = \frac{\delta \Gamma}{\delta \varphi(x)} - (-\nabla_\Lambda^2 + r) \varphi(x). \quad (11.25)$$

The relation shows that, up to explicit subtractions affecting only the $\langle \phi^3 \phi \rangle$ vertex function, the insertion of ϕ^3 is equivalent to the insertion of ϕ itself.

The diagrammatic interpretation of equation (11.25) is simple: the insertion of ϕ^3 is indistinguishable from the addition of a ϕ^4 vertex with one of the lines attached to the vertex being an external line. However, diagrams without a ϕ^4 vertex cannot be generated, and this explains the subtractions (see Fig. 11.2).



Fig. 11.2 ϕ^3 insertion

We now introduce the generating functional of renormalized vertex functions

$$\Gamma_r(\varphi) = \Gamma(\varphi\sqrt{Z})$$

and, thus,

$$\frac{\delta \Gamma_r(\varphi)}{\delta \varphi(x)} = \sqrt{Z} \frac{\delta \Gamma(\varphi\sqrt{Z})}{\delta \varphi(x)}.$$

We insert this relation into equation (11.25),

$$(-\nabla_\Lambda^2 + r) Z \varphi(x) + g\sqrt{Z} \Gamma_{\phi^3}(\varphi\sqrt{Z}, x) = \frac{\delta \Gamma_r(\varphi)}{\delta \varphi(x)}. \quad (11.26)$$

The right-hand side is finite in the infinite cut-off limit. We now also introduce in the equation the renormalization constants of the $\phi_{d=4}^4$ QFT, as defined by equations (11.1) and (11.2), and obtain

$$[-Z\nabla^2 + m_r^2 + Z\delta r] Z\varphi(x) + g_r(Z_g/Z^{3/2})\Gamma_{\phi^3}(\varphi\sqrt{Z}, x) = \frac{\delta\Gamma_r(\varphi)}{\delta\varphi(x)}. \quad (11.27)$$

This relation shows that all ϕ -field vertex functions with one insertion of the operator $Z_g\phi_r^3(x)$ are finite, except the $\langle\phi^3\phi\rangle$ vertex function which needs two additional subtractions. We determine the corresponding renormalization constants by imposing, in the Fourier representation,

$$\begin{aligned} \left\{ \tilde{\Gamma}_{\phi^3}^{(1)} \right\}_r(p, -p) \Big|_{p=0} &= 0, \\ \frac{\partial}{\partial p^2} \left\{ \tilde{\Gamma}_{\phi^3}^{(1)} \right\}_r(p, -p) \Big|_{p=0} &= 0. \end{aligned}$$

The coefficient of degree n in φ in equation (11.27) then yields explicitly (q is the argument of ϕ^3),

$$\tilde{\Gamma}_r^{(n+1)}(q; p_1, \dots, p_n) = g_r \left\{ \tilde{\Gamma}_{\phi^3}^{(n)} \right\}_r(q; p_1, \dots, p_n) + \delta_{n1} (p^2 + m_r^2). \quad (11.28)$$

With this definition $\{\tilde{\Gamma}_{\phi^3}^{(3)}\}_r$ satisfies the renormalization condition

$$\left\{ \tilde{\Gamma}_{\phi^3}^{(3)} \right\}_r(0; 0, 0, 0) = 1. \quad (11.29)$$

Equation (11.23) also contains information about multiple insertions of ϕ^3 . For example, after some algebraic manipulations, for two insertions, one finds

$$\begin{aligned} g^2 \Gamma_{\phi^3\phi^3}(\varphi, x_1, x_2) + g\delta^{(4)}(x_1 - x_2)\Gamma_{\phi^2}(\varphi, x_1) + (-\nabla_\Lambda^2 + r)\delta^{(4)}(x_1 - x_2) \\ = \frac{\delta^2 \Gamma}{\delta\varphi(x_1)\delta\varphi(x_2)}. \end{aligned} \quad (11.30)$$

The equation relates two insertions of ϕ^3 to two insertions of ϕ , again with subtraction terms, which now involve the insertion of ϕ^2 .

11.2.2 Other relations: Renormalization of operators of dimension 4

We have shown in Section 7.5 that more general equations are obtained by performing infinitesimal changes of variables. We can use them to establish relations between operators. For example, in the change of field variables,

$$\phi(x) \mapsto \phi'(x), \text{ with } \phi'(x) = \phi(x) + \varepsilon(x)\phi(x),$$

the variation of the action (11.1) in the presence of a source is

$$\delta \left[\mathcal{S}(\phi) - \int d^4y J(y)\phi(y) \right] = \int d^4x \varepsilon(x) \left[\phi(x) (-\nabla_\Lambda^2 + r) \phi(x) + \frac{g}{3!} \phi^4(x) - J(x)\phi(x) \right].$$

We define the operator $O(x) \equiv \phi(x) (-\nabla_\Lambda^2 + r) \phi(x)$. From the invariance of the field integral, it follows that

$$\mathcal{W}_O(J, x) + \frac{g}{3!} \mathcal{W}_{\phi^4}(J, x) = J(x) \frac{\delta \mathcal{W}}{\delta J(x)}. \quad (11.31)$$

The discussion of the large cut-off limit of the equation, or the corresponding one obtained after Legendre transformation,

$$\Gamma_O(\phi, x) + \frac{g}{3!} \Gamma_{\phi^4}(\phi, x) = \varphi(x) \frac{\delta \Gamma}{\delta \varphi(x)}, \quad (11.32)$$

is more subtle than in the case of equation (11.25): the operator $O(x)$, which in Pauli–Villars regularization is

$$O(x) \equiv \phi(x) (-\nabla^2 + r + \alpha_1 \nabla^4/\Lambda^2 - \alpha_2 \nabla^6/\Lambda^3 + \dots) \phi(x),$$

contains operators of canonical dimensions larger than 4 divided by powers of the cut-off. We have discussed in Section 11.1.4 the insertion of non-renormalizable (irrelevant) operators in four dimensions in the large cut-off limit (we return to the problem in dimension $d = 4 - \varepsilon$ in Section 17.3). The conclusion is that, in the large cut-off limit, the operator $\phi(-\nabla_\Lambda^2 + r)\phi$ is equivalent, at leading order, to a linear combination of all operators of dimensions 4 and 2.

Equation (11.32) implies, after renormalization, the identity satisfied by the operators $\{\mathcal{O}_i(\phi)\}_r$ as defined in equations (11.8) and (11.9),

$$\sum_{i=1}^4 C_i(g_r) \left\{ \tilde{\Gamma}_{\mathcal{O}_i}^{(n)} \right\}_r (q; p_1, \dots, p_n) = \sum_{m=1}^n \tilde{\Gamma}_r^{(n)}(p_1, \dots, p_m + q, \dots, p_n). \quad (11.33)$$

To explicitly define the insertions of the operators \mathcal{O}_i , in the spirit of Chapter 9 we impose the renormalization conditions,

$$\begin{aligned} \left\{ \tilde{\Gamma}_{\mathcal{O}_1}^{(2)} \right\}_r (0; 0, 0) &= m_r^2, \\ \left\{ \tilde{\Gamma}_{\mathcal{O}_3}^{(2)} \right\}_r (q; p_1, p_2) &= -p_1 \cdot p_2 + O(p^4), \\ \left\{ \tilde{\Gamma}_{\mathcal{O}_3}^{(4)} \right\}_r (0; 0, 0, 0, 0) &= 0, \\ \left\{ \tilde{\Gamma}_{\mathcal{O}_4}^{(2)} \right\}_r (q; p_1, p_2) &= O(p^4), \\ \left\{ \tilde{\Gamma}_{\mathcal{O}_4}^{(4)} \right\}_r (0; 0, 0, 0, 0) &= 1. \end{aligned} \quad (11.34)$$

We can then calculate the coefficients $C_i(g_r)$ only from $\{\Gamma_{\mathcal{O}_1}^{(n)}\}_r$ and its derivatives at zero momentum.

11.3 Short-distance expansion of operator products

Several chapters are devoted to the discussion of the large or small momentum behaviour of correlation functions. Our essential tools are the CS or RG equations. However, these equations lead to a direct characterization of the behaviour of

$$\tilde{\Gamma}^{(n)}(\rho p_1 + r_1, \rho p_2 + r_2, \dots, \rho p_n + r_n),$$

for $\rho \rightarrow \infty$, only for *non-exceptional momenta*, that is, provided all, not empty strictly subsets of momenta p_i have a non-vanishing sum. If we denote by I the indices of momenta in such a subset, the condition can be written as

$$\sum_{i \in I \neq \emptyset} p_i \neq 0, \quad \forall I \not\equiv \{1, 2, \dots, n\}.$$

For exceptional momenta, a new tool is required: the SDE of product of operators [84, 85, 87]. In this section, we mainly discuss the SDE of the product of two fields at leading order, but we shall indicate how the method can be generalized to more complicated examples. It is entirely based on the theory of renormalization of local monomials we have just described.

Definition. We consider the vertex function

$$\Gamma^{(n+2)}(x+v/2, x-v/2, y_1, \dots, y_n) = \langle \phi(x+v/2)\phi(x-v/2)\phi(y_1)\cdots\phi(y_n) \rangle_{\text{1PI}}, \quad (11.35)$$

in which all arguments are fixed, except the vector v . We want to investigate the $|v| \rightarrow 0$ limit. In a QFT that is sufficiently regularized (see Section A8.3), one can expand the product of fields $\phi(x+v/2)\phi(x-v/2)$ in powers of v ,

$$\begin{aligned} \phi(x+v/2)\phi(x-v/2) &= \phi^2(x) + \frac{1}{4} \sum_{\mu_1, \mu_2} v_{\mu_1} v_{\mu_2} [\phi(x)\partial_{\mu_1}\partial_{\mu_2}\phi(x) - \partial_{\mu_1}\phi(x)\partial_{\mu_2}\phi(x)] \\ &\quad + \sum_{n=2}^{\infty} \frac{1}{2^{2n}} \frac{1}{2n!} \sum_{\mu_1, \dots, \mu_{2n}} v_{\mu_1} \cdots v_{\mu_{2n}} \mathcal{O}_{\mu_1 \dots \mu_{2n}} [\phi(x)], \end{aligned} \quad (11.36)$$

in which $\mathcal{O}_{\mu_1 \dots \mu_{2n}}(\phi)$ is a local operator quadratic in ϕ , with $2n$ derivatives.

If we now insert expansion (11.36) into a correlation function, even after field renormalization, all terms of the expansion in general diverge when the cut-off is removed. As we have extensively discussed in Section 11.1, the various local monomials appearing in expansion (11.36) have to be renormalized. Therefore, we have to expand each bare operator on a basis of renormalized operators. In terms of renormalized operators $\{\mathcal{O}^\alpha(\phi)\}_r$, the expansion (11.36) takes the form

$$\phi_r(x+v/2)\phi_r(x-v/2) = \sum_{\alpha} C_{\alpha}(v, \Lambda) \{\mathcal{O}^\alpha(x)\}_r, \quad (11.37)$$

in which, at cut-off Λ fixed, the coefficients $C_{\alpha}(v, \Lambda)$ are regular, even functions of v . An operator $\{\mathcal{O}^\alpha(x)\}_r$ receives contributions from bare operators of equal or higher dimensions. We conclude that, for $|v|$ small, the coefficient functions $C_{\alpha}(v, \Lambda)$ behave like

$$C_{\alpha}(\lambda v, \Lambda) \sim \lambda^{[\mathcal{O}_{\alpha}] - 2[\phi]}, \quad \text{for } \lambda \rightarrow 0. \quad (11.38)$$

When the cut-off becomes infinite, the coefficients of the expansion of C_{α} in powers of v , being renormalization constants, diverge, and if C_{α} has a limit, the limiting function is singular at $v = 0$. The coefficients are functions of the cut-off, which, for $v = 0$, diverge in a way predicted by power counting. Since v is small but non-vanishing, the coefficients grow with the cut-off until Λ is of order $1/|v|$. In this range, all contributions to a given coefficient are of the same order, up to powers of logarithms, because powers of v compensate the powers of Λ . Therefore, at least in perturbation theory, the ordering of operators consequence of equation (11.38) survives, the operators of lowest dimensions dominate the expansion (11.37) for $|v|$ small, and the behaviour of the limiting coefficients $C_{\alpha}(v)$ is given by equation (11.38), up to powers of $\ln v$.

The expansion (11.37) converges towards the SDE of the product of two operators ϕ .

11.3.1 SDE at leading order

Further insight into the structure of the SDE is gained by realizing that the product $\phi_r(x + v/2)\phi_r(x - v/2)$ can be considered as a regularization by point splitting of the local monomials $\{\phi^2(x)\}_r$. Let us investigate this point more precisely in the framework of the $\phi_{d=4}^4$ QFT. We then expect

$$\phi_r(x + v/2)\phi_r(x - v/2) \underset{|\mathbf{v}| \rightarrow 0}{\sim} C_1(v) \{\phi^2(x)\}_r. \quad (11.39)$$

The singularities of $C_1(v)$ for $|\mathbf{v}|$ small should be directly related to the divergences of the renormalization constant Z_2 that renders $\{\phi^2(x)\}_r$ finite.

In what follows, it becomes convenient to treat the product $\frac{1}{2}\phi_r(x + v/2)\phi_r(x - v/2)$ as a local monomial, depending on the point x and a parameter v , in particular, from the point of view of connectivity and one-particle irreducibility. To make this explicit, we set

$$\Xi(x, v) = \frac{1}{2}\phi_r(x + v/2)\phi_r(x - v/2). \quad (11.40)$$

However, the relation between connected correlation functions is then affected, for example,

$$\begin{aligned} \langle \Xi(x, v)\phi_r(y_1) \cdots \phi_r(y_n) \rangle_c &= \frac{1}{2} \langle \phi_r(x + v/2)\phi_r(x - v/2)\phi_r(y_1) \cdots \phi_r(y_n) \rangle_c \\ &\quad + \frac{1}{2} \sum_{I \cup J = \{1, \dots, n\}} \langle \phi_r(x + v/2)\phi_r(y_{i_1}) \cdots \phi_r(y_{i_p}) \rangle_c \\ &\quad \times \langle \phi_r(x - v/2)\phi_r(y_{j_1}) \cdots \phi_r(y_{j_{n-p}}) \rangle_c, \end{aligned} \quad (11.41)$$

in which I and J are all non-empty partitions of $\{1, \dots, n\}$.

Rather than writing explicitly the corresponding relations between vertex functions, we give the relation in terms of generating functionals. Denoting the generating functional of correlation functions with one insertion of $\Xi(x, v)$ by

$$\mathcal{Z}_\Xi(x, v; J) = \frac{1}{2} \frac{\delta^2 \mathcal{Z}(J)}{\delta J(x + v/2) \delta J(x - v/2)}, \quad (11.42)$$

we find for, connected correlation functions with obvious notation,

$$\begin{aligned} \mathcal{W}_\Xi(x, v; J) &= \mathcal{Z}_\Xi(x, v; J) \mathcal{Z}^{-1}(J) \\ &= \frac{1}{2} \frac{\delta^2 \mathcal{W}(J)}{\delta J(x + v/2) \delta J(x - v/2)} + \frac{1}{2} \frac{\delta \mathcal{W}}{\delta J(x + v/2)} \frac{\delta \mathcal{W}}{\delta J(x - v/2)}, \end{aligned} \quad (11.43)$$

and, finally, for the generating functional of vertex functions,

$$\Gamma_\Xi(x, v; \varphi) = -\frac{1}{2}\varphi(x + v/2)\varphi(x - v/2) - \frac{1}{2} \left[\frac{\delta^2 \Gamma}{\delta \varphi(x + v/2) \delta \varphi(x - v/2)} \right]^{-1}. \quad (11.44)$$

where inverse means inverse in the sense of kernels.

Note that the equation is analogous to equation (7.107), which arises in the proof the one-line irreducibility of $\Gamma(\varphi)$.

To now ensure the limit (11.39), we determine $C_1(v)$ by imposing that the insertion of the operator $-\frac{1}{2}C_1^{-1}(v)\Xi(x, v)$ in the two-point function satisfies, for any v , the renormalization condition (9.33). Defining

$$\begin{aligned} C_1^{-1}(v) \int d^4x d^4y_1 d^4y_2 e^{ipx+iq_1y_1+iq_2y_2} \frac{\delta^2\{\Gamma_\Xi(x, v; \varphi)\}_r}{\delta\varphi(y_1)\delta\varphi(y_2)} \Big|_{\varphi=0} \\ = -(2\pi)^4 \delta^{(4)}(p + q_1 + q_2) \tilde{\Gamma}_r^{(1,2)}(v; p; q_1, q_2), \end{aligned} \quad (11.45)$$

we impose

$$\tilde{\Gamma}_r^{(1,2)}(v; 0; 0, 0) = 1. \quad (11.46)$$

Setting $q_1 = q_2 = 0$ in equation (11.45), we derive

$$C_1(v) = - \frac{\delta^2\{\Gamma_\Xi(0, v; \varphi)\}_r}{\delta\tilde{\varphi}(0)\delta\tilde{\varphi}(0)} \Big|_{\varphi=0}, \quad (11.47)$$

in which $\tilde{\varphi}(q)$ is the Fourier transform of $\varphi(x)$.

By differentiating equation (11.44) twice with respect to φ , we relate the right-hand side of equation (11.47) to the φ -field vertex functions, and find

$$C_1(v) = 1 - \frac{1}{2} \int \frac{d^4k}{(2\pi)^4} e^{-ikv} \left[\widetilde{W}_r^{(2)}(k) \right]^2 \tilde{\Gamma}_r^{(4)}(k, -k, 0, 0). \quad (11.48)$$

The equation can also be rewritten as

$$C_1(v) = 1 + \frac{1}{2} m_r^4 \langle \phi_r(v/2) \phi_r(-v/2) \tilde{\phi}_r(0) \tilde{\phi}_r(0) \rangle_c. \quad (11.49)$$

We have introduced a mixed connected correlation function, $\tilde{\phi}(p)$ being the Fourier transform of the field $\phi(x)$ and used the renormalization conditions (9.30).

The coefficient $C_1(v)$ is defined in such a way that the renormalized operator $\phi_r(x + v/2)\phi_r(x - v/2)C_1^{-1}(v)$ then converges towards the operator $\{\phi_r^2(x)\}_r$, correctly normalized.

For example, the implication for the four-point function is

$$\begin{aligned} C_1(v) \tilde{\Gamma}_r^{(1,2)}(p; q_1, q_2) &\underset{|v| \rightarrow 0}{\sim} 1 - \frac{1}{2} \int \frac{d^4k}{(2\pi)^4} e^{-ikv} \widetilde{W}_r^{(2)}(p/2 - k) \widetilde{W}_r^{(2)}(p/2 + k) \\ &\times \tilde{\Gamma}_r^{(4)}(p/2 + k, p/2 - k, q_1, q_2). \end{aligned} \quad (11.50)$$

The neglected contributions are of order v^2 , up to powers of $\ln v$, at any finite order in the perturbative expansion.

Renormalization theory tells us that equation (11.39) is valid as long as the replacement of the operator product by $\phi^2(x)$ does not generate new renormalizations. Therefore, we can use equation (11.39) in all $\phi(x)$ and $\phi^2(x)$ correlation functions except for:

- (i) the two-point function $\langle \phi(x + v/2)\phi(x - v/2) \rangle$, which leads to $\langle \phi^2(x) \rangle$ and
- (ii) the four-point function $\langle \phi(x+v/2)\phi(x-v/2)\phi^2(y) \rangle$, which leads to the ϕ^2 two-point function $\langle \phi^2(x)\phi^2(y) \rangle$.

Both require an additional additive renormalization. The strategy in such cases is to first apply the CS differential operator on the correlation function to generate additional $\int \phi^2(x)dx$ insertions until relation (11.39) can be used. As a consequence, the SDE is modified by contributions which are solutions of the homogeneous CS equations.

11.3.2 One-loop calculation of the leading coefficient of the SDE

Equation (11.48) can be used to calculate the coefficient function $C_1(v)$ in perturbation theory. At one-loop order, it is sufficient to replace correlation functions by their tree level values, and the bare parameters by renormalized parameters:

$$C_1(v) = 1 - \frac{g}{2} \int \frac{d^4 k}{16\pi^2} \frac{e^{-ikv}}{(k^2 + m_r^2)^2} + O(g^2). \quad (11.51)$$

It is clear from this expression that, at one-loop order, $C_1(v)$ has the form

$$C_1(v) = A \ln(m_r |v|) + B + O(v^2). \quad (11.52)$$

A frequently useful idea to extract an asymptotic expansion of this form is to calculate the Mellin transform $\mu(\alpha)$ of the function:

$$\mu(\alpha) = \int_0^\infty dv v^{\alpha-1} C_1(v), \quad (11.53)$$

in which v is the length $|\mathbf{v}|$ of the vector \mathbf{v} . For $\alpha \rightarrow 0$, the expansion (11.52) then implies

$$\mu(\alpha) = -\frac{A}{\alpha^2} + (A \ln m_r + B) \frac{1}{\alpha} + O(1). \quad (11.54)$$

Applying this technique, one has to evaluate

$$f(\alpha) = \int_0^\infty dv v^{\alpha-1} \int \frac{d^4 k}{16\pi^2} \frac{e^{-ikv}}{(k^2 + m_r^2)^2}. \quad (11.55)$$

As usual, we rewrite the integral as

$$f(\alpha) = \int_0^\infty dv v^{\alpha-1} \int_0^\infty t dt \frac{d^4 k}{16\pi^2} e^{-t(k^2 + m_r^2) - ikv}. \quad (11.56)$$

Finally, integrating over k , v , and t , in this order, we obtain

$$f(\alpha) = \frac{1}{8} \left(\frac{2}{m_r} \right)^\alpha \frac{\Gamma^2(1 + \alpha/2)}{\alpha^2}. \quad (11.57)$$

An expansion for $\alpha \sim 0$ yields the coefficients A and B :

$$C_1(v) = 1 - \frac{g}{16} \left[-\ln \left(\frac{|v|m_r}{2} \right) + \psi(1) \right] + O(g^2), \quad (11.58)$$

in which $\psi(z)$ is the logarithmic derivative of the function $\Gamma(z)$.

11.4 Large-momentum expansion of the SDE coefficients: CS equations

To the behaviour at short distance of the product of fields $\phi(x + v/2)\phi(x - v/2)$ is associated, after Fourier transformation, the behaviour at large relative momentum k of the product $\tilde{\phi}(p/2 - k)\tilde{\phi}(p/2 + k)$. However, some information is lost in the transformation. The large-momentum behaviour is only sensitive to the singular part of the short-distance behaviour. For instance, after Fourier transformation, the constant terms in the asymptotic expansion of $C_1(v)$ yield terms proportional to $\delta^{(4)}(k)$, which do not contribute to the large-momentum behaviour. As a consequence, the algebraic structure is, for the same reason, somewhat simplified.

We consider the example of equation (11.50). We introduce the Fourier transform

$$\tilde{C}_1(k) = \int e^{ikv} C_1(v) d^4v. \quad (11.59)$$

After Fourier transformation, for k large, equation (11.50) yields

$$\tilde{\Gamma}_r^{(4)}(p/2 + k, p/2 - k, q_1, q_2) \sim -2\tilde{C}_1(k) \left[\tilde{\Gamma}_r^{(2)}(k) \right]^2 \tilde{\Gamma}_r^{(1,2)}(p; q_1, q_2), \quad (11.60)$$

in which the neglected terms are of order $1/k^2$ up to powers of $\ln k$ at any finite order in perturbation theory.

More generally, due to momentum conservation, the disconnected contributions in equation (11.43) coming from $\delta\mathcal{W}/\delta J(x + v/2)\delta\mathcal{W}/\delta J(x - v/2)$ do not contribute to the large-momentum behaviour.

In addition, the expansion of the right-hand side of equation (11.44) in powers of φ yields two types of contributions: one contribution which becomes one-line irreducible after amputation of the lines corresponding to the fields $\phi(x + v/2)$ and $\phi(x - v/2)$, and the other contributions, which remain reducible. Fig. 11.3 gives the example of the six-point function.

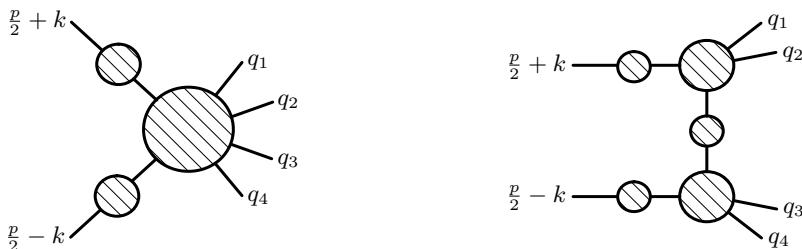


Fig. 11.3 Contributions from the six-point function

Due to momentum conservation in the reducible terms, the propagator that connects the vertex functions carries a momentum of order k , for k large. The corresponding contributions are thus suppressed by a factor $1/k^2$ (up to powers of $\ln k$), and can be neglected at leading order.

Therefore, at leading order, one finds, for $n > 0$,

$$\tilde{\Gamma}_r^{(n+2)}\left(\frac{p}{2} + k, \frac{p}{2} - k, q_1, \dots, q_n\right) \underset{|k| \rightarrow \infty}{\sim} -2\tilde{C}_1(k) \left[\tilde{\Gamma}_r^{(2)}(k) \right]^2 \tilde{\Gamma}_r^{(1,n)}(p; q_1, \dots, q_n). \quad (11.61)$$

As explained in Section 11.3, the equation generalizes to ϕ^2 insertions, for $n > 0$ or $l > 1$,

$$\begin{aligned} \tilde{\Gamma}_r^{(l,n+2)}(p_1, \dots, p_l; p/2+k, p/2-k, q_1, \dots, q_n) \\ \sim -2\tilde{C}_1(k) \left[\tilde{\Gamma}_r^{(2)}(k) \right]^2 \tilde{\Gamma}_r^{(l+1,n)}(p_1, \dots, p_l, p; q_1, \dots, q_n). \end{aligned} \quad (11.62)$$

Callan–Symanzik equation for the first coefficient of the SDE [88]. From now on, all quantities are assumed to be renormalized and we omit the subscript r .

By comparing equations (11.48) and (11.61), we note that in the momentum representation the relevant function is $\tilde{\Gamma}^{(4)}(k, -k, 0, 0)$, which we call $B(k)$ in what follows:

$$B(k) \equiv \tilde{\Gamma}^{(4)}(k, -k, 0, 0) \underset{|k| \rightarrow \infty}{\sim} -2C_1(k) \left[\tilde{\Gamma}^{(2)}(k) \right]^2. \quad (11.63)$$

Equation (11.61) becomes

$$\tilde{\Gamma}^{(n+2)}(p/2+k, p/2-k, q_1, \dots, q_n) \sim B(k) \tilde{\Gamma}^{(1,n)}(p; q_1, \dots, q_n). \quad (11.64)$$

Similarly,

$$\tilde{\Gamma}^{(1,n+2)}(0; p/2+k, p/2-k, q_1, \dots, q_n) \sim B(k) \tilde{\Gamma}^{(2,n)}(0, p; q_1, \dots, q_n). \quad (11.65)$$

We introduce for the CS differential operator the notation

$$D \equiv m_r \frac{\partial}{\partial m_r} + \beta(g) \frac{\partial}{\partial g}. \quad (11.66)$$

We write the CS equations for $\tilde{\Gamma}^{(n+2)}$ as

$$[D - \frac{1}{2}(n+2)\eta(g)] \tilde{\Gamma}^{(n+2)}(\dots) = m_r^2 \sigma(g) \tilde{\Gamma}^{(1,n+2)}(0; \dots). \quad (11.67)$$

We now use equations (11.64) and (11.65) in the large $|k|$ limit:

$$[D - \frac{1}{2}(n+2)\eta(g)] \left(B(k) \tilde{\Gamma}^{(1,n)}(p; q_1, \dots, q_n) \right) = m_r^2 \sigma(g) B(k) \tilde{\Gamma}^{(2,n)}(0, p; q_1, \dots, q_n).$$

Using the CS equation for $\tilde{\Gamma}^{(1,n)}$,

$$D \tilde{\Gamma}^{(1,n)} = [\frac{1}{2}n\eta(g) + \eta_2(g)] \tilde{\Gamma}^{(1,n)} + m_r^2 \sigma(g) \tilde{\Gamma}^{(2,n)}(0, \dots), \quad (11.68)$$

we finally obtain an equation for $B(k)$:

$$[D - \eta(g) + \eta_2(g)] B(k) \sim 0. \quad (11.69)$$

We can also write an equation for $\tilde{C}_1(k)$, using relation (11.63):

$$[D + \eta(g) + \eta_2(g)] \tilde{C}_1(k/m_r, g) \sim 0. \quad (11.70)$$

We can compare this equation with equations (9.38) and (9.39), which imply

$$[D - \eta(g) - \eta_2(g)] Z_2(\Lambda/m_r, g) = 0. \quad (11.71)$$

These equations confirm that $\tilde{C}_1(k/m_r, g)$ and $Z_2^{-1}(\Lambda/m_r, g)$ play similar roles.

11.5 SDE beyond leading order. General operator product

In this section also, all quantities except the operators are assumed to be renormalized.

The $\phi\phi$ product. The product $\phi(x+v/2)\phi(x-v/2)$ is not only a regularization of $\phi^2(x)$ by point splitting, but also of operators of higher dimensions which can be obtained by differentiation.

At next order, which means taking in expansion (11.37) all operators of dimensions 2 and 4 into account, the SDE of $\phi(x+v/2)\phi(x-v/2)$ is an expansion of a regularized bare operator of dimension 4 on a basis of renormalized operators of dimensions 2 and 4, as discussed in previous sections. Indeed, differentiating the expansion (11.37) twice with respect to v , one obtains

$$\frac{1}{4} \left[\phi\left(x + \frac{v}{2}\right) \partial_\mu \partial_\nu \phi\left(x - \frac{v}{2}\right) + \phi\left(x - \frac{v}{2}\right) \partial_\mu \partial_\nu \phi\left(x + \frac{v}{2}\right) - \partial_\mu \phi\left(x + \frac{v}{2}\right) \partial_\nu \phi\left(x - \frac{v}{2}\right) - \partial_\mu \phi\left(x - \frac{v}{2}\right) \partial_\nu \phi\left(x + \frac{v}{2}\right) \right] = \sum_\alpha \partial_\mu \partial_\nu C_\alpha(v) \{ \mathcal{O}^\alpha(\phi(x)) \}_r. \quad (11.72)$$

The product in the right-hand side can be considered as a form regularized by point splitting of $\frac{1}{2} (\phi(x) \partial_\mu \partial_\nu \phi(x) - \partial_\mu \phi(x) \partial_\nu \phi(x))$, which is a linear combination of operators of dimension 4, and spins 0 and 2. In Section 11.1, we have discussed the renormalization of operators of dimension 4 and spin 0. The operators of spin 2 introduce two new linearly independent operators, which can be chosen to be the traceless part of $\partial_\mu \phi(x) \partial_\nu \phi(x)$ and $\partial_\mu \partial_\nu (\phi^2(x))$. Rotation invariance in space implies that operators of different spins do not mix under renormalization. Therefore, in addition to the relations (11.9), we have

$$\begin{aligned} \mathcal{O}_5(\phi(x)) &\equiv \partial_\mu \phi(x) \partial_\nu \phi(x) - \frac{1}{4} \delta_{\mu\nu} (\nabla \phi(x))^2 \\ &= Z_5^{-1} \{ (\mathcal{O}_5(\phi(x))) \}_r - B_5 (\partial_\mu \partial_\nu - \frac{1}{4} \delta_{\mu\nu} \nabla^2) \{ \phi^2(x) \}_r. \end{aligned} \quad (11.73)$$

We can impose the renormalization conditions

$$\begin{aligned} \{ \tilde{\Gamma}_{\mathcal{O}_5}^{(2)} \}_r (q; p_1, p_2) &= - (p_1^\mu p_2^\nu + p_1^\nu p_2^\mu - \frac{1}{2} \delta_{\mu\nu} p_1 p_2) + O(p^4), \\ \{ \tilde{\Gamma}_{\mathcal{O}_5}^{(4)} \}_r (0; 0, 0, 0, 0, 0) &= 0. \end{aligned} \quad (11.74)$$

We can now use exactly the same strategy as for the leading term in the SDE. We insert the expansion (11.72), truncated by omitting all operators of dimension larger than 4, which correspond to vanishing coefficients, in the two- and four-point correlation functions. We go over to the vertex functions in the way explained in Section 11.3, that is, considering the product of fields in (11.72) as a local monomial. Finally, we use the renormalization conditions (11.34) and (11.74) to determine all the coefficients $\partial_\mu \partial_\nu C_\alpha(v)$ in the truncated SDE.

Note that we have lost no information by differentiating twice, since only $C_1(v)$, which has been already determined has a constant part when $|v|$ goes to zero. In this limit, the coefficients $\partial_\mu \partial_\nu C_\alpha(v)$ have singularities in v that are similar to the divergences in Λ/m of the renormalization constants which appear in the expansion of operators \mathcal{O}_5 on a basis of dimension 4 renormalized operators. To determine the asymptotic behaviour of the coefficients $C_\alpha(v)$, we then have to establish RG equations for them, by introducing the SDE in the CS equations for vertex functions. We have to worry about the SDE expansion in the presence of a $\phi^2(x)$ insertion. We have to use the analogue of equation (11.10), the difference being that the new renormalization constant, which appears in front of the contact term, is now a function of v .

We conclude that the coefficients of the SDE satisfy RG equations which are formally identical to the relations between the renormalization constants and the finite RG functions which arise in the CS equations for the operators of dimension 4 and spins 0 and 2, in complete analogy with the correspondence between equations (11.70) and (11.71).

SDE of products of arbitrary local operators. For general local operators A and B , we expect

$$A(x+v/2)B(x-v/2) = \sum_{\alpha} C_{AB}^{\alpha}(v) \mathcal{O}_{\text{r}}^{\alpha}(x), \quad (11.75)$$

in which at any finite order in perturbation theory

$$C_{AB}^{\alpha}(\lambda v) \sim \lambda^{[\mathcal{O}^{\alpha}] - [A] - [B]} \quad \text{for } \lambda \rightarrow 0, \quad (11.76)$$

up to powers of $\ln \lambda$ and the $\{\mathcal{O}^{\alpha}(x)\}_{\text{r}}$ form a complete basis of local operators.

Let us consider the example of $A \equiv B \equiv \{\phi^2(x)\}_{\text{r}}$. The product $\phi^2(x+v/2)\phi^2(x-v/2)$ is a regularized form of $\phi^4(x)$; however, $\phi^4(x)$ by renormalization mixes with all operators of dimension 4 and $\phi^2(x)$. Power counting implies that, among these operators, $\phi^2(x)$ has the most divergent coefficient. Therefore, at leading order,

$$\{\phi^2(x+v/2)\}_{\text{r}} \{\phi^2(x-v/2)\}_{\text{r}} \underset{|v| \rightarrow 0}{\sim} C_{\phi^2\phi^2}^1(v) \{\phi^2(x)\}_{\text{r}}, \quad (11.77)$$

in which the coefficient can be determined by using the renormalization condition for $\{\phi^2\}_{\text{r}}$:

$$\tilde{\Gamma}^{(1,2)}(q; p_1, p_2) = 1, \quad \text{for } q = p_1 = p_2 = 0,$$

and expressing that equation (11.77) should be exact when inserted into the two-point function at the subtraction point where all momenta vanish.

It is then simple to derive RG equations for this new coefficient by inserting the relation into the CS equations for the vertex functions $\tilde{\Gamma}^{(l,n)}$.

We do not go into further detail, since the discussion becomes rather technical. The most important idea to keep in mind is the complete parallelism between the SDE of operator products and the renormalization equations of the corresponding local monomials.

11.6 Light-cone expansion of operator products

After analytic continuation to real time, the length of the vector squared, x^2 , may vanish, while the vector x_{μ} remains finite. In such a situation the relevant expansion for a product of fields is no longer the SDE but, instead, the light-cone expansion (LCE) [89].

It is necessary to classify all operators not only according to their canonical dimensions, but also to their spin s , which characterizes their transformation properties under space rotations. The LCE takes the form

$$\phi(x+v/2)\phi(x-v/2) = \sum_{\alpha,s} C_{\alpha}^s(v^2) P_{\mu_1 \dots \mu_s}^s(v) \mathcal{O}_{\mu_1 \dots \mu_s}^{s,\alpha}(x). \quad (11.78)$$

The polynomial $P_{\mu_1 \dots \mu_s}^s(v)$ is a homogeneous, traceless for $s > 0$, polynomial of the vector v_{μ} , and the operators $\mathcal{O}_{\mu_1 \dots \mu_s}^{s,\alpha}$ form a complete basis of local operators.

For example,

$$P_{\mu_1, \mu_2}^2(v) = v_{\mu_1} v_{\mu_2} - \delta_{\mu_1 \mu_2} v^2/d.$$

When v^2 goes to zero with v_μ finite, the polynomials P^s have a finite limit and, therefore, the coefficients $C_\alpha^s(v^2)$ contain the whole non-trivial dependence on v^2 . The analysis, already performed for the SDE, can be extended, and shows that in perturbation theory $C_\alpha^s(v^2)$ behaves as

$$C_\alpha^s(\lambda^2 v^2) \sim \lambda^{[\mathcal{O}^{s,\alpha}] - 2[\phi] - s}, \quad \text{up to powers of } \ln \lambda, \text{ for } \lambda \rightarrow 0. \quad (11.79)$$

Therefore, the important parameter is no longer the dimension of the operator $\mathcal{O}^{s,\alpha}$, but the quantity

$$\tau = [\mathcal{O}^{s,\alpha}] - s, \quad (11.80)$$

called the *twist*.

The operators of lowest twist dominate the LCE of product of operators.

In the expansion (11.78), the lowest twist is 2, which is the twist of $\phi^2(x)$. Each new factor $\phi(x)$ increases the twist by one unit, while additional derivatives either increase the twist or leave it unchanged.

Therefore, the most general operator of twist 2 has the form

$$\phi(x) (\partial_{\mu_1} \cdots \partial_{\mu_2} \cdots \partial_{\mu_n} - \text{traces}) \phi(x),$$

or is a combination of derivatives of twist 2 operators.

Operators of twist 2 and spin s , since they are the operators of lowest dimension for a given spin, renormalize among themselves. Using previous considerations about the SDE, it is simple to write RG equations for the corresponding coefficients $C_\alpha^s(v^2)$.

12 Relativistic fermions: Introduction

So far, we have discussed only scalar boson quantum field theories (QFTs). We now consider theories with relativistic fermions, as required for particle physics.

In Chapter 4, we have introduced some basic concepts needed for the discussion of Fermi fields: quantum mechanics with Grassmann variables in Section 4.6 and, in Section 4.7, a representation by field integrals of the statistical operator $e^{-\beta H}$ for the non-relativistic Fermi gas, in the formalism of second quantization. In Section 5.6, we have then obtained an expression for the evolution operator.

To discuss relativistic fermions, we must recall how fermions transform under the *spin group* [92]. Therefore, we begin with real time (and Minkovski metric), in the spirit of Section 5.6.

We analyse the free action for Dirac fermions, explain the relation between fields and particles, derive an expression for the scattering matrix, and discuss the non-relativistic limit [93] of a model of self-coupled massive Dirac fermions.

We introduce the formalism of Euclidean relativistic fermions. In the Euclidean formalism, fermions transform under the fundamental representation of the spin group $\text{Spin}(d)$ associated with the $SO(d)$ rotation group (spin 1/2 fermions for $d = 4$).

Since we have devoted Chapter 7 to perturbation theory and functional methods, we outline here only the aspects that are specific to fermions. As for the scalar field theory, we first calculate the Gaussian integral, which corresponds to a free field theory. Then, adding a source term to the action, we obtain the generating functional of correlation functions. The field integral corresponding to a general action with an interaction expandable in powers of the field, can be expressed in terms of a series of Gaussian integrals, which can be calculated, for example, with the help of Wick's theorem.

In Section 12.1, we note the connection between spin and statistics for free fermions. In Section 12.6, we verify the property in a simple perturbative calculation.

For completeness, in Section A12, we describe a few additional properties of the spin group, the algebra of γ matrices, and the corresponding spinors for Euclidean fermions.

12.1 Massive Dirac fermions

We consider massive relativistic fermions, first in even space-time dimensions d , which generalize the four-dimensional Dirac fermions, because, otherwise, the fermion mass breaks space reflection symmetry. In real time, relativistic invariance involves the spin group $\text{Spin}(1, d-1)$ associated with the relativistic group $SO(1, d-1)$ (generalized Lorentz group) of space-time pseudo-orthogonal transformations. We do not discuss these groups extensively, because it is beyond the scope of this work. We describe the structure of fermion states and spectrum in the free field theory. We construct the S -matrix in the interacting theory.

12.1.1 The free massive Dirac fermion

In a formalism based on field integrals, fermion fields $\bar{\psi}_\alpha(x), \psi_\alpha(x)$ are generators of a Grassmann algebra. Moreover, in a relativistic invariant theory, they are *spinors*, vectors with $2^{d/2}$ components when the total space-time dimension d is even, transforming under the action of the spin group.

In real time, and $(d - 1)$ -dimensional space, the free action for a massive Dirac field has the form

$$\mathcal{A}_0(\bar{\psi}, \psi) = \int dt d^{d-1}x \sum_{\alpha, \beta} \bar{\psi}_\alpha(t, x) \left[\frac{1}{i} (\gamma_0)_{\alpha\beta} \partial_t + \boldsymbol{\gamma}_{\alpha\beta} \cdot \nabla_x + m \delta_{\alpha\beta} \right] \psi_\beta(t, x), \quad (12.1)$$

where the $d^{2d/2} \times d^{2d/2}$ γ matrices, γ_0 and $\boldsymbol{\gamma} \equiv \{\gamma_1, \dots, \gamma_{d-1}\}$, are chosen Hermitian, and are defined more precisely by equation (A12.1) in Section A12.1.4. They satisfy

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2 \delta_{\mu\nu} \mathbf{1}. \quad (12.2)$$

In this section, we denote by γ_0 the matrix γ_d associated with the time coordinate $t \equiv x_0 \equiv -ix_d$. In the following, we also need the Hermitian matrix (A12.23),

$$\gamma_S \equiv \gamma_{d+1} = i^{-d/2} \prod_{\mu=1}^d \gamma_\mu, \quad (12.3)$$

which generalizes γ_5 to arbitrary even dimensions and satisfies (equation (A12.24)),

$$\gamma_S^2 = 1, \quad \gamma_S \gamma_\mu + \gamma_\mu \gamma_S = 0, \quad \text{for } \mu \leq d. \quad (12.4)$$

Relativistic invariance. Lets Λ be an element of the fundamental representation of the spin group $\text{Spin}(1, d-1)$, and lets the $d \times d$ matrix $\mathbf{R}(\Lambda)$ be its image in the $SO(1, d-1)$ representation. The elements of the group $SO(1, d-1)$ satisfy (\mathbf{M}^T denotes the *transpose* of the matrix \mathbf{M})

$$\mathbf{R} \mathbf{g} \mathbf{R}^T = \mathbf{g},$$

where \mathbf{g} is the diagonal matrix with coefficients $(+1, -1, \dots, -1)$.

The action (12.1) is relativistic invariant, that is, it is invariant under the transformations

$$(\psi_\Lambda)_\alpha(t, x) = \sum_\beta \Lambda_{\alpha\beta}^{-1} \psi_\beta(\mathbf{R}(t, x)), \quad (\bar{\psi}_\Lambda)_\alpha(t, x) = \sum_\beta \bar{\psi}_\beta(\mathbf{R}(t, x)) \Lambda_{\beta\alpha}. \quad (12.5)$$

After the linear change of variables $\mathbf{R}(t, x) \mapsto (t', x')$, the invariance of the action follows from the identity

$$R_{\mu\nu} \Lambda \tilde{\gamma}_\nu \Lambda^{-1} = \tilde{\gamma}_\mu, \quad \text{with } \tilde{\gamma}_0 = \gamma_0, \quad \tilde{\gamma} = i\boldsymbol{\gamma}.$$

12.1.2 Hamiltonian, spectrum, and particle content

In the relativistic framework, one introduces a field $\bar{\psi}$ that is not the Hermitian conjugate of ψ , because $\bar{\psi}$ has simpler transformation properties under the spin group. Indeed, matrices Λ belonging to the spin group $\text{Spin}(1, d-1)$ satisfy $\gamma_0 \Lambda^\dagger \gamma_0 = \Lambda^{-1}$, as one verifies by adapting the expressions of Section A12.1 to real time.

However, to identify the action (12.1) with an action of the form (5.48) resulting from a Hamiltonian formalism, one has to express $\bar{\psi}$ in terms of ψ^\dagger . With our conventions,

$$\psi^\dagger = -\bar{\psi} \gamma_0. \quad (12.6)$$

Then,

$$\mathcal{A}_0(\psi, \psi^\dagger) = \int dt d^{d-1}x \psi^\dagger(t, x) [i\partial_t - \gamma_0(\boldsymbol{\gamma} \cdot \nabla_x + m)] \psi(t, x).$$

Since the γ matrices are Hermitian, $\mathcal{A}_0 = \mathcal{A}_0^\dagger$, and thus, the corresponding Hamiltonian is Hermitian.

The quadratic form can then be diagonalized by Fourier transformation. After introduction of the Fourier representation,

$$\psi(t, x) = \int d^{d-1}p e^{ipx} \tilde{\psi}(t, p), \quad \psi^\dagger(t, x) = \int d^{d-1}p e^{-ipx} \tilde{\psi}^\dagger(t, p),$$

the free action becomes

$$\mathcal{A}_0(\psi, \psi^\dagger) = (2\pi)^{d-1} \int dt d^{d-1}p \tilde{\psi}^\dagger(t, p) [i\partial_t + h(p)] \tilde{\psi}(t, p),$$

where we have introduced the matrix

$$h(p) = -\gamma_0(i\boldsymbol{\gamma} \cdot \mathbf{p} + m) = h^\dagger(p). \quad (12.7)$$

In contrast with the scalar case, due to the spin structure, the Hamiltonian is not completely diagonalized, but the diagonalization has been reduced to a matrix problem. One verifies

$$h^2(p) = \omega^2(p), \quad \text{with} \quad \omega(p) = \sqrt{p^2 + m^2}. \quad (12.8)$$

The matrix h thus has two eigenvalues $\pm\omega(p)$. We introduce here the matrix γ_S defined by equation (12.3). Then, since

$$\gamma_0 \gamma_S h(p) = -h(p) \gamma_0 \gamma_S,$$

if a spinor $u(p)$ is an eigenvector with eigenvalue $\omega(p)$, $\gamma_0 \gamma_S u(p)$ ($\gamma_0 \gamma_S$ is associated with time reversal) is an eigenvector with eigenvalue $-\omega(p)$. The two corresponding subspaces have equal dimensions.

The two orthogonal, Hermitian projectors P_\pm on the two subspaces are

$$P_\pm = \frac{1}{2} [1 \pm h(p)/\omega(p)], \quad \Rightarrow \quad P_+ + P_- = 1, \quad P_\pm^2 = P_\pm, \quad P_+ P_- = 0. \quad (12.9)$$

With the assignment of ψ as the Grassmann field associated with creation operators, states are created with both positive and negative energies. This means that we have misidentified the reference state, which must be the ground state. As we explain in Section 5.6, by exchanging, after diagonalization, the role of the two conjugate fields we change the sign of the one-particle energy (a property specific to fermions). As in the scalar case, and in contrast to what one might naively have guessed, the fields ψ and ψ^\dagger must be decomposed into a sum of analytic and anti-analytic components to ensure that one-particle states have positive energy.

We thus set

$$\chi_-^*(t, p) = [2\omega(p)]^{1/2} P_- \tilde{\psi}(t, p), \quad \varphi_+(t, p) = [2\omega(p)]^{1/2} P_+ \tilde{\psi}(t, p), \quad (12.10a)$$

$$\varphi_+^\dagger(t, p) = [2\omega(p)]^{1/2} \tilde{\psi}^\dagger(t, p) P_+, \quad \chi_-^T(t, p) = [2\omega(p)]^{1/2} \tilde{\psi}^\dagger(t, p) P_-, \quad (12.10b)$$

where χ_- and φ_+ are two-component spinors.

Conversely,

$$\tilde{\psi}(t, p) = \frac{1}{\sqrt{2\omega(p)}} (\chi_-^*(t, p) + \varphi_+(t, p)), \quad (12.11a)$$

$$\tilde{\psi}^\dagger(t, p) = \frac{1}{\sqrt{2\omega(p)}} (\chi_-^T(t, p) + \varphi_+^\dagger(t, p)). \quad (12.11b)$$

The action becomes

$$\begin{aligned} \mathcal{A}_0(\chi_-, \varphi_+) = (2\pi)^{d-1} \int dt \frac{d^{d-1}p}{2\omega(p)} & \left[\varphi_+^\dagger(t, p) (i\partial_t + \omega(p)) \varphi_+(t, p) \right. \\ & \left. + \chi_-^\dagger(t, p) (i\partial_t + \omega(p)) \chi_-(t, p) \right]. \end{aligned}$$

We find two particles with $2^{d/2-1}$ components and the same mass m , φ_+ transforming under the fundamental representation of the static spin group $\text{Spin}(d-1)$ (see Section A12), χ_- under the conjugated representation. These two representations are equivalent. To verify it, we set

$$\varphi_-(t, p) = C^\dagger \chi_-(t, p), \quad (12.12)$$

where C is a unitary matrix, $C^\dagger C = 1$. The field χ_- satisfies $P_-^T \chi_- = \chi_-$. If we can find a matrix C such that

$$C^\dagger P_-^T C = P_+,$$

then $P_+ \varphi_- = \varphi_-$, and the two fields φ_\pm have the same transformation properties. Since C is unitary, this equation reduces to

$$C^\dagger h^T(p) C = -h(p) \Leftrightarrow C^\dagger (i\gamma^T \cdot p + m) \gamma_0^T C = -\gamma_0 (i\gamma \cdot p + m).$$

In Section A12.1.7, we construct a charge conjugation matrix C that satisfies (equation (A12.39))

$$C^{-1} \gamma_\mu^T C = -\gamma_\mu,$$

and thus, has the required property.

The action then reads

$$\mathcal{A}_0(\varphi_+, \varphi_-) = (2\pi)^{d-1} \int dt \frac{d^{d-1}p}{2\omega(p)} \sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t, p) (i\partial_t + \omega(p)) \varphi_\epsilon(t, p). \quad (12.13)$$

The final form of the action shows that the Dirac field carries two particles transforming under the fundamental representation of the spin group $\text{Spin}(d-1)$ (spin 1/2 particles in 1 + 3 dimensions) related by charge conjugation (in the case of charged particles they have opposite charge).

Fock's space. The construction of Fock's space follows the lines of Section 4.7, with the same kinematic modifications as in the scalar case of Section 6.1.2. The generating functional of n -particle wave functions, which is an element of Fock's space, is a general Grassmann analytic function of φ_\pm . In terms of the fields $\varphi_{\pm\alpha}$, where $\alpha = 1, 2$ denotes the spinor component, one defines the functional

$$\begin{aligned} \Psi(\varphi_+, \varphi_-) = \sum_{n=0} \frac{1}{n!} \sum_{\{\epsilon_i\}, \{\alpha_i\}} & \left(\int \prod_{i=1}^n \frac{d^{d-1}p_i}{2\omega(p_i)} \varphi_{\epsilon_i \alpha_i}(p_i) \right) \\ & \times \psi_{n; \epsilon_1 \alpha_1, \dots, \epsilon_n \alpha_n}(p_1, \dots, p_n), \end{aligned} \quad (12.14)$$

which generates antisymmetric fermion wave functions.

The scalar product then takes the form (4.112), modified to take a covariant form as in equation (6.16).

In functional form, the free Hamiltonian reads

$$\mathbf{H}_0 = (2\pi)^{d-1} \int d^{d-1}p \omega(p) \left(\varphi_+(p) \frac{\delta}{\delta \varphi_+(p)} - \frac{\delta}{\delta \varphi_-(p)} \varphi_-(p) \right).$$

The Hamiltonian written in normal order becomes

$$\mathbf{H}_0 = (2\pi)^{d-1} \int d^{d-1}p \omega(p) \left(\varphi_+(p) \frac{\delta}{\delta \varphi_+(p)} + \varphi_-(p) \frac{\delta}{\delta \varphi_-(p)} \right) + E_0(\text{Dirac}),$$

where the ground state (vacuum) energy $E_0(\text{Dirac})$ is negative, and proportional to the free scalar vacuum energy (6.20):

$$E_0(\text{Dirac}) = -2^{d/2} E_0(\text{scalar}).$$

The particle number operators for both particles commute with the Hamiltonian:

$$\mathbf{N}_\pm = (2\pi)^{d-1} \int d^{d-1}p \varphi_\pm(p) \frac{\delta}{\delta \varphi_\pm(p)}, \quad [\mathbf{N}_\pm, \mathbf{H}_0] = 0,$$

a property that no longer holds, in general, for a local interacting theory.

Remarks.

(i) By adding $2^{d/2}$ scalar bosons of the same mass m to one Dirac fermion, one can construct a theory with vanishing vacuum energy. One can show that this boson–fermion free theory then has a special fermion-type symmetry called *supersymmetry* (see Chapter 27).

(ii) The possibility of solving the problem of the negative energy states depends crucially on the anticommuting character of fermions. This is the reflection of the *connection between spin and statistics*, a property specific to local relativistic QFT.

12.1.3 Interacting theory and the S -matrix

The expression of the S -matrix in an interacting theory generalizes the expression for a non-relativistic Fermi gas of Section 5.6, with the kinematic structure of the scalar case of Section 6.2. One has only to be careful of the signs. One verifies that the S -matrix is given by

$$S(\varphi, \bar{\varphi}) = \int [d\bar{\psi} d\psi] \exp i\mathcal{A}(\bar{\psi}\sqrt{Z} + \bar{\psi}_c, \psi\sqrt{Z} + \psi_c), \quad (12.15)$$

that is, again a field integral in a background field, where the classical anticommuting fields $\bar{\psi}_c, \psi_c$ are solutions to the free field equations, which can be parametrized in the form (12.11) and (12.12):

$$\tilde{\psi}_c(p) = C^* \varphi_-(p) + \varphi_+(p), \quad \tilde{\psi}_c^\dagger(p) = \varphi_-(p) + \bar{\varphi}_+(p)C^T, \quad \bar{\psi}_c = -\psi_c^\dagger \gamma_0,$$

and $P_+ \varphi_\pm = \varphi_\pm$. A renormalization constant Z is also required here to obtain S -matrix elements with the proper normalization.

In the same notation, the unitarity of the S -matrix takes the form

$$\begin{aligned} & \int [d\bar{\varphi}'(p) d\varphi'(p)] \mathcal{S}^*(\varphi', \bar{\varphi}) \mathcal{S}(\varphi', \bar{\varphi}) \exp \left[(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \sum_{\varepsilon=\pm} \bar{\varphi}'_\varepsilon(p) \varphi'_\varepsilon(p) \right] \\ &= \exp \left[(2\pi)^{d-1} \int \frac{d^{d-1}p}{2\omega(p)} \sum_{\varepsilon=\pm} \varphi_\varepsilon(p) \bar{\varphi}_\varepsilon(p) \right]. \end{aligned} \quad (12.16)$$

12.2 Self-interacting massive fermions: Non-relativistic limit

To establish a relation between relativistic and non-relativistic quantum theories, we now investigate the low energy, low momentum of a theory with massive fermions, with arguments analogous to those for scalar bosons of Section 6.4. Since a fermion mass breaks space reflection symmetry in odd dimensions, we consider again only even dimensions.

As an example, we study the non-relativistic limit of the QFT corresponding to the action of a self-interacting fermion field,

$$\begin{aligned} \mathcal{A}(\bar{\psi}, \psi) = & \int dt d^{d-1}x \left[\bar{\psi}(t, x) \left(-i\gamma_0 \partial_t + \boldsymbol{\gamma} \cdot \nabla_x + m \right) \psi(t, x) \right. \\ & \left. + \frac{1}{2} G(\bar{\psi}(t, x) \psi(t, x))^2 \right], \end{aligned} \quad (12.17)$$

which we express in terms of ψ, ψ^\dagger (equation (12.6)) as

$$\begin{aligned} \mathcal{A}(\psi, \psi^\dagger) = & \int dt d^{d-1}x \left[\psi^\dagger(t, x) [i\partial_t - \gamma_0(\boldsymbol{\gamma} \cdot \nabla_x + m)] \psi(t, x) \right. \\ & \left. + \frac{1}{2} G(\psi^\dagger(t, x) \gamma_0 \psi(t, x))^2 \right]. \end{aligned}$$

Due to the spin structure and the linearity in ∇_x of the action, extracting the non-relativistic limit requires slightly more work than in the scalar case. However, one can use the transformations that lead to the action (12.13). In the kinematic part, one then expands the one-particle energy

$$\omega(p) = \sqrt{p^2 + m^2} = m + p^2/2m + O(m^{-3}).$$

In the interaction terms, one takes the non-relativistic limit, neglecting all momenta compared to the mass. The projectors P_\pm defined by equation (12.9) then reduce to

$$P_\pm = \frac{1}{2}(1 \pm \gamma_0).$$

The transformation between fields ψ and φ_\pm becomes local. Choosing a different normalization, one finds

$$\begin{aligned} \varphi_-(t, x) &= P_+ C^\dagger \psi^*(t, x), & \varphi_+(t, x) &= P_+ \psi(t, x), \\ \varphi_+^\dagger(t, x) &= \psi^\dagger(t, x) P_+, & \varphi_-^\dagger(t, x) &= \psi^T(t, x) C P_+, \end{aligned}$$

or, conversely,

$$\psi(t, x) = \varphi_+(t, x) + C^* \varphi_-^*(t, x).$$

For γ matrices, it is convenient to choose a basis in which γ_0 is diagonal, and we restrict, below, the spinor indices to the non-vanishing components of φ_\pm . Using the relation

$$-\bar{\psi}(t, x) \psi(t, x) = \psi^\dagger(t, x) \gamma_0 \psi(t, x) = \varphi_+^\dagger(t, x) \varphi_+(t, x) + \varphi_-^\dagger(t, x) \varphi_-(t, x),$$

one obtains the action

$$\begin{aligned} \mathcal{A}(\varphi_\pm, \varphi_\pm^\dagger) = & \int dt d^{d-1}x \left[\sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t, x) \left[i\partial_t + m - \nabla_x^2/2m \right] \varphi_\epsilon(t, x) \right. \\ & \left. + \frac{1}{2} G \left(\sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t, x) \varphi_\epsilon(t, x) \right)^2 \right]. \end{aligned}$$

One then proceeds in analogy with the boson case, as in Section 6.4. One translates the one-particle energy by the mass m , setting

$$\varphi_{\pm}(t, x) \mapsto e^{imt} \varphi_{\pm}(t, x), \quad \varphi_{\pm}^{\dagger}(t, x) \mapsto e^{-imt} \varphi_{\pm}^{\dagger}(t, x).$$

In the interaction, one neglects all terms that depend explicitly on time. Here, one finds

$$\begin{aligned} \mathcal{A}(\varphi_{\pm}, \varphi_{\pm}^{\dagger}) \\ = \int dt d^{d-1}x \left[\sum_{\epsilon=\pm} \varphi_{\epsilon}^{\dagger}(t, x) [i\partial_t - \nabla_x^2/2m] \varphi_{\epsilon}(t, x) + \frac{1}{2} G \left(\sum_{\epsilon=\pm} \varphi_{\epsilon}^{\dagger}(t, x) \varphi_{\epsilon}(t, x) \right)^2 \right]. \end{aligned}$$

This action describes a many-body theory of two non-relativistic fermions with equal mass and with spin, the spin playing the role of an external quantum number decoupled from space time.

Borrowing the result of Section 5.6, and comparing with the action (4.118), one infers the non-relativistic Hamiltonian \mathbf{H} , up to an infinite energy shift. One again verifies that, in the limiting non-relativistic theory, the number of particles is conserved and, therefore, sectors with different particle number decouple:

$$\mathbf{N}_{\pm\alpha} = \int d^{d-1}x \varphi_{\pm\alpha}(x) \frac{\delta}{\delta \varphi_{\pm\alpha}(x)} \quad \Rightarrow \quad [\mathbf{N}_{\pm\alpha}, \mathbf{H}] = 0.$$

A general n -particle contribution to the generating functional of wave functions can be written as

$$\begin{aligned} \Phi(\varphi) = \sum_n \frac{1}{n!} \int d^{d-1}x_1 \cdots d^{d-1}x_n \sum_{\{\alpha_i\}, \{\epsilon_i\}} & \varphi_{\epsilon_1\alpha_1}(x_1) \varphi_{\epsilon_2\alpha_2}(x_2) \cdots \varphi_{\epsilon_n\alpha_n}(x_n) \\ & \times \phi_{n;\epsilon_1\alpha_1, \epsilon_2\alpha_2, \dots, \epsilon_n\alpha_n}(x_1, \dots, x_n), \end{aligned}$$

where $\phi_{n;\epsilon_1\alpha_1, \epsilon_2\alpha_2, \dots, \epsilon_n\alpha_n}(x_1, \dots, x_n)$ is a totally antisymmetric wave function.

In the n -particle sector, the Hamiltonian reads

$$H_n = -\frac{1}{2m} \sum_{i=1}^n \nabla_{x_i}^2 - G \sum_{i < j} \delta^{(d)}(x_i - x_j). \quad (12.18)$$

The fermions interact through a two-body $\delta^{(d)}(x)$ function potential which can, here, be repulsive or attractive. The spin acts only through the Pauli principle, which dictates the possible symmetries of the wave function Φ .

A final remark is in order: in two dimensions, the fermion QFT that we have considered here, is equivalent to the well-known massive Thirring model (see Chapter 30). The non-relativistic limit, the δ -function model, can be exactly solved by the Bethe ansatz, that is, a complete set of wave-functions is provided by a superposition of a finite number of plane waves in each of the $n!$ sectors corresponding to all possible ordering of particle positions. Its relativistic generalization is also integrable, because particle production does not arise.

12.3 Free Euclidean relativistic fermions

We now perform the analytic continuation to Euclidean time. This makes a study of quantum statistics of relativistic fermions (see Chapter 33) possible and, generally, simplifies perturbative calculations. Therefore, we first discuss the symmetries of the free relativistic fermion action, like invariance under the spin group, under other continuous symmetries like phase rotation or chiral transformations, and under several discrete symmetries like Hermiticity, reflection, and charge conjugation, which also determine the couplings to other fields. The technical basis for the discussion, like properties of the spin group and the definition of γ matrices, as well as our conventions and notation, can be found in Section A12.1.

Note that some of these symmetries have a form somewhat different from what one is familiar with in real-time QFT. After continuation to imaginary time, symmetries that involve a complex conjugation are no longer directly symmetries of the Euclidean action: Hermiticity is lost, and time reversal has another natural definition that makes it indistinguishable from space reflections. We thus describe here their Euclidean analogues.

Euclidean Dirac fermions. The free fermion action $\mathcal{S}_0(\bar{\psi}, \psi)$ for Euclidean massive Dirac fermions, continuation to imaginary time of the standard action for spinor fields (12.1), can be written as (now $x \equiv (x_1, \dots, x_d)$, and $\gamma_0 \mapsto \gamma_d$)

$$\mathcal{S}_0(\bar{\psi}, \psi) = - \int d^d x \sum_{\alpha, \beta} \bar{\psi}_\alpha(x) \left[(\not{\partial})_{\alpha\beta} + m\delta_{\alpha\beta} \right] \psi_\beta(x), \quad (12.19)$$

where the fields $\bar{\psi}_\alpha(x), \psi_\alpha(x)$ are still generators of a Grassmann algebra. In expression (12.19), we have introduced the traditional notation $\not{\partial}$ to represent the matrix $\sum_\mu \partial_\mu \gamma_\mu$.

Chiral components. We show in Section A12.1.6 that, in even dimensions d , the spinor representation can be reduced. We thus define chiral components ψ_\pm of the fermion field

$$\psi_\pm(x) = \frac{1}{2}(1 \pm \gamma_S)\psi(x). \quad (12.20)$$

and correspondingly $\bar{\psi}_\pm(x)$:

$$\bar{\psi}_\pm(x) = \bar{\psi}(x) \frac{1}{2}(1 \pm \gamma_S), \quad (12.21)$$

often denoted by $\psi_R(x), \psi_L(x), \bar{\psi}_R(x), \bar{\psi}_L(x)$ for right and left components, by reference to the propagation in real time.

However, with two of these spinors it is possible to construct only a massless theory:

$$\mathcal{S}_0(\bar{\psi}_-, \psi_+) = - \int d^d x \bar{\psi}_-(x) \not{\partial} \psi_+(x), \quad (12.22)$$

because $\bar{\psi}_- \psi_+ = \bar{\psi}_+ \psi_- = 0$. To construct an action for a massive propagating fermion, the four spinors are required.

12.3.1 Hermitian conjugation

According to the discussion of Section 4.6, Hermiticity of the Hamiltonian is equivalent to invariance of the Euclidean action under complex conjugation followed by Euclidean time reversal. However, we have here to take into account a peculiarity of the relativistic formalism, the Hermitian conjugate of ψ is not $\bar{\psi}$ but, instead, (equation (12.6)) $\psi^\dagger = -\bar{\psi}\gamma_d$, where γ_d is the γ matrix associated with the Euclidean time component.

When one combines these two transformations, one verifies that they can be realized differently, in a way that no longer singles out the time variable. One defines $\bar{\psi}$ now as the Hermitian conjugate of ψ , instead of ψ^\dagger , and, after Hermitian conjugation, performs the transformation

$$\psi(x) \mapsto \gamma_\mu \psi(\tilde{x}), \quad \bar{\psi}(x) \mapsto \bar{\psi}(\tilde{x})\gamma_\mu, \quad \text{with } \tilde{x} = P_\mu x, \quad (12.23)$$

where P_μ is the space reflection along the μ axis: acting on a space vector x , P_μ changes the sign of its component μ :

$$P_\mu x = \tilde{x}, \quad \text{with } \tilde{x} : \begin{cases} \tilde{x}_\mu = -x_\mu, \\ \tilde{x}_\lambda = x_\lambda \text{ for } \lambda \neq \mu. \end{cases} \quad (12.24)$$

We have chosen a generic component μ to emphasize that all Euclidean components are equivalent. The symmetry corresponding to the product of these two transformations is called *reflection Hermiticity*.

Let us apply the transformation on the action (12.19). After Hermitian conjugation we find

$$\mathcal{S}_0^\dagger(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (-\not{\partial} + m) \psi(x),$$

because ∂_μ is anti-Hermitian. In the transformation (12.23) the mass term is invariant, and in $\not{\partial}$ the contribution $\sum_{\lambda \neq \mu} \gamma_\lambda \partial_\lambda$ changes sign as a consequence of the anticommutation with γ_μ while the remaining term changes sign from $\partial_\mu \mapsto -\partial_\mu$.

The determinant resulting from the integration over ψ and $\bar{\psi}$ in the field integral is thus real. Eigenvalues of the operator $\not{\partial} + m$ are real or appear as complex conjugate pairs.

The action (12.22) has also reflection Hermiticity as a symmetry. Indeed,

$$\mathcal{S}_0^\dagger(\bar{\psi}_-, \psi_+) = \int d^d x \bar{\psi}_+(x) \not{\partial} \psi_-(x),$$

since γ_S is Hermitian. But then, in the second transformation,

$$(1 - \gamma_S)\gamma_\mu = \gamma_\mu(1 + \gamma_S),$$

in such a way that the initial chirality is recovered.

12.3.2 Spin group and reflections

We first show that both actions (12.19) and (12.22) are invariant under the transformations of the Euclidean spin group. We transform the spinors ψ and $\bar{\psi}$ as

$$\begin{aligned} (\psi_\Lambda)_\alpha(x) &= \sum_\beta \Lambda_{\alpha\beta}^\dagger \psi_\beta(\mathbf{R}x), \\ (\bar{\psi}_\Lambda)_\alpha(x) &= \sum_\beta \bar{\psi}_\beta(\mathbf{R}x) \Lambda_{\beta\alpha}, \end{aligned} \quad (12.25)$$

where Λ belong to the spin group $\text{Spin}(d)$ (equation (A12.32)), and the matrix $\mathbf{R}(\Lambda)$ is its image in $SO(d)$ (equation (A12.33)).

After the change of variables $\mathbf{R}x \mapsto x'$, the invariance of the action follows from the identity

$$\sum_{\nu} R_{\mu\nu} \Lambda \gamma_{\nu} \Lambda^{\dagger} = \gamma_{\mu},$$

which is implied by equation (A12.33).

Reflections. In even dimensions, a reflection along the $\mu = 1$ axis corresponds to the transformation Π_1 defined by (Section A12.1.6):

$$\Pi_1 : \quad \psi_{\Pi_1}(x) = \gamma \gamma_1 \psi(\tilde{x}), \quad \bar{\psi}_{\Pi_1}(x) = \bar{\psi}(\tilde{x}) \gamma_1 \gamma^{\dagger}, \quad \text{with } \tilde{x} = (-x_1, x_2, \dots, x_d). \quad (12.26)$$

The mass term in action (12.19) is clearly invariant. In the term \not{D} , the space reflection changes ∂_1 in $-\partial_1$, but then $\gamma \gamma_1$ anticommutes with γ_1 and commutes with all other γ_{μ} matrices. The total action (12.19) is thus invariant.

In contrast with the action (12.19), the action (12.22) is not invariant under reflection, since reflection exchanges chiral components.

In odd dimensions, total reflection $\tilde{x} = -x$ can be implemented by

$$\Pi : \quad \psi_{\Pi}(x) = \psi(\tilde{x}), \quad \bar{\psi}_{\Pi}(x) = -\bar{\psi}(\tilde{x}), \quad (12.27)$$

a transformation that does not belong to the spin group but commutes with it. However, then the fermion mass term is not invariant under reflection: in odd dimensions a *mass term* for a spinor fermion *violates parity conservation*.

12.3.3 Charge conjugation, charge conservation, and chiral symmetry

Charge conjugation. We introduce a unitary matrix C and transform spinors as

$$\psi_{\alpha}(x) = \sum_{\beta} \bar{\psi}'_{\beta}(x) C_{\beta\alpha}^{\dagger}, \quad \bar{\psi}_{\alpha}(x) = - \sum_{\beta} C_{\alpha\beta} \psi'_{\beta}(x). \quad (12.28)$$

As a function of the new fields ψ' and $\bar{\psi}'$, the action (12.19) now reads

$$\mathcal{S}_0(\bar{\psi}', \psi') = - \int d^d x \bar{\psi}'(x) (-C^{\dagger} \not{D}^T C + m) \psi'(x). \quad (12.29)$$

The action (12.19) is thus invariant if the matrix C satisfies

$$C^{\dagger} \gamma_{\mu}^T C = -\gamma_{\mu}.$$

We recognize the definition of the charge conjugation matrix (A12.39).

We also consider the action (12.22). The transformation (12.28) leads to

$$\begin{aligned} \not{D}(1 + \gamma_S) &\mapsto -C^{\dagger}(1 + \gamma_S)^T \not{D}^T C = -C^{\dagger}(1 + \gamma_S) \not{D}^T C = -C^{\dagger} \not{D}^T (1 - \gamma_S) C \\ &= \not{D} C^{\dagger}(1 - \gamma_S) C. \end{aligned}$$

(The matrix γ_S is symmetric.) We show in Section A12.1.7 that if the dimension d is of the form $d = 2 \pmod{4}$ then,

$$C^{\dagger}(1 + \gamma_S) C = (1 - \gamma_S),$$

and, therefore, the action (12.22) is invariant. By contrast, if the dimension is a multiple of four then,

$$\not{D}(1 + \gamma_S) \mapsto \not{D}(1 - \gamma_S),$$

and charge conjugation is not a symmetry. However, charge conjugation multiplied by space reflection (CP), which exchanges chiral components, is a symmetry (the situation in the lepton sector if neutrinos would be massless).

Finally, to explain the denomination *charge conjugation*, we consider charged fields ψ and $\bar{\psi}$ with charges $\mp e$, coupled to an external electromagnetic field $A_\mu(x)$. The action then takes the form

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (\not{d} + m + ie\mathcal{A}(x)) \psi(x).$$

After charge conjugation, as a consequence of equation (A12.39), the sign of the charge e has changed.

Odd dimensions. In odd dimensions, reflection symmetry forbids a mass. Charge conjugation can then be implemented by the matrix \tilde{C} (definition (A12.41)), which satisfies

$$\tilde{C}^\dagger \gamma_\mu^T \tilde{C} = \gamma_\mu,$$

and the transformations

$$\psi_\alpha(x) = \sum_\beta \bar{\psi}'_\beta(x) \tilde{C}_{\beta\alpha}^\dagger, \quad \bar{\psi}_\alpha(x) = \sum_\beta \tilde{C}_{\alpha\beta} \psi'_\beta(x). \quad (12.30)$$

Self-conjugate spinors. For some dimensions (this includes three and four), it is possible to write a consistent theory for self-conjugate spinors, that is, that satisfy $\bar{\psi}^T = C\psi$. They correspond to neutral fermion fields and are called Majorana spinors. The existence of Majorana neutrinos has been proposed for theoretical and experimental reasons but is not established (2020).

Fermion number conservation. If we assign a fermion number $+1$ to ψ and -1 to $\bar{\psi}$, we note that the action (12.19) conserves this fermion number. The fermion number conservation corresponds to a $U(1)$ invariance of the action:

$$\psi_\theta(x) = e^{i\theta} \psi(x), \quad \bar{\psi}_\theta(x) = e^{-i\theta} \bar{\psi}(x). \quad (12.31)$$

For charged fermions, the fermion number is proportional to the electric charge.

Note that the existence of *Majorana particles* leads to fermion number violation (the existence of Majorana neutrinos would lead to neutrinoless double β decays).

Chiral symmetry. In even dimensions, the massless free fermion action (12.19) possesses an important additional $U(1)$ symmetry, called *chiral symmetry*, defined by

$$\psi_\theta(x) = e^{i\theta\gamma_5} \psi(x), \quad \bar{\psi}_\theta(x) = \bar{\psi}(x) e^{i\theta\gamma_5}. \quad (12.32)$$

Such a symmetry has deep consequences, in particular implying an important difference between boson and fermion fields. In contrast with bosons, *the property for fermions to be massless can be enforced by a symmetry of the action*.

12.4 Partition function. Correlations

In Section 4.7, we have derived an expression for the statistical operator of a system of non-relativistic fermions in the form of a field integral. Here, we generalize the expression to a relativistic QFT, using the formalism of Euclidean fermions introduced in Section 12.3. We define the partition function and introduce the generating functional of correlation functions, continuation to imaginary time of the Green's fermions, which lead to the S -matrix. We show how to calculate them in a perturbative expansion.

The partition function. From the combined analyses of Sections 4.6 and 12.3, we infer that the partition function for self-interacting massive Dirac fermions is given by an integral over Grassmann fields of the form

$$\mathcal{Z}(\beta) = \int [d\psi(t, x)d\bar{\psi}(t, x)] \exp \left[-\mathcal{S}(\bar{\psi}, \psi) + \mu \int_0^\beta dt \int d^{d-1}x \bar{\psi}(t, x) \gamma_d \psi(t, x) \right], \quad (12.33)$$

with

$$\mathcal{S}(\bar{\psi}, \psi) = - \int_0^\beta dt \int d^{d-1}x [\bar{\psi}(t, x)(\not{\partial} + m)\psi(t, x) + V(\bar{\psi}(t, x), \psi(t, x))], \quad (12.34)$$

where t is the Euclidean time, β the inverse temperature, and μ the chemical potential coupled to the conserved fermion charge (see equation (12.31) and Section A13.1). The term coupled to the chemical potential μ is Hermitian, corresponds to a quantum operator which commutes with the Hamiltonian, and has the correct non-relativistic limit.

Moreover, fermion fields satisfy anti-periodic boundary conditions in the time direction:

$$\psi(t \equiv x_d = 0, x) = -\psi(\beta, x), \quad \bar{\psi}(0, x) = -\bar{\psi}(\beta, x).$$

In what follows, we work for simplicity at zero temperature and in zero chemical potential, the finite temperature field theory being discussed in Chapter 33. In this limit, the boundary conditions play no role, and we no longer distinguish between space and time.

From now on, x thus denotes all d coordinates.

We now introduce Grassmann sources $\bar{\eta}, \eta$ and consider the more general field integral,

$$\mathcal{Z}(\bar{\eta}, \eta) = \int [d\psi(x)d\bar{\psi}(x)] \exp \left\{ -\mathcal{S}(\bar{\psi}, \psi) + \int d^d x [\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)] \right\}, \quad (12.35)$$

where $\mathcal{Z}(\bar{\eta}, \eta)$ is the generating functional of $\psi, \bar{\psi}$ field correlation functions, since

$$\mathcal{Z}^{-1}(0, 0) \prod_{i=1}^n \frac{\delta}{\delta \eta(x_i)} \prod_{j=1}^n \frac{\delta}{\delta \bar{\eta}(y_j)} \mathcal{Z}(\eta, \bar{\eta}) \Big|_{\eta=\bar{\eta}=0} = (-1)^n \langle \bar{\psi}(x_1) \cdots \bar{\psi}(x_n) \psi(y_1) \cdots \psi(y_n) \rangle,$$

where $\langle \bullet \rangle$ means expectation value with the weight $e^{-\mathcal{S}}$. Because the sources $\eta(x), \bar{\eta}(x)$ are generators of a Grassmann algebra, correlation functions are antisymmetric in their arguments, in agreement with Fermi–Dirac statistics.

The Gaussian integral. We first calculate the Gaussian integral with external sources,

$$\mathcal{Z}_G(\bar{\eta}, \eta) = \int [d\psi(x)d\bar{\psi}(x)] \exp \left[-\mathcal{S}_0(\bar{\psi}, \psi) + \int d^d x (\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)) \right], \quad (12.36)$$

where $\mathcal{S}_0(\bar{\psi}, \psi)$ is the free action (12.19):

$$\mathcal{S}_0(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (\not{\partial} + m) \psi(x). \quad (12.37)$$

As usual, we shift variables $\psi \mapsto \psi'$ to eliminate the linear terms:

$$\psi(x) + (\not{\partial} + m)^{-1} \eta(x) = \psi'(x), \quad \bar{\psi}(x) + \bar{\eta}(x) (\not{\partial} + m)^{-1} = \bar{\psi}'(x). \quad (12.38)$$

Normalizing the field integral (12.36) by $\mathcal{Z}(0, 0) = 1$, we obtain

$$\mathcal{Z}_G(\bar{\eta}, \eta) = \exp \left[- \int d^d x d^d y \bar{\eta}(y) \Delta_F(y, x) \eta(x) \right], \quad (12.39)$$

in which the fermion propagator Δ_F has the Fourier representation,

$$\Delta_F(y, x) = \frac{1}{(2\pi)^d} \int d^d p e^{-ip(x-y)} \tilde{\Delta}_F(p), \quad \text{with} \quad \tilde{\Delta}_F(p) = \frac{(m - i\cancel{p})}{p^2 + m^2}. \quad (12.40)$$

On mass-shell ($p^2 = -m^2$), $m - i\cancel{p}$ is a projector on a space of dimension $2^{[d/2]-1}$. This reflects the property that physical massive fermion states can be classified according to the static spin group $\text{Spin}(d-1)$, the subgroup of $\text{Spin}(d)$ that leaves the momentum p invariant.

The fermion two-point correlation function in a free or Gaussian theory is given by

$$\langle \bar{\psi}_\alpha(x) \psi_\beta(y) \rangle_0 = -\frac{\delta}{\eta_\alpha(x)} \frac{\delta}{\bar{\eta}_\beta(y)} \mathcal{Z}_G(\bar{\eta}, \eta) = (\Delta_F)_{\beta\alpha}(y, x), \quad (12.41)$$

where $\langle \bullet \rangle_0$ means expectation value with the weight e^{-S_0} . A generalization of equation (1.80) to the action (12.34) yields the identity

$$\mathcal{Z}(\bar{\eta}, \eta) = \exp \left[\int d^d x V \left(-\frac{\delta}{\delta \eta(x)}, \frac{\delta}{\delta \bar{\eta}(x)} \right) \right] \mathcal{Z}_G(\bar{\eta}, \eta). \quad (12.42)$$

The identity leads to the perturbative expansion of a field theory with self-interacting fermions. Alternatively, the field integral (12.35) can be expanded in powers of V , η , $\bar{\eta}$, and all terms evaluated using Wick's theorem for fermion fields, a simple generalization of the form (1.78), which can be written in terms of free field expectation values as

$$\left\langle \prod_{i=1,n} \bar{\psi}_{\alpha_i}(x_i) \psi_{\beta_i}(y_i) \right\rangle_0 = \sum_{\substack{\text{permutations} \\ P \text{ of } \{1, 2, \dots, n\}}} \text{sgn}(P) \prod_{i=1,n} \langle \bar{\psi}_{\alpha_{P(i)}}(x_{P(i)}) \psi_{\beta_i}(y_i) \rangle_0, \quad (12.43)$$

where $\text{sgn}(P)$ is the signature of the permutation P .

12.5 Generating functionals

We have only discussed connected functions and vertex functions in the case of a boson field theory, but the extension to fermions is straightforward. Let $\bar{\psi}, \psi$ be a Dirac fermion field and $\mathcal{S}(\bar{\psi}, \psi)$ the corresponding local action. It is clear that $\mathcal{W}(\eta, \bar{\eta}) = \ln \mathcal{Z}(\eta, \bar{\eta})$ is still the generating functional of connected correlation functions (we have called $\bar{\eta}$ and η the sources for ψ and $\bar{\psi}$). If, following the conventions of Section 12.4, we write the source terms in the field integral $\bar{\eta}\psi + \bar{\psi}\eta$, then we define the Legendre transform of \mathcal{W} by

$$\Gamma(\bar{\psi}, \psi) + \mathcal{W}(\eta, \bar{\eta}) = \int d^d x [\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)], \quad (12.44a)$$

$$\psi(x) = \frac{\delta \mathcal{W}}{\delta \bar{\eta}(x)}, \quad \bar{\psi}(x) = -\frac{\delta \mathcal{W}}{\delta \eta(x)}. \quad (12.44b)$$

The equations (12.44b) are equivalent to

$$\eta(x) = \frac{\delta\Gamma}{\delta\bar{\psi}(x)}, \quad \bar{\eta}(x) = -\frac{\delta\Gamma}{\delta\psi(x)}.$$

With these conventions, one can verify that $\Gamma(\bar{\psi}, \psi) = \mathcal{S}(\bar{\psi}, \psi)$ in the tree approximation. All the other algebraic properties derived for bosons generalize to the fermion case. However, we recall here that a Gaussian integration over fermion fields yields a determinant instead of the inverse of a determinant for a complex scalar. This implies that, in Feynman diagrams, fermion loops are affected by an additional minus sign compared to boson loops.

12.5.1 Boson–fermion action: An example

Many field theories involve both fermions and bosons. An example, which is studied in Section 20.2.1, is

$$\begin{aligned} & \mathcal{Z}(\bar{\eta}, \eta, J) \\ &= \int [d\psi d\bar{\psi} d\phi] \exp \left[-\mathcal{S}(\bar{\psi}, \psi, \phi) + \int d^d x [\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x) + J(x)\phi(x)] \right], \end{aligned} \quad (12.45)$$

in which $\psi, \bar{\psi}, \eta, \bar{\eta}$ are Grassmann fields and ϕ, J real scalar fields, and in which the action $\mathcal{S}(\bar{\psi}, \psi, \phi)$ has the form

$$\mathcal{S}(\bar{\psi}, \psi, \phi) = \int d^d x \{ -\bar{\psi}(x) [\partial + M + g\phi(x)] \psi(x) \} + \mathcal{S}_B(\phi), \quad (12.46)$$

with

$$\mathcal{S}_B(\phi) = \int d^d x \left[\frac{1}{2} (\nabla\phi(x))^2 + \frac{1}{2} r\phi^2(x) + \frac{1}{24} \lambda\phi^4(x) \right].$$

The functional $\mathcal{Z}(\bar{\eta}, \eta, J)$ in equation (12.45) generates both ϕ field and $\psi, \bar{\psi}$ field correlation functions. The perturbative expansion is generated by

$$\begin{aligned} & \mathcal{Z}(\bar{\eta}, \eta, J) \\ &= \exp \left[- \int d^d x \left(\frac{\lambda}{24} \left(\frac{\delta}{\delta J(x)} \right)^4 + g \frac{\delta}{\delta J(x)} \frac{\delta}{\delta \eta(x)} \frac{\delta}{\delta \bar{\eta}(x)} \right) \right] \mathcal{Z}_G(\bar{\eta}, \eta, J), \end{aligned} \quad (12.47)$$

in which $\mathcal{Z}_G(\bar{\eta}, \eta, J)$ is the product of free fermion and free boson functionals:

$$\mathcal{Z}_G(\bar{\eta}, \eta, J) = \exp \left[\int d^d x d^d y \left(\frac{1}{2} J(x) \Delta(x, y) J(y) - \bar{\eta}(x) \Delta_F(x, y) \eta(y) \right) \right]. \quad (12.48)$$

All the algebraic transformations that we have performed on expression (7.17) can be generalized to the representation (12.47), in particular, field equations can be derived, or infinitesimal change of variables justified in perturbation theory. Moreover, a functional δ -function can be defined for fermion fields.

12.5.2 One-loop calculation

As an example, we now calculate the one-loop contribution corresponding to the field integral (12.45) with the boson-fermion action (12.46).

We first look again for the solutions $\phi_c, \bar{\psi}_c, \psi_c$ of the field equations. After shifting the fields, we then have to calculate a Gaussian field integral over boson and fermion fields. The result can be inferred from the expressions (1.61–1.63) for the determinant of a mixed matrix involving bosons and fermions or can be obtained directly by integrating over fermions first, and then over bosons.

After a short calculation, one finds for connected correlation functions

$$\begin{aligned} \mathcal{W}(J, \bar{\eta}, \eta) = & -\mathcal{S}(\phi_c, \bar{\psi}_c, \psi_c) + \int d^d x [J(x)\phi_c(x) + \bar{\eta}(x)\psi_c(x) + \bar{\psi}_c(x)\eta(x)] \\ & - \frac{1}{2} \text{tr} \ln \left[\frac{\delta^2 \mathcal{S}_B}{\delta \phi_c(x) \delta \phi_c(y)} + 2g^2 \bar{\psi}_c [\not{\partial} + g\phi_c]^{-1} \psi_c \right] + \text{tr} \ln [\not{\partial} + g\phi_c], \end{aligned}$$

where the fields $\phi_c, \bar{\psi}_c, \psi_c$ are solutions of the field equations

$$\frac{\delta \mathcal{S}_B}{\delta \phi_c(x)} - g\bar{\psi}_c(x)\psi_c(x) - J(x) = 0, \quad (12.49a)$$

$$[\not{\partial} + g\phi_c(x)]\psi_c(x) + \eta(x) = 0, \quad (12.49b)$$

$$\bar{\psi}_c(x)[\not{\partial} + g\phi_c(x)] + \bar{\eta}(x) = 0 \quad (12.49c)$$

(by convention the operator ∂_μ in (12.49c) acts on the left with a minus sign).

A Legendre transformation then yields the generating functional of vertex functions at one-loop order:

$$\begin{aligned} \Gamma(\varphi, \bar{\psi}, \psi) = & \mathcal{S}(\varphi, \bar{\psi}, \psi) - \text{tr} \ln (\not{\partial} + g\varphi) \\ & + \frac{1}{2} \text{tr} \ln \left\{ \frac{\delta^2 \mathcal{S}_B}{\delta \varphi(x) \delta \varphi(y)} + 2g^2 \bar{\psi} [\not{\partial} + g\phi]^{-1} \psi \right\}. \end{aligned} \quad (12.50)$$

The expansion of $\Gamma(\varphi, 0, 0)$ in powers of φ , generates a set of fermion one-loop Feynman diagrams (see Section 7.9). A similar integral over boson fields would have generated a contribution of the form $+ \text{tr} \ln$. Again, we verify that, compared to boson loops, *fermion loops* are multiplied by an additional *minus sign*.

12.6 Connection between spin and statistics

We have noticed in Section 12.1 that fields transforming under the fundamental representation of the spin group (spin 1/2 in four dimensions) could only be quantized as fermions. This connection between spin and statistics is a deep consequence of locality, Hermiticity of the Hamiltonian, and relativistic invariance: fermions transform under representations of odd degrees of the spin group, while bosons transform under the $SO(d)$ group. This implies that, in four dimensions, bosons must have integer spin while fermions must have half-integer spin.

We illustrate this property here, which can be proven with a great deal of generality [94], by an explicit calculation.

We have shown in Section A2.2 that, as a consequence of the Hermiticity of the Hamiltonian, the two-point function has a spectral representation in terms of a positive measure. We have translated this result into the relativistic kinematics in Section 6.6.

All possible intermediate states contribute with the same sign. This result can be generalized to the discontinuity in the physical domain of diagonal scattering amplitudes. We now show that the sign of fermion loops implies a relation between spin and statistics.

Boson contribution. We consider the leading order contribution from a scalar field ϕ to the two-point function of a coupled scalar field χ . The ϕ -field action is

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{2} g \phi^2(x) \chi(x) \right].$$

The integration over ϕ yields a (non-local) contribution to the χ action ($\ln \det = \text{tr} \ln$):

$$\delta \mathcal{S}(\chi) = \frac{1}{2} \text{tr} \ln(-\nabla_x^2 + m^2 + g\chi).$$

The expansion of $\delta \mathcal{S}$ to order g^2 yields a term linear in χ , which shifts the χ field expectation value, and a quadratic term, which modifies the two-point function in the Gaussian approximation. The corresponding contribution to $\tilde{\delta}W_{\chi^2}^{(2)}$, the χ two-point function in Fourier representation, is

$$\tilde{\delta}W_{\chi^2}^{(2)} = \frac{1}{2} g^2 \Delta_{\chi}^2(p) B_d(p, m) + \text{const.},$$

with

$$B_d(p, m) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(q^2 + m^2) [(p + q)^2 + m^2]}. \quad (12.51)$$

A convenient integral representation of the diagram is obtained by using Feynman's parametrization (see Section A10) in the form

$$\begin{aligned} B_d(p, m) &= \frac{1}{(2\pi)^d} \int_0^1 ds \int \frac{d^d q}{\left[(1-s)q^2 + m^2 + s[(p+q)^2 + m^2] \right]^2} \\ &= N_d (1-d/2) \frac{\pi}{2 \sin(\pi d/2)} \int_0^1 ds [s(1-s)p^2 + m^2]^{d/2-2}, \end{aligned}$$

where N_d is the loop factor (10.12), which indeed shows that the function has a cut for $s = -p^2 > (2m)^2$, the region of physical ϕ particle emission, with a positive imaginary part,

$$\frac{1}{2i} [B_d(p, m)|_{s+i\epsilon} - B_d(p, m)|_{s-i\epsilon}] = \frac{\pi^{(3-d)/2} 2^{3-2d}}{\Gamma[\frac{1}{2}(d-1)]} s^{-1/2} (s-4m^2)^{(d-3)/2},$$

a result which is consistent with the representation (6.60) for $\delta W_{\chi^2}^{(2)}$.

By contrast, if the contribution would come from scalar fermions, the $\text{tr} \ln$ would be replaced by $-\text{tr} \ln$, and the contribution have the opposite sign, which is inconsistent with Hermiticity.

Spinor fermions. Let us calculate now the contribution of spinor fermions. After a Gaussian integration, the contribution to the χ action is

$$\delta S(\chi) = -\text{tr} \ln(\not{\partial} + m + g\chi),$$

which expanded to order g^2 yields the contribution to the χ two-point function,

$$\begin{aligned} \delta \tilde{W}_{\chi^2}^{(2)}(p) &= -\frac{g^2}{(2\pi)^d} \text{tr} \int \frac{d^d q (-iq + m)(-ip - iq + m)}{(q^2 + m^2) [(p+q)^2 + m^2]} \\ &= -\frac{g^2}{(2\pi)^d} \text{tr}_\gamma \mathbf{1} \int \frac{d^d q (m^2 - pq - p^2)}{(q^2 + m^2) [(p+q)^2 + m^2]}. \end{aligned}$$

We then use the identity

$$m^2 - pq - q^2 = 2m^2 + \frac{1}{2}p^2 - \frac{1}{2}[(p+q)^2 + m^2 + q^2 + m^2].$$

The two terms inside the brackets cancel a denominator and thus yield a constant (in general divergent, see Chapter 8) contribution, which has no discontinuity. Thus,

$$\delta\tilde{W}_{\chi^2}^{(2)}(p) = -g^2 \operatorname{tr}_\gamma \mathbf{1}(p^2 + 4m^2)\Delta_\chi^2(p)B_d(p, m) + \text{const.} \quad (12.52)$$

In the region of physical emission $-p^2 > (2m)^2$ the factor $(p^2 + 4m^2)$, which reflects the spin structure, is negative and compensates the negative sign due to the fermion loop.

12.7 Divergences and momentum cut-off

For the Dirac fermions $\psi, \bar{\psi}$ discussed here, the expression (12.40) of the fermion propagator in the field Fourier representation yields the ultraviolet (UV) dimension of ψ and $\bar{\psi}$. Indeed (see also Section 8.2 with $\sigma = 1$),

$$\tilde{\Delta}_F(\lambda p) \propto 1/\lambda, \quad \text{for } \lambda \rightarrow \infty \Rightarrow [\psi] = [\bar{\psi}] = \frac{1}{2}(d-1). \quad (12.53)$$

The $(\bar{\psi}\psi)^2$ self-interaction has dimension $2d-2-d = d-2$, and thus is only renormalizable for $d = 2$. Other renormalizable interactions involve a scalar field (or a gauge field in a renormalizable gauge). An interaction of the form $\bar{\psi}\psi\phi^n$ has dimension $d-1+n(d-2)/2-d$. For $d = 2$, it is super-renormalizable for all n . For $d = 3$, it is super-renormalizable for $n = 1$ and renormalizable for $d = 2$. Finally, for $d = 4$ only $\bar{\psi}\psi\phi$ is renormalizable. Again, no stable local field theory is renormalizable for $d > 4$. One may wonder whether this property is related to the empirical fact that physical space-time has dimension four.

12.7.1 Momentum cut-off regularization

Since the perturbative expansions of all QFTs involving fermions exhibit UV divergences, theories with fermions also require a regularization.

The free action (12.19),

$$\mathcal{S}_0(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x)(\not{D} + m)\psi(x),$$

can be replaced by the regularized expression

$$\mathcal{S}_{0\text{ reg.}}(\bar{\psi}, \psi) = \int d^d x \bar{\psi}(x)(\not{D} + m) \prod_{i=1}^s (1 - \nabla^2/M_i^2)\psi(x), \quad (12.54)$$

where $M_i/\Lambda = O(1)$. The propagator $(m + i\not{p})^{-1}$ then is replaced by

$$\Delta_F(p) = (m + i\not{p})^{-1} \prod_{i=1}^s (1 + p^2/M_i^2)^{-1}. \quad (12.55)$$

Note that for $m = 0$, this modification preserves, in even dimensions, chiral symmetry (defined by (12.32)). In four dimensions, the chiral transformations are

$$\psi_\theta(x) = e^{i\theta\gamma_5} \psi(x), \quad \bar{\psi}_\theta(x) = \bar{\psi}(x) e^{i\theta\gamma_5}.$$

12.7.2 Regulator fields

The fermion inverse propagator (12.55) can be written as

$$\Delta_F^{-1}(p) = (m + i\cancel{p}) \prod_{k=1}^n (1 + i\cancel{p}/M_k)(1 - i\cancel{p}/M_k).$$

This indicates that again the same form can be obtained by adding a set of regulator fields $\bar{\psi}_{k\pm}$, $\psi_{k\pm}$. One replaces the action (12.19) by

$$\mathcal{S}_{0\text{ reg}} = - \int d^d x \bar{\psi}(x) (\cancel{\partial} + m) \psi(x) - \sum_{k,\epsilon=\pm} \frac{1}{z_{k\epsilon}} \int d^d x \bar{\psi}_{k\epsilon}(x) (\cancel{\partial} + \epsilon M_k) \psi_{k\epsilon}(x). \quad (12.56)$$

In the same way, in the interaction, the fields ψ and $\bar{\psi}$ are replaced by the sums

$$\psi \mapsto \psi + \sum_{k,\epsilon} \psi_{k\epsilon}, \quad \bar{\psi} \mapsto \bar{\psi} + \sum_{k,\epsilon} \bar{\psi}_{k\epsilon}.$$

For a suitable choice of the constants z_k , after integration over the regulator fields, the form (12.55) is recovered. In particular, note that for $m = 0$ one finds $z_{k+} = z_{k-}$. This indicates how chiral symmetry is preserved by the regularization, although the regulators are massive: by *fermion doubling*. The fermions ψ_+ and ψ_- are chiral partners. For a pair $\psi \equiv (\psi_+, \psi_-)$, $\bar{\psi} \equiv (\bar{\psi}_+, \bar{\psi}_-)$ the action can be written as

$$\int d^d x \bar{\psi}(x) (\cancel{\partial} \otimes \mathbf{1} + M \mathbf{1} \otimes \sigma_3) \psi(x),$$

where the first matrix $\mathbf{1}$ and the Pauli matrix σ_3 act in \pm space. The spinors then transform like

$$\psi_\theta(x) = e^{i\theta\gamma_S \otimes \sigma_1} \psi(x), \quad \bar{\psi}_\theta(x) = \bar{\psi}(x) e^{i\theta\gamma_S \otimes \sigma_1},$$

because σ_1 anticommutes with σ_3 .

Determinants and regularization. A potential weakness of the momentum cut-off methods has again to be stressed: the generating functional of correlation function $\mathcal{Z}(J)$ obtained by adding to the action a source term for fields can be written, for instance, in the case of the self-coupled scalar field (expression (7.17)), as

$$\mathcal{Z}(J) = \det^{1/2}(\Delta_B) \exp[-\mathcal{V}_I(\delta/\delta J)] \exp\left(\frac{1}{2} \int d^d x d^d y J(x) \Delta_B(x-y) J(y)\right),$$

where the determinant is generated by the Gaussian integration. None of the cut-off methods, in the form explained previously, regularizes the determinant. As long as the determinant is a divergent constant that cancels in normalized correlation functions, this is not a problem but, in the case of a determinant in an external field (which generates a set of one-loop diagrams), this may become a serious issue. Such problems arise even in simple quantum mechanics ($d = 1$) in models that have divergences or ambiguities due to problem of order between quantum operators (see Section 3.3): a class of one-loop Feynman diagrams then cannot be regularized by this method. QFTs where the problem occurs include models with non-linear (Chapter 19) or gauge symmetries (Chapter 22).

A possible solution is to introduce additional regulator fields with the *wrong spin-statistics connection*, like bosons with Dirac fermion spin in the case of the fermion determinant (Section 21.8.1), or supersymmetry (see Chapter 27).

12.8 Dimensional regularization

For fermions belonging to the fundamental representation of the spin group $\text{Spin}(d)$, the spin problem can be reduced to the calculation of traces of γ matrices. Therefore, only an additional prescription for the trace of the unit matrix is needed. There is no natural continuation since odd and even dimensions behave differently. However, we show in Section A12.4 that no algebraic manipulation depends on the explicit value of the trace. Thus any smooth continuation in the neighbourhood of the relevant dimension can be used. A convenient choice, which we always adopt, is to take the trace constant. In even dimensions, as long as only γ_μ matrices are involved, no other problem arises.

The problem of γ_S . When a field theory involves the matrix γ_S , problems may appear, because no dimensional continuation preserves all properties of γ_S . This is the case when it becomes necessary to use the identity (see equation (A12.49))

$$d! \gamma_S = i^{-d/2} \sum_{\mu_1, \dots, \mu_d} \epsilon_{\mu_1 \dots \mu_d} \gamma_{\mu_1} \cdots \gamma_{\mu_d}, \quad (12.57)$$

where $\epsilon_{\mu_1 \dots \mu_d}$ is the completely antisymmetric symbol with $\epsilon_{12\dots d} = 1$. This difficulty is the source of *chiral anomalies*.

Since we have to calculate traces, one possibility is to define γ_S in terms of γ_μ matrices by a generalization of the expression in the initial dimension; for example, if we start from four dimensions we define $\gamma_5 \equiv \gamma_S$ in terms of a completely antisymmetric tensor $e_{\mu\nu\rho\sigma}$ by

$$4! \gamma_5 = - \sum_{\mu, \nu, \rho, \sigma} e_{\mu\nu\rho\sigma} \gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma.$$

It is easy then to verify that, with this definition, γ_5 anticommutes with all other γ_μ matrices in only four dimensions. If, for example, we start from dimension n (n even) and evaluate the product $\gamma_\nu \gamma_S \gamma_\nu$ in d dimensions, we find

$$\gamma_\nu \gamma_S \gamma_\nu = (d - 2n) \gamma_S,$$

instead of $-d \gamma_S$ if γ_S would anticommute with γ_μ . An alternative definition such that γ_S anticommutes with other γ matrices is inconsistent because it implies that, for generic dimensions, the trace of γ_5 with any product of γ_μ matrices vanishes.

Finally, another possibility, useful for supersymmetric theories, consists in keeping the spinors and γ -matrices of the initial dimension, and thus breaking $SO(d)$ rotation invariance.

12.9 Lattice fermions and the doubling problem

To regularize the free fermion action (12.19) by lattice methods, one can think about replacing $\bar{\psi}(x) \partial \psi(x)$ by

$$\bar{\psi}(x) \gamma_\mu \nabla_\mu^{\text{lat}} \psi(x) \equiv \bar{\psi}(x) \gamma_\mu [\psi(x + a n_\mu) - \psi(x - a n_\mu)] / 2a,$$

where a is the lattice spacing and n_μ the unit vector in the μ direction. In the Fourier representation, the Dirac operator becomes

$$D_W(p) = m + i \sum_\mu \gamma_\mu \frac{\sin ap_\mu}{a}. \quad (12.58)$$

A problem then arises: the equations relevant to the small lattice spacing limit,

$$\sin(a p_\mu) = 0,$$

have, for each value of μ , two solutions $p_\mu = 0$ and $p_\mu = \pi/a$, within one period, that is, within the Brillouin zone. Therefore, the propagator $\tilde{\Delta}_F(p) = D_W^{-1}(p)$ propagates 2^d fermions [95].

Wilson's fermions. The degeneracy can be removed by adding to the regularized action a contribution involving second derivatives, like

$$\delta\mathcal{S}(\bar{\psi}, \psi) = \frac{1}{2}M \sum_{x,\mu} [2\bar{\psi}(x)\psi(x) - \bar{\psi}(x+an_\mu)\psi(x) - \bar{\psi}(x)\psi(x+an_\mu)]. \quad (12.59)$$

After Fourier transformation, the regularized Dirac operator D_W reads

$$D_W(p) = m + M \sum_\mu (1 - \cos ap_\mu) + \frac{i}{a} \sum_\mu \gamma_\mu \sin ap_\mu. \quad (12.60)$$

The fermion propagator becomes

$$\Delta_F(p) = D_W^\dagger(p) \left[D_W(p) D_W^\dagger(p) \right]^{-1},$$

with

$$D_W(p) D_W^\dagger(p) = \left(m + M \sum_\mu (1 - \cos ap_\mu) \right)^2 + \frac{1}{a^2} \sum_\mu \sin^2 ap_\mu.$$

Therefore, the degeneracy between the different states has been removed. For each component p_μ which takes the value π/a , the mass is increased by M . If M is of order $1/a$, the spurious states are eliminated in the continuum limit. This is the recipe for *Wilson's fermions* [96].

The problem of chiral fermions. A new problem arises if one wants to construct a theory with massless fermions and preserve chiral symmetry, corresponding to the transformations (equations (12.32)),

$$\psi'(x) = e^{i\theta\gamma_5} \psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x) e^{i\theta\gamma_5}.$$

The Dirac operator \mathbf{D} must then anticommute with γ_5 :

$$\mathbf{D}\gamma_5 + \gamma_5\mathbf{D} = 0. \quad (12.61)$$

Thus, both the mass term and the term (12.59) are excluded. It remains possible to add various counter-terms and try to adjust them to recover chiral symmetry in the continuum limit. However, there is no *a priori* guarantee that this is indeed possible and, moreover, calculations may be plagued by fine-tuning problems or cancellations of unnecessary UV divergences.

One could also think about modifying the fermion propagator by adding terms connecting fermions separated by more than one lattice spacing. A chiral propagator then takes the general form

$$\Delta_F^{-1}(p) = \sum_\mu \gamma_\mu f_\mu(p),$$

in which $f_\mu(p)$ is a smooth function (singularities lead to a violation of locality), periodic with period $2\pi/a$, which vanishes linearly for $|p|$ small:

$$f_\mu(p) \sim p_\mu, \quad \text{for } |p| \rightarrow 0.$$

In one dimension, it is easy to understand that this modification does not solve the problem: if $f(p)$ is periodic and continuous, it has to vanish linearly an even number of times in each period. This argument can be generalized to any dimension [95]. This doubling of the number of fermion degrees of freedom is also related to the problem of anomalies (see Chapter 23).

Staggered fermions. Since the most naive form of the propagator yields 2^d fermion states, for $d > 1$ one can try to reduce this number to a smaller multiple of two. This is the idea of staggered fermions [97]: first, by modifying the action one is able to decrease the multiplication factor from 2^d to $2^{d/2}$. Then the remaining degeneracy can be interpreted as the reflection of an internal symmetry $SU(2^{d/2})$ of the action. The discussion is slightly involved and will not be given here.

12.9.1 Ginsparg–Wilson relation and overlap fermion

Notation. For convenience, we now set the lattice spacing $a = 1$, and use for the fields the notation $\psi(x) \equiv \psi_x$.

A different, and theoretically important, solution to the problem of chiral fermions has been found. It had been noted that a potential way to avoid fermion doubling, while still retaining chiral properties in the continuum limit, was to construct a lattice Dirac operator \mathbf{D} , satisfying, instead of condition (12.61), the Ginsparg–Wilson relation [98],

$$\mathbf{D}^{-1}\gamma_S + \gamma_S\mathbf{D}^{-1} = \gamma_S\mathbf{1}, \quad (12.62)$$

where $\mathbf{1}$ in the right-hand side means δ_{xy} , the Kronecker δ for lattice sites.

However, finding consistent solutions to the relation (12.62) is not straightforward, because the combined demands that \mathbf{D} and the anticommutator $\{\mathbf{D}^{-1}, \gamma_S\}_+$ should be local, are difficult to satisfy, especially in the most interesting case of a gauge-covariant operator in gauge theories (Chapter 25).

Let us briefly explain the main idea. Using the relation, a consequence of Euclidean Hermiticity and reflection symmetry

$$\mathbf{D}^\dagger = \gamma_S\mathbf{D}\gamma_S,$$

one can rewrite the relation (12.62) as

$$\mathbf{D}^{-1} + (\mathbf{D}^{-1})^\dagger = \mathbf{1} \Rightarrow \mathbf{D} + \mathbf{D}^\dagger = \mathbf{D}\mathbf{D}^\dagger = \mathbf{D}^\dagger\mathbf{D}.$$

This implies that the lattice operator \mathbf{D} has an index, and, in addition, the operator

$$\mathbf{S} = \mathbf{1} - \mathbf{D} \quad (12.63)$$

is unitary:

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{1}. \quad (12.64)$$

The eigenvalues of \mathbf{S} lie on the unit circle, the eigenvalue 1 corresponding to the pole of the Dirac propagator.

Moreover,

$$\gamma_S\mathbf{S} = \mathbf{S}^\dagger\gamma_S \Rightarrow (\gamma_S\mathbf{S})^2 = \mathbf{1}. \quad (12.65)$$

The matrix $\gamma_S\mathbf{S}$ is Hermitian and $\frac{1}{2}(\mathbf{1} \pm \gamma_S\mathbf{S})$ are two orthogonal projectors that, if \mathbf{D} is a Dirac operator in a gauge background, depend on the gauge field (Chapter 25).

An explicit solution can be derived from a Wilson–Dirac operator without doublers like \mathbf{D}_W in equation (12.60). Setting

$$\mathbf{A} = \mathbf{1} - \mathbf{D}_W,$$

one takes

$$\mathbf{S} = \mathbf{A} (\mathbf{A}^\dagger \mathbf{A})^{-1/2} \Rightarrow \mathbf{D} = \mathbf{1} - \mathbf{A} (\mathbf{A}^\dagger \mathbf{A})^{-1/2}. \quad (12.66)$$

With this ansatz \mathbf{D} has an eigenvalue 0 when $\mathbf{A} (\mathbf{A}^\dagger \mathbf{A})^{-1/2}$ has the eigenvalue 1. This can happen when \mathbf{A} and \mathbf{A}^\dagger have the same eigenvector with a *positive* eigenvalue. In the case of the Wilson–Dirac operator (12.60), a necessary condition is

$$\sin p_\mu = 0.$$

The presence of doublers thus depends on the value of the terms coming from second derivatives. By choosing $2M > 1$, one keeps the desired $p_\mu = 0$ mode, but eliminates all doublers, which then correspond to the eigenvalue 2 for \mathbf{D} , and the doubling problem is, in principle, solved [99].

It is then possible to construct lattice actions which have a chiral symmetry that corresponds to local, but non-point-like, transformations of the infinitesimal form [100],

$$\psi'_x = \psi_x + \sum_y (\boldsymbol{\theta} \gamma_S \mathbf{S})_{xy} \psi_y, \quad \bar{\psi}'_x = \bar{\psi}_x + \bar{\psi}_x \boldsymbol{\theta} \gamma_S, \quad (12.67)$$

where $\boldsymbol{\theta}$ commutes with \mathbf{D} . Note that the role of γ_S for $\bar{\psi}$ is played by $\gamma_S \mathbf{S}$ for ψ . Thus, $\frac{1}{2}(\mathbf{1} \pm \gamma_S \mathbf{S})\psi$ can be considered as the chiral components of ψ .

The problem is that these transformations no longer leave the integration measure over the fermion fields $\prod_x d\psi_x d\bar{\psi}_x$ automatically invariant. Indeed, the infinitesimal change of variables $\psi, \bar{\psi} \mapsto \psi', \bar{\psi}'$, lead to

$$\frac{\partial \psi'_x}{\partial \psi_y} = \delta_{xy} + \boldsymbol{\theta} \gamma_S \mathbf{S}_{xy}, \quad \frac{\partial \bar{\psi}'_x}{\partial \bar{\psi}_y} = \delta_{xy} (\mathbf{1} + \boldsymbol{\theta} \gamma_S),$$

and, thus, to a Jacobian given by

$$\ln J \sim \text{tr } \gamma_S \boldsymbol{\theta} \sum_x (2 - \mathbf{D}_{xx}).$$

This leaves the possibility of generating the expected anomalies (see Section 23.6), when the Wilson–Dirac operator is a covariant operator in the background of a gauge field.

A12 Euclidean fermions, spin group and γ matrices

We want to briefly describe here the formalism of Euclidean fermions with spin, analytic continuation to imaginary time of relativistic fermions (for more systematic presentations, see Refs. [92]). In Section 12.1, we have described the role of the spin group in real time. Some elements about Euclidean fermions have already been given in Section 12.3. As we have noticed, in the continuation to imaginary time, the relativistic pseudo-orthogonal group $O(1, d-1)$ transforms into the orthogonal group $O(d)$, d being the Euclidean space dimension. Similarly, Euclidean fermions transform under the spin group associated with the group $O(d)$.

The appendix is organized as follows: we first define an abstract Clifford algebra, show that it is invariant under $O(d)$ transformations, and use it to construct the spin group $\text{Spin}(d)$, a group that is only locally isomorphic to $SO(d)$: in fact, the group $SO(d)$ is an orthogonal representation of the spin group.

We then exhibit Hermitian matrices, generalizing Dirac γ matrices, which represent the algebra. A unitary representation of the spin group follows.

A section is devoted to the special example of dimension 4. Finally, we discuss the calculation of traces of products of γ -matrices, quantities that appear in perturbation theory with fermions, and define the Fierz transformation.

A12.1 Spin group. Dirac γ matrices

In this section, we begin with general considerations about the relation between Clifford algebras and orthogonal groups. We then construct an explicit representation of the Clifford algebra, where the generators are Hermitian matrices, generalizing Dirac γ matrices. Note that we always discuss even space dimensions first, because the situation is simpler, and then comment about the possible extension to odd dimensions.

A12.1.1 Clifford algebra. Orthogonal groups

Let γ_μ , $\mu = 1, \dots, d$, be the generators of an associative algebra on \mathbb{R} , satisfying the commutation relations

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2 \delta_{\mu\nu} \mathbf{1}, \quad (\text{A12.1})$$

where $\mathbf{1}$ is the unit element. They generate a Clifford algebra $\mathcal{C}(d)$ isomorphic to the algebra generated by the operators $(\theta_i + \partial/\partial\theta_i)$, in the notation of Section 1.6.1, acting on Grassmann algebras (see equation (1.49)).

It follows from the relations (A12.1) that the elements of $\mathcal{C}(d)$ form a real vector space of dimension 2^d , spanned by $\mathbf{1}$ and the products $\gamma_{\mu_1} \gamma_{\mu_2} \dots \gamma_{\mu_p}$, with $\mu_1 < \mu_2 < \dots < \mu_p$.

Automorphism. As in the example of Grassmann algebras, we define an automorphism P in $\mathcal{C}(d)$ by

$$P(\gamma_\mu) = -\gamma_\mu. \quad (\text{A12.2})$$

It splits $\mathcal{C}(d)$ into two vector spaces, $\mathcal{C}_-(d)$ and $\mathcal{C}_+(d)$, containing odd and even elements, respectively,

$$P(\mathcal{C}_\pm) = \pm \mathcal{C}_\pm,$$

where only $\mathcal{C}_+(d)$ is a subalgebra.

Clifford algebra and orthogonal group. We perform a linear transformation acting on the generators,

$$\gamma \mapsto \gamma', \quad \text{with} \quad \gamma'_\mu = \sum_\nu R_{\mu\nu} \gamma_\nu, \quad \det \mathbf{R} \neq 0, \quad (\text{A12.3})$$

where \mathbf{R} of elements $R_{\mu\nu}$ is a real matrix.

The elements γ' form an equivalent set of generators if they satisfy the relations (A12.1):

$$\gamma'_\mu \gamma'_\nu + \gamma'_\nu \gamma'_\mu = \sum_{\rho, \sigma} R_{\mu\rho} R_{\nu\sigma} (\gamma_\rho \gamma_\sigma + \gamma_\sigma \gamma_\rho) = 2 \sum_\rho R_{\mu\rho} R_{\nu\rho} \mathbf{1} = 2\delta_{\mu\nu} \mathbf{1}.$$

Therefore, the matrix \mathbf{R} must be orthogonal,

$$\sum_\rho R_{\mu\rho} R_{\nu\rho} = \delta_{\mu\nu}. \quad (\text{A12.4})$$

The relations (A12.1) are invariant under the orthogonal group $O(d)$ (rotations–reflections in Euclidean d -dimensional space).

Remark. If, in the right-hand side of equation (A12.1), the tensor $\delta_{\mu\nu}$ is replaced by another metric tensor $g_{\mu\nu}$, the symmetry group becomes the group that leaves the metric $g_{\mu\nu}$ invariant. In the case of the diagonal metric $+1, -1, \dots, -1$, one obtains the relativistic group $O(1, d-1)$, generalization of the Lorentz group $O(1, 3)$.

Product of all generators. In $\mathcal{C}(d)$, one element plays a special role, the product of all generators. We thus define

$$\tilde{\gamma} = \gamma_1 \gamma_2 \cdots \gamma_d. \quad (\text{A12.5})$$

Then,

$$\tilde{\gamma}^2 = (-1)^{[d/2]} \mathbf{1}, \quad (\text{A12.6})$$

where $[d/2]$ is the integer part of $d/2$.

An orthogonal transformation of matrix \mathbf{R} acting on the generators transforms $\tilde{\gamma}$ into $\tilde{\gamma} \det \mathbf{R}$.

If d is even, $\tilde{\gamma}$ commutes with the elements of \mathcal{C}_+ and anticommutes with those of \mathcal{C}_- , that is, for all elements c of $\mathcal{C}(d)$,

$$\tilde{\gamma}c = P(c)\tilde{\gamma}. \quad (\text{A12.7})$$

By contrast, if d is odd, $\tilde{\gamma}$ commutes with all elements c of $\mathcal{C}(d)$,

$$\tilde{\gamma}c = c\tilde{\gamma}. \quad (\text{A12.8})$$

Centre of the algebra. One verifies that, when d is even, the centre of the algebra, that is, the set of elements that commute with $\mathcal{C}(d)$, contains only $r\mathbf{1}$, $r \in \mathbb{R}$. When d is odd, the elements that commute with $\mathcal{C}(d)$ are linear combinations of $\mathbf{1}$ and $\tilde{\gamma}$.

Dimension d odd. Since $\tilde{\gamma}$ commutes with all elements of $\mathcal{C}(d)$, one can construct an algebra homomorphic to $\mathcal{C}(d)$ by imposing to the generators, in addition to the relations (A12.1), the relation

$$\tilde{\gamma} = i^{(d-1)/2} \mathbf{1}, \quad (\text{A12.9})$$

consistent with the identity (A12.6). This has several consequences:

(i) The symmetry group of the relations (A12.1) and (A12.9) is reduced to orthogonal matrices with determinant 1, thus belonging to the group $SO(d)$ (rotations) and the transformation P (equation (A12.2)) is no longer an automorphism.

(ii) If $d = 1 \pmod{4}$, then $\tilde{\gamma} = \pm \mathbf{1}$ and the algebra is isomorphic to $\mathcal{C}(d-1)$ (a real vector space of dimension 2^{d-1}). If $d = 3 \pmod{4}$, then $\tilde{\gamma} = \pm i\mathbf{1}$ and the algebra is isomorphic to the complexified form of $\mathcal{C}(d-1)$ (still a vector space of dimension 2^d on \mathbb{R}).

A12.1.2 Clifford algebra and group structure

We now consider the group $\mathfrak{G}(d)$ of invertible elements Λ of $\mathcal{C}(d)$ that satisfy

$$\Lambda^{-1}\gamma_\mu\Lambda = \sum_\nu R_{\mu\nu}\gamma_\nu, \quad (\text{A12.10})$$

where \mathbf{R} of coefficients $R_{\mu\nu}$ is a real matrix. This defining relation induces a homomorphism of groups. Indeed, if

$$\Lambda_1 \mapsto \mathbf{R}_1, \quad \Lambda_2 \mapsto \mathbf{R}_2,$$

to the product $\Lambda_1\Lambda_2$ corresponds the product of real matrices $\mathbf{R}_1\mathbf{R}_2$.

Moreover, the relation (A12.10) implies

$$\Lambda^{-1}\gamma_\mu\gamma_\nu\Lambda = \sum_{\rho,\sigma} R_{\mu\rho}R_{\nu\sigma}\gamma_\rho\gamma_\sigma.$$

Adding the symmetric relation $\mu \leftrightarrow \nu$, and using the relations (A12.1), one obtains

$$\sum_\rho R_{\mu\rho}R_{\nu\rho} = \delta_{\mu\nu}.$$

The real matrices \mathbf{R} thus form a group, subgroup of the orthogonal group $O(d)$ (rotation–reflection) of transformations (A12.3).

Remarks.

(i) If Λ corresponds to \mathbf{R} , then $\lambda\Lambda$, where λ is an invertible element of the centre, corresponds the same matrix \mathbf{R} . For d even, the centre reduces to $r \in \mathbb{R}^*$ ($r \neq 0 \in \mathbb{R}$).

(ii) If Λ belongs to the group $\mathfrak{G}(d)$ and is associated with the matrix \mathbf{R} , then a short calculation shows

$$\Lambda^{-1}\tilde{\gamma}\Lambda = \det \mathbf{R} \tilde{\gamma} \Leftrightarrow \tilde{\gamma}\Lambda = \det \mathbf{R} \Lambda \tilde{\gamma}. \quad (\text{A12.11})$$

Comparing with the two properties (A12.7) and (A12.8), we conclude:

For d even, even elements of $\mathfrak{G}(d)$ generate orthogonal matrices with determinant 1, that is, belonging to the subgroup $SO(d)$ (rotations) of $O(d)$. Instead, odd elements generate matrices with determinant -1 .

For d odd, $\tilde{\gamma}$ commutes with all other elements of $\mathcal{C}(d)$, and thus the relation (A12.11) always implies $\det \mathbf{R} = 1$. Reflections cannot be generated, and the orthogonal matrices belong to $SO(d)$.

(iii) The generators γ_ρ belong to the group $\mathfrak{G}(d)$, and correspond to the matrices \mathbf{R}^ρ :

$$R_{\mu\nu}^\rho = 2\delta_{\rho\mu}\delta_{\rho\nu} - \delta_{\mu\nu}, \quad (\text{A12.12})$$

which have determinant $(-1)^{d-1}$.

When the dimension d is even, the generators γ_μ provide examples of group elements associated with $O(d)$ matrices of determinant -1 .

(iv) The element $\tilde{\gamma}$ also belongs to the group $\mathfrak{G}(d)$ and, for d even, corresponds to the rotation matrix $-\mathbf{1}$, since from equation (A12.7),

$$(\tilde{\gamma})^{-1}\gamma_\mu\tilde{\gamma} = -\gamma_\mu. \quad (\text{A12.13})$$

Finally, since γ_μ is associated with the orthogonal matrix (A12.12), the product

$$\Pi_\mu = \tilde{\gamma}\gamma_\mu, \quad (\text{A12.14})$$

corresponds to a reflection P_μ along the μ axis, $x_\mu \mapsto -x_\mu$ (as defined by equation (12.24)). Note that

$$P_\mu^2 = \mathbf{1}, \quad \text{but} \quad (\Pi_\mu)^2 = (-1)^{d/2+1}\mathbf{1}.$$

A12.1.3 Spin group and Lie algebra

By explicit construction, we now show that the whole group $SO(d)$ can be generated by a subgroup of $\mathfrak{G}(d)$, the spin group $\text{Spin}(d)$, obtained by dividing $\mathfrak{G}(d)$ by Abelian factors.

It is easy to define a topology in the Clifford algebra, since it has the form of a finite-dimensional vector space. With such a topology, the groups $\mathfrak{G}(d)$ or $\text{Spin}(d)$ are Lie groups, and we can discuss their Lie algebras.

We consider the elements

$$\tilde{\sigma}_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]. \quad (A12.15)$$

Only $d(d - 1)/2$ elements $\tilde{\sigma}_{\mu\nu}$ are linearly independent. One can choose as a basis, for example, $\tilde{\sigma}_{\mu\nu}$, with $\mu < \nu$. Using the relations (A12.1), one verifies that for $\mu \neq \nu$,

$$\tilde{\sigma}_{\mu\nu}^2 = (\gamma_\mu \gamma_\nu)^2 = -\mathbf{1}.$$

To $\tilde{\sigma}_{\mu\nu}$ ($\mu \neq \nu$), we associate the elements

$$\begin{aligned} \Lambda(\theta) &= \exp[-\frac{1}{2}\theta \tilde{\sigma}_{\mu\nu}], \quad \theta \in \mathbb{R}, \\ &= \cos(\theta/2) \mathbf{1} - \sin(\theta/2) \tilde{\sigma}_{\mu\nu}. \end{aligned} \quad (A12.16)$$

The elements $\Lambda(\theta)$, for a given matrix $\tilde{\sigma}_{\mu\nu}$, generate a group isomorphic to the Abelian groups $U(1)$ or $SO(2)$, and correspond to rotations of angle $\theta/2$.

A straightforward calculation leads to

$$\Lambda^{-1}(\theta)\gamma_\rho\Lambda(\theta) = \begin{cases} \gamma_\rho & \text{for } \rho \neq \mu \text{ and } \rho \neq \nu, \\ \cos\theta\gamma_\mu - \sin\theta\gamma_\nu & \text{for } \rho = \mu, \\ \cos\theta\gamma_\nu + \sin\theta\gamma_\mu & \text{for } \rho = \nu. \end{cases}$$

Therefore, $\Lambda(\theta)$ is an element of $\mathfrak{G}(d)$ and the corresponding orthogonal matrix \mathbf{R} represents a rotation of angle θ leaving the space orthogonal to the (μ, ν) plane invariant.

The whole group $SO(d)$ can be generated by a product of such rotations.

Lie algebras and groups. We now introduce the generators of the Lie algebra of the group $SO(d)$ in the defining representation, $d \times d$ antisymmetric matrices $\mathbf{T}^{\rho\sigma}$ with elements

$$(\mathbf{T}^{\rho\sigma})_{\alpha\beta} = \delta_{\rho\alpha}\delta_{\sigma\beta} - \delta_{\rho\beta}\delta_{\sigma\alpha}, \quad (A12.17)$$

where only $d(d - 1)/2$ are independent. With this normalization of $\mathbf{T}^{\rho\sigma}$, if $\theta_{\rho\sigma}$ is an arbitrary antisymmetric matrix, then

$$\frac{1}{2} \sum_{\rho, \sigma} (\mathbf{T}^{\rho\sigma})_{\alpha\beta} \theta_{\rho\sigma} = \theta_{\alpha\beta}. \quad (A12.18)$$

A general matrix \mathbf{R} of $SO(d)$ can thus be written as

$$\mathbf{R} = e^\theta = \exp \left[\frac{1}{2} \sum_{\rho, \sigma} \mathbf{T}^{\rho\sigma} \theta_{\rho\sigma} \right]. \quad (A12.19)$$

Spin group. It follows from the homomorphism between groups that the matrices $\tilde{\sigma}_{\mu\nu}/2$ satisfy the commutation relations of the generators $\mathbf{T}^{\mu\nu}$ of the Lie algebra of the group $SO(d)$.

Exponentiating, one finds a general representation of the elements of the *spin group* $\text{Spin}(d)$, subgroup of the group $\mathfrak{G}(d)$,

$$\Lambda = \exp\left(\frac{1}{4} \sum_{\mu,\nu} \theta_{\mu\nu} \tilde{\sigma}_{\mu\nu}\right). \quad (\text{A12.20})$$

The only non-trivial element of the centre that is contained in the spin group is $-\mathbf{1}$, which corresponds in equation (A12.16) to $\theta = 2\pi$. Therefore, the spin group and $SO(d)$ are not isomorphic, since the two elements $\pm\Lambda$ of the spin group correspond to the same rotation matrix.

Finally, as we have shown previously, for d even, the addition of one reflection $\Pi_\mu = \tilde{\gamma}\gamma_\mu$ makes it possible to generate the whole $O(d)$ group.

Therefore, the transformations (A12.10) implement the transformations (A12.3) provided that, in the odd case, one restricts the Clifford algebra by the relation (A12.9).

A12.1.4 The γ matrices: A Hermitian representation

We now construct, inductively, an explicit representation of the Clifford algebra $\mathcal{C}(d)$ generated by Hermitian (and thus unitary) matrices of minimal size. We use *the same notation* γ_μ , $\mu = 1, \dots, d$ for the matrices representing the generators. First, we deal with even-dimensional spaces, and then generalize to odd-dimensional spaces, in the framework of the constraint (A12.9).

Space of even dimensions d . Since the dimension of $\mathcal{C}(d)$ is 2^d , the Clifford algebra cannot be represented by matrices of dimension smaller than $2^{d/2}$. We now give an inductive construction ($d \mapsto d+2$) of Hermitian matrices γ_μ satisfying the defining relations (A12.1).

For $d = 2$, the standard Pauli matrices realize the algebra $\mathcal{C}(2)$:

$$\gamma_1^{(d=2)} \equiv \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_2^{(d=2)} \equiv \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (\text{A12.21})$$

The third Pauli matrix σ_3 is proportional to the matrix $\tilde{\gamma}$ (equation (A12.5)). We define

$$\gamma_S^{(d=2)} \equiv \gamma_3^{(d=2)} = -i\tilde{\gamma}^{(d=2)} = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A12.22})$$

The three matrices are Hermitian:

$$\gamma_i = \gamma_i^\dagger.$$

The matrices γ_1 and γ_3 are symmetric, and γ_2 is antisymmetric:

$$\gamma_1 = \gamma_1^T, \quad \gamma_3 = \gamma_3^T, \quad \gamma_2 = -\gamma_2^T,$$

where we have denoted by γ^T the transpose of the matrix γ .

Quite generally, we define

$$\gamma_S^{(d)} \equiv \gamma_{d+1}^{(d)} = i^{-d/2} \tilde{\gamma}^{(d)}, \quad (\text{A12.23})$$

where $\tilde{\gamma}$ is the product of all generators (definition (A12.5)). The matrix γ_S then satisfies

$$\gamma_S^2 = \mathbf{1}, \quad \gamma_S \gamma_\mu + \gamma_\mu \gamma_S = 0, \quad \text{for } \mu \leq d. \quad (\text{A12.24})$$

Note that the matrix γ_S belongs to the representation of $\mathcal{C}(d)$ only for $d = 0 \pmod{4}$.

To construct the γ matrices for higher even dimensions, we then proceed by induction, setting

$$\gamma_i^{(d+2)} = \sigma_1 \otimes \gamma_i^{(d)} = \begin{pmatrix} 0 & \gamma_i^{(d)} \\ \gamma_i^{(d)} & 0 \end{pmatrix}, \quad 1 \leq i \leq d+1, \quad (A12.25)$$

$$\gamma_{d+2}^{(d+2)} = \sigma_2 \otimes \mathbf{1}_d = \begin{pmatrix} 0 & -i\mathbf{1}_d \\ i\mathbf{1}_d & 0 \end{pmatrix}, \quad (A12.26)$$

in which $\mathbf{1}_d$ is the unit matrix in $2^{d/2}$ dimensions.

From the definition (A12.23), it follows that $\gamma_S^{(d+2)}$ is then given by

$$\gamma_S^{(d+2)} \equiv \gamma_{d+3}^{(d+2)} = \sigma_3 \otimes \mathbf{1}_d = \begin{pmatrix} \mathbf{1}_d & 0 \\ 0 & -\mathbf{1}_d \end{pmatrix}. \quad (A12.27)$$

The matrices $\gamma_i^{(d+2)}$ are tensor products of the matrices $\gamma_i^{(d)}$ and $\mathbf{1}_d$ by the matrices σ_i . A straightforward calculation shows that if the matrices $\gamma_i^{(d)}$ satisfy relations (A12.1), the $\gamma_i^{(d+2)}$ matrices satisfy the same relations.

By inspection, we see that the γ matrices are all Hermitian. In addition,

$$\gamma_i^{(d+2)T} = \begin{pmatrix} 0 & \gamma_i^{(d)T} \\ \gamma_i^{(d)T} & 0 \end{pmatrix}, \quad 1 \leq i \leq d+1.$$

Therefore, if $\gamma_i^{(d)}$ is symmetric or antisymmetric, $\gamma_i^{(d+2)}$ has the same property. The matrix $\gamma_{d+2}^{(d+2)}$ is antisymmetric, and $\gamma_S^{(d+2)}$, which is also $\gamma_{d+3}^{(d+2)}$, is symmetric. It follows immediately that, in this representation, all γ matrices with an odd index are symmetric, all matrices with an even index are antisymmetric:

$$\gamma_i^T = (-1)^{i+1} \gamma_i, \quad 1 \leq i \leq d+1. \quad (A12.28)$$

Finally, the relations (A12.1) and (A12.24) can be summarized by

$$\gamma_i \gamma_j + \gamma_j \gamma_i = 2\delta_{ij} \mathbf{1}, \quad \text{for } i, j = 1, \dots, d, d+1. \quad (A12.29)$$

We will use Greek letters μ, ν and so on, to indicate that we exclude the value $(d+1)$ for the index.

Space of odd dimensions d. The relation (A12.29) shows that the set of γ matrices defined in even dimension $(d-1)$ together with the matrix $\gamma_S \equiv \gamma_d$ form a representation of the generators of the Clifford algebra.

As a consequence of the relation (A12.23), $\tilde{\gamma}$, which is the product of γ matrices, satisfies the relation (A12.9),

$$\tilde{\gamma} = \gamma_1 \gamma_2 \cdots \gamma_d = i^{d/2}.$$

A12.1.5 Spin group: A unitary representation

When the generators of the Clifford algebra are represented by Hermitian matrices, the generators $\tilde{\sigma}_{\mu\nu}$ of the spin group $\text{Spin}(d)$ are represented by anti-Hermitian, traceless matrices (for which we will use the same notation). The complex vectors on which the representation acts are called *spinors*.

Instead of the anti-Hermitian matrices $\tilde{\sigma}_{\mu\nu}$, one usually defines the Hermitian matrices

$$\sigma_{\mu\nu} = \frac{1}{i} \tilde{\sigma}_{\mu\nu} = \frac{1}{2i} [\gamma_\mu, \gamma_\nu]. \quad (\text{A12.30})$$

Then, for $\mu \neq \nu$, the matrices $\sigma_{\mu\nu}$ have the property

$$\sigma_{\mu\nu}^2 = \mathbf{1}. \quad (\text{A12.31})$$

The matrices belonging to the representation of the spin group $\text{Spin}(d)$ can be written as

$$\Lambda = \exp\left(\frac{i}{4} \sum_{\mu,\nu} \sigma_{\mu\nu} \theta_{\mu\nu}\right), \quad (\text{A12.32})$$

where $\theta_{\mu\nu}$ is a real antisymmetric matrix. Since Λ is the exponential of an anti-Hermitian traceless matrix, it is a unitary matrix of determinant 1: the representation of the spin group $\text{Spin}(d)$ is a unitary group, subgroup of the unitary group $SU(2^{[d/2]})$.

The relation (A12.10) can be written as

$$\Lambda \gamma_\mu \Lambda^\dagger = \sum_\nu \gamma_\nu R_{\nu\mu}. \quad (\text{A12.33})$$

If p_μ is a space vector, the equation can be written in an equivalent form as

$$\sum_\mu \Lambda p_\mu \gamma_\mu \Lambda^\dagger = \sum_\mu (p_R)_\mu \gamma_\mu, \quad \text{with} \quad (p_R)_\mu = \sum_\nu R_{\mu\nu} p_\nu. \quad (\text{A12.34})$$

In this form, the equation shows explicitly that the group $SO(d)$ is isomorphic to the adjoint representation of the spin group $\text{Spin}(d)$. As we have seen, the spin group and $SO(d)$ have the same Lie algebra but are not isomorphic, because $\pm \Lambda$ correspond to the same rotation matrix.

Note that the matrix $R_{\mu\nu}$ can be calculated explicitly from equation (A12.33) in terms of Λ by taking a trace:

$$R_{\mu\nu} = \text{tr} (\Lambda^\dagger \gamma_\mu \Lambda \gamma_\nu) / \text{tr } \mathbf{1}.$$

Examples. For $d = 2$, the spin group is isomorphic to a group $SO(2)$, but as we have seen, a rotation of angle $\theta/2$ in the spin group corresponds to a rotation of angle θ in the adjoint representation which is also isomorphic to $SO(2)$, a peculiarity of Abelian groups.

For $d = 3$, the group $SO(3)$ is associated with $SU(2)$, for $d = 4$, $SO(4)$ with $SU(2) \times SU(2)$.

A12.1.6 Reflections and chiral components

Reducibility. In even dimensions, from equation (A12.7) we infer that γ_S , which is proportional to $\tilde{\gamma}$, commutes with all transformations of the spin group:

$$[\Lambda, \gamma_S] = 0. \quad (\text{A12.35})$$

Therefore, the unitary representation of $\text{Spin}(d)$ can be reduced by projecting spinors ψ onto the two eigenspaces of γ_S using the projectors $(\mathbf{1} \pm \gamma_S)/2$. This defines two spinors ψ_{\pm} , the chiral components of the spinor ψ :

$$\psi_{\pm} = \frac{1}{2}(\mathbf{1} \pm \gamma_S)\psi. \quad (\text{A12.36})$$

Space reflections and chiral components. To obtain the full orthogonal group, we have still to represent reflection transformations. We have seen that this can be achieved with elements of the Clifford algebra only in spaces of even dimensions. Then, the elements $\pm \Pi_{\mu}$, ($\Pi_{\mu} = \tilde{\gamma}\gamma_{\mu}$, see equation (A12.14)) correspond to reflections P_{μ} (equation (12.24)) that act on a position x by changing the sign of the coordinate x_{μ} . The anticommutation relation,

$$\gamma_S \Pi_{\mu} = -\gamma_S \Pi_{\mu}, \quad (\text{A12.37})$$

implies

$$\Pi_{\mu} \frac{1}{2}(1 + \gamma_S) = \frac{1}{2}(1 - \gamma_S)\Pi_{\mu}. \quad (\text{A12.38})$$

A reflection exchanges chiral components. The representation of the spin group associated with the group $O(d)$ is thus irreducible.

Remark. In odd dimensions, the whole group $O(d)$ factorizes into $SO(d) \times Z_2$, because $-\mathbf{1}$ is a reflection matrix. The transformations of spinors corresponding to $-\mathbf{1}$ can thus be represented by external transformations that commutes with the whole spin group.

A12.1.7 Charge conjugation

We now exhibit unitary matrices C such that

$$C^{\dagger} \gamma_{\mu}^T C = -\gamma_{\mu} \quad \Leftrightarrow \quad C \gamma_{\mu} C^{\dagger} = -\gamma_{\mu}^T. \quad (\text{A12.39})$$

Note that since the matrices γ_{μ} are Hermitian $\gamma_{\mu}^T = \gamma_{\mu}^*$.

Even dimensions. In the representation of Section A12.1.4, for d even we can take

$$C = \pm (\tilde{\gamma})^{d/2} \prod_{\text{all } \mu \text{ odd}} \gamma_{\mu}, \quad \Rightarrow C^{\dagger} C = \mathbf{1}. \quad (\text{A12.40})$$

For example,

$$\text{for } d = 2, \quad C = \sigma_2, \quad \text{for } d = 4, \quad C = \gamma_1 \gamma_3.$$

Note that the unitary matrix

$$\tilde{C} = C \tilde{\gamma}, \quad (\text{A12.41})$$

then satisfies

$$\begin{aligned} \tilde{C}^{\dagger} \gamma_{\mu}^T \tilde{C} &= \tilde{\gamma}^{\dagger} C^{\dagger} \gamma_{\mu}^T C \tilde{\gamma} = -\tilde{\gamma}^{\dagger} \gamma_{\mu} \tilde{\gamma} \\ &= \gamma_{\mu}. \end{aligned} \quad (\text{A12.42})$$

The matrix γ_S is symmetric. Under C or \tilde{C} , it transforms as

$$C^{\dagger} \gamma_S^T C = (-1)^{d/2} \gamma_S.$$

Odd dimensions. In odd dimensions, γ_S becomes γ_d . We see that the property (A12.39) of the matrix C extends to dimensions $d = 3 \pmod{4}$, but not $d = 1 \pmod{4}$. The converse is true for the matrix \tilde{C} (equation (A12.42)).

Spin group: Conjugate representation. We now apply the transformation (A12.39) to an element Λ of the unitary representation of the spin group, using the form (A12.32),

$$C \exp \left(\sum_{\mu,\nu} \frac{1}{8} \theta_{\mu\nu} [\gamma_\mu, \gamma_\nu] \right) C^\dagger = \exp \left(\frac{1}{8} \sum_{\mu,\nu} \theta_{\mu\nu} [\gamma_\mu^*, \gamma_\nu^*] \right),$$

and thus

$$\Lambda^* = C \Lambda C^\dagger.$$

This identity shows that the unitary representation and the representation obtained by complex conjugation are equivalent. The same property holds for the matrix \tilde{C} and this extends the property to the odd dimensions, in which C does not exist.

Finally, for d even,

$$C \tilde{\gamma} C^\dagger = \tilde{\gamma}^*,$$

and this extends the property to reflections.

These transformations are called *charge conjugation* (see Section 12.3.3).

Charge conjugation and chiral components. The transformation properties of γ_S imply

$$\begin{cases} C^\dagger (1 + \gamma_S^T) C = (1 + \gamma_S) & \text{for } d = 0 \pmod{4}, \\ C^\dagger (1 + \gamma_S^T) C = (1 - \gamma_S) & \text{for } d = 2 \pmod{4}. \end{cases}$$

If the dimension is a multiple of 4, C respects chirality, otherwise it exchanges chiral components. Charge conjugation multiplied by a reflection like $C\Pi_\mu$ has the opposite property.

A12.2 The example of dimension 4

Since the dimension 4 plays a special role in physics, we specialize some of the previous results to the example of the spin group $\text{Spin}(4)$, which is homomorphic to $SO(4)$. With the conventions of Section A12.1.4, the 4×4 γ matrices take the form

$$\gamma_{i=1,2,3} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \gamma_4 = \begin{pmatrix} 0 & -i\mathbf{1}_2 \\ i\mathbf{1}_2 & 0 \end{pmatrix}$$

(σ_i are the three Pauli matrices) and $\gamma_S \equiv \gamma_5$,

$$\gamma_5 = -\tilde{\gamma} = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix}.$$

The matrices $\sigma_{\mu\nu}$ then become

$$\sigma_{ij} = \sum_k \epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \text{ for } i, j, k \leq 3, \quad \sigma_{i4} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \text{ for } i \leq 3.$$

We recognize in the matrices

$$\sigma_i^\pm = \frac{1}{4} \sum_{j,k} \epsilon_{ijk} \sigma_{jk} \pm \frac{1}{2} \sigma_{i4},$$

the generators of the group $SU(2) \times SU(2)$. The projectors $\frac{1}{2}(1 \pm \gamma_5)$ decompose a Dirac spinor into the sum of two vectors transforming as the $(1/2, 0)$ and $(0, 1/2)$ representations of the group (Weyl spinors).

Note that a reflection exchanges the two vectors (as expected, since the representation then is no longer reducible). In terms of Weyl spinors, the construction of invariants with respect to the spinor group reduces to considerations about $SU(2)$. A useful remark in this context, is that the representation and its complex conjugate are equivalent, since

$$U^* = \sigma_2 U \sigma_2, \quad \forall U \in SU(2),$$

(see also Section A12.1.7) and thus if φ and χ are two $SU(2)$ spinors the combination,

$$\sum_{\alpha,\beta} \varphi_\alpha (\sigma_2)_{\alpha\beta} \chi_\beta = -i \sum_{\alpha,\beta} \epsilon_{\alpha\beta} \varphi_\alpha \chi_\beta,$$

where $\epsilon_{\alpha\beta}$ is the antisymmetric tensor ($\epsilon_{12} = 1$), is an $SU(2)$ invariant.

We recall that for charge conjugation we can take $\gamma_1 \gamma_3 = \mathbf{1}_2 \otimes \sigma_2$.

A12.3 The Fierz transformation

Within the algebra of γ matrices, it is possible to define a basis of 2^d Hermitian matrices orthogonal by the trace. We denote these matrices by Γ^A . Any fermion two-point correlation function can then be expanded on such a basis. A four-point fermion correlation function can be expanded on a basis formed by the tensor products of these matrices. However, in this case one has to first separate the four fermion fields into two pairs of fields, and there are three ways of doing it. A connection between these different bases can be found through a Fierz transformation. We use the property that any $2^{d/2} \times 2^{d/2}$ matrix \mathbf{X} can be expanded on the basis of Γ matrices:

$$X_{ab} = N^{-1} \sum_A \text{tr } \mathbf{X} \Gamma^A \Gamma_{ab}^A, \quad (A12.43)$$

in which N is the trace of $\mathbf{1}$. We choose a matrix \mathbf{X} of the form

$$X_{ab} = \Gamma_{cb}^B \Gamma_{ad}^C. \quad (A12.44)$$

The identity (A12.43) becomes

$$\Gamma_{cb}^B \Gamma_{ad}^C = N^{-1} \sum_A (\Gamma^B \Gamma^A \Gamma^C)_{cd} \Gamma_{ab}^A. \quad (A12.45)$$

By expanding the product $\Gamma^B \Gamma^A \Gamma^C$ on the basis of Γ matrices,

$$\Gamma^B \Gamma^A \Gamma^C = N \sum_D M_{AD}^{BC} \Gamma^D,$$

we obtain the decomposition of any element of one basis onto another:

$$\Gamma_{cb}^B \Gamma_{ad}^C = \sum_{A,D} M_{AD}^{BC} \Gamma_{cd}^D \Gamma_{ab}^A.$$

Examples.

(i) For $d = 2$, a basis is $\mathbf{1}$ and σ_i . We leave as an exercise for the reader to verify that the subset $\mathbf{1} \otimes \mathbf{1}$ and $\sigma_i \otimes \sigma_i$ transforms into itself with the matrix

$$\mathbf{M}_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix}.$$

As expected the square of the matrix \mathbf{M}_2 is the unit matrix.

(ii) For $d = 4$ a basis is

$$\mathbf{1}, \gamma_\mu, \gamma_S, i\gamma_5\gamma_\mu, \boldsymbol{\sigma}_{\mu\nu}.$$

We also leave as an exercise for the reader to verify that the subset

$$\mathbf{1} \otimes \mathbf{1}, \gamma_\mu \otimes \gamma_\mu, \gamma_S \otimes \gamma_S, i\gamma_S\gamma_\mu \otimes i\gamma_S\gamma_\mu, \boldsymbol{\sigma}_{\mu\nu} \otimes \boldsymbol{\sigma}_{\mu\nu},$$

transforms into itself with the matrix of square $\mathbf{1}$,

$$\mathbf{M}_4 = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 4 & -2 & -4 & 2 & 0 \\ 1 & -1 & 1 & -1 & 1 \\ 4 & 2 & -4 & -2 & 0 \\ 6 & 0 & 6 & 0 & -2 \end{pmatrix}.$$

A12.4 Traces of products of γ matrices

Perturbative calculations involving relativistic fermions often require the calculation of traces of products of γ matrices, which, therefore, we explain in detail. It is possible to calculate traces within an explicit matrix representation, but here we define the trace as a linear mapping of the Clifford algebra (A12.1) to real or complex numbers (to account for all dimensions) which satisfies the cyclic condition. In addition, we normalize the trace by the value of the trace of the unit matrix, the only quantity that depends explicitly on the representation. We define

$$\text{tr } \mathbf{1} = N, \quad N \in \mathbb{R}_+. \quad (\text{A12.46})$$

Since we work in the framework of the relation (A12.9), we can restrict the analysis to *even dimensions*.

Traces of odd elements. We first consider odd elements, which are such that $\text{P}(c) = -c$ (definition (A12.2)). By using the relations (A12.1) they can be written as linear combinations of basis vectors $\gamma_{\mu_1}\gamma_{\mu_2} \cdots \gamma_{\mu_{2p+1}}$, with $\mu_1 < \mu_2 < \cdots < \mu_{2p+1}$, and it is thus sufficient to calculate the traces of these vectors that we denote by Γ^A . Since d is even, at least one generator γ_μ is absent from the product Γ^A . Then,

$$\gamma_\mu \Gamma^A + \Gamma^A \gamma_\mu = 0.$$

Using $\gamma_\mu^2 = \mathbf{1}$, this anticommutation relation and the cyclic property of the trace, we find the chain of identities,

$$\text{tr } \Gamma^A = \text{tr } \gamma_\mu^2 \Gamma^A = -\text{tr } \gamma_\mu \Gamma^A \gamma_\mu = -\text{tr } \Gamma^A \Rightarrow \text{tr } \Gamma^A = 0.$$

Therefore, the traces of all odd elements of the Clifford algebra vanish.

Product of even numbers of generators γ_μ . To calculate the trace of the product of an even number $2n$ of generators, $\text{tr } \gamma_{\mu_1} \cdots \gamma_{\mu_{2n}}$, we successively commute $\gamma_{\mu_{2n}}$ through all other factors $\gamma_{\mu_1}, \dots, \gamma_{\mu_{2n-1}}$, using the commutation relations (A12.1). We then generate a linear combination of traces of the products of $(2n-2)$ generators. At each commutation the sign changes. After all commutations, as a consequence of the cyclic property of the trace, we recover the opposite of the initial expression. As a consequence, we find

$$\begin{aligned} \text{tr } \gamma_{\mu_1} \cdots \gamma_{\mu_{2n}} &= \delta_{\mu_1 \mu_{2n}} \text{tr} (\gamma_{\mu_2} \cdots \gamma_{\mu_{2n-1}}) - \delta_{\mu_2 \mu_{2n}} \text{tr} (\gamma_{\mu_1} \gamma_{\mu_3} \cdots \gamma_{\mu_{2n-1}}) \\ &\quad + \cdots + \delta_{\mu_{2n-1} \mu_{2n}} \text{tr} (\gamma_{\mu_1} \cdots \gamma_{\mu_{2n-2}}). \end{aligned} \quad (A12.47)$$

Therefore, we prove by induction Wick's theorem for the trace of a product of an even number of generators γ_μ :

$$\text{tr } \gamma_{\mu_1} \cdots \gamma_{\mu_{2n}} = N \sum_{\substack{\text{all possible pairings} \\ \text{of } (1, 2, \dots, 2n)}} \text{sgn}(P) \delta_{\mu_{P_1} \mu_{P_2}} \cdots \delta_{\mu_{P_{2n-1}} \mu_{P_{2n}}}, \quad (A12.48)$$

in which $\text{sgn}(P)$ is the signature of the permutation P when $P_{2m-1} < P_{2m}$ for $1 \leq m \leq n$.

The element γ_S . The calculations of traces of products of the form $\text{tr } \gamma_S \gamma_{\mu_1} \cdots \gamma_{\mu_{2n}}$ (the trace vanishes for an odd number of γ_μ generators) are of direct interest in even dimensions and, because $\gamma_S = \gamma_{d+1}$, are also of interest for odd dimensions. From Wick's theorem, it follows immediately that

$$\text{tr } \gamma_S \gamma_{\mu_1} \cdots \gamma_{\mu_{2n}} = 0, \text{ for } 2n < d.$$

For $2n = d$, from $\gamma_S = i^{-d/2} \tilde{\gamma}$, one infers

$$\text{tr } \gamma_S \gamma_{\mu_1} \cdots \gamma_{\mu_d} = N i^{d/2} \epsilon_{\mu_1 \cdots \mu_d}, \quad (A12.49)$$

in which $\epsilon_{\mu_1 \cdots \mu_d}$ is the completely antisymmetric tensor normalized by

$$\epsilon_{12 \dots d} = 1. \quad (A12.50)$$

The relation (A12.49), which depends explicitly on the number of space dimensions, has deep consequences. In particular, dimensional regularization does not preserve this relation and this is the source of possible *anomalies* in field theories that are chiral invariant in the classical approximation (see Section 23.6).

13 Symmetries, chiral symmetry breaking, and renormalization

Most quantum field theories (QFTs) with physics applications exhibit symmetries, exact symmetries or symmetries with soft (*e.g.* linear) breaking. A particularly interesting situation corresponds to *spontaneous symmetry breaking* (SSB), a phenomenon encountered in the ordered phase in phase transitions (see, *e.g.* Chapter 14), and in particle physics.

In this chapter, we consider *linear continuous symmetries* corresponding to *compact Lie groups*; consequences of discrete symmetries can also be studied but with somewhat different methods (see the remark at the end of Section 13.8).

To construct a QFT with a symmetry, one first exhibits a classical, or tree order, action with the symmetry, and then tries to quantize it, for example, using the field integral formalism. However, such a straightforward construction generates a perturbative expansion with ultraviolet (UV) divergences. As we have discussed in Chapter 9, it is necessary to regularize the quantum theory, for example, by modifying its short distance structure. When the initial action is symmetrical, it is convenient to introduce a symmetrical regularization.

It is then possible to study the symmetry of the regularized theory. This can be achieved by deriving relations between correlation functions, called Ward–Takahashi (WT) identities, which describe the physical consequences of the symmetry. Their derivation involves only infinitesimal group transformations and, therefore, group transformations continuously connected to the identity. The usual Legendre transformation then leads to relations between vertex (or 1PI) functions.

These relations can be used to determine the specific structure of divergences beyond simple power counting, and thus of the counter-terms that render the QFT finite. This analysis is based on a loop expansion. The procedure generates a renormalized action, whose structure from the viewpoint of symmetry can be analysed. For the QFTs discussed in this chapter, a symmetrical regularization can always be found, but this is not always possible, like in the example of chiral gauge theories (see Section 23.6). Moreover, the general form of the renormalized action in the case of the addition of *soft symmetry breaking terms* (relevant contributions in the terminology of renormalization group, see Chapter 15), or in the case of SSB, is not intuitive.

Note that, in some examples, the regularized theory may be less symmetrical than the initial theory, without affecting the symmetry of the renormalized theory. A simple example is provided by lattice regularization with hypercubic symmetry, which leads to a continuum renormalized theory with translation and rotation symmetry. More generally, this happens when the symmetry of the renormalizable part of the action is automatically larger than the symmetry of the complete regularized action.

In this chapter, we specifically discuss the examples of linear symmetry breaking and the very important limiting case of SSB, and quadratic symmetry breaking.

Finally, in Sections 13.5–13.7, we apply the formalism to study the physics implications of broken chiral symmetry in low-energy effective models of hadron physics.

In the appendix, we outline the relation between WT identities and current conservation. We derive the expressions of the energy–momentum tensor and dilatation current.

13.1 Lie groups and algebras: Preliminaries

We first describe our notation and conventions for orthogonal and unitary matrix representations of compact Lie groups and algebras.

13.1.1 Orthogonal and unitary representations: Conventions and notation

Orthogonal representations. Let ϕ be a vector with N real components ϕ_i , which transforms under an orthogonal representation $\mathcal{R}(G)$ of a Lie group G as

$$\phi \mapsto \phi_R \text{ with } \phi_R = \mathbf{R}(\mathbf{g})\phi, \quad \Leftrightarrow \quad [\phi_R]_i = \sum_j R_{ij}(\mathbf{g})\phi_j, \text{ for } \mathbf{g} \in G, \quad (13.1)$$

where the $N \times N$ matrix \mathbf{R} satisfies $\mathbf{R}\mathbf{R}^T = \mathbf{1}$ (T means transposed).

A Lie algebra $\mathcal{L}(G)$, whose generators can be represented by real antisymmetric $N \times N$ matrices \mathbf{t}^α with elements t_{ij}^α , corresponds to the group G . The trace of the product of two antisymmetric matrices defines a scalar product. It can be used to normalize the matrices by

$$\mathrm{tr} \mathbf{t}^\alpha \mathbf{t}^\beta = -N\delta_{\alpha\beta}. \quad (13.2)$$

The structure constants $f_{\alpha\beta\gamma}$ of the Lie algebra are defined by the commutation relations

$$[\mathbf{t}^\alpha, \mathbf{t}^\beta] = \sum_\gamma f_{\alpha\beta\gamma} \mathbf{t}^\gamma. \quad (13.3)$$

Then, the structure constants are given by

$$f_{\alpha\beta\gamma} = \frac{2}{N} \mathrm{tr}(\mathbf{t}^\gamma \mathbf{t}^\beta \mathbf{t}^\alpha),$$

and are completely antisymmetric in the three indices. The basis of the Lie algebra is fixed up to an orthogonal transformation.

Unitary representations. In the case of a unitary representation by $N \times N$ complex matrices, we can represent the generators of the Lie algebra by Hermitian or anti-Hermitian matrices (this is a matter of convenience). Choosing anti-Hermitian matrices, we can normalize them by

$$\mathrm{tr} \mathbf{t}^\alpha \mathbf{t}^\beta = -N\delta_{\alpha\beta}.$$

Then, as a consequence, as in the orthogonal case, the structure constants defined by

$$[\mathbf{t}^\alpha, \mathbf{t}^\beta] = \sum_\gamma f_{\alpha\beta\gamma} \mathbf{t}^\gamma,$$

and, thus, given by

$$f_{\alpha\beta\gamma} = \frac{2}{N} \mathrm{Re} \mathrm{tr} (\mathbf{t}^\gamma \mathbf{t}^\beta \mathbf{t}^\alpha),$$

are again completely antisymmetric.

The $SU(2)$ example. The generators of the group $SU(2)$ of 2×2 unitary matrices with determinant 1 are the Hermitian Pauli matrices (defined by equations (A12.21) and (A12.22)), which satisfy the commutation relations

$$[\boldsymbol{\sigma}_\alpha, \boldsymbol{\sigma}_\beta] = 2i \sum_\gamma \epsilon_{\alpha\beta\gamma} \boldsymbol{\sigma}_\gamma,$$

where $\epsilon_{\alpha\beta\gamma}$ is the totally antisymmetric tensor with $\epsilon_{123} = 1$.

Adjoint representation. Casimir operator. Structure constants satisfy the *Jacobi identity* in the form

$$\sum_\delta [f_{\alpha\beta\delta} f_{\delta\gamma\epsilon} + f_{\beta\gamma\delta} f_{\delta\alpha\epsilon} + f_{\gamma\alpha\delta} f_{\delta\beta\epsilon}] = 0.$$

The Jacobi identity implies that the matrices \mathbf{F}^α with elements $F_{\beta\gamma}^\alpha = f_{\alpha\beta\gamma}$, form a representation of the Lie algebra, called *an adjoint representation*.

An element of the enveloping algebra, $\sum_\alpha \mathbf{t}^\alpha \mathbf{t}^\alpha$, which commutes with all generators and is called *Casimir operator*, corresponds to each simple component of the Lie algebra.

Infinitesimal transformations. In this chapter, we mainly make use of infinitesimal group transformations. We parametrize a group element \mathbf{R} in $\mathcal{R}(G)$ close to the identity as $\mathbf{R} = \mathbf{1} + \boldsymbol{\omega} + O(|\boldsymbol{\omega}|^2)$, with $\boldsymbol{\omega} \in \mathcal{L}(G)$. The corresponding infinitesimal variation of a vector ϕ , at first order in $\boldsymbol{\omega}$, is

$$\delta\phi \equiv \phi_R - \phi = \boldsymbol{\omega}\phi \equiv \sum_\alpha \omega_\alpha \mathbf{t}^\alpha \phi,$$

(the real constants ω_α parametrize the element of the Lie algebra) or, in component form,

$$\delta\phi_i = \sum_{\alpha,j} \omega_\alpha t_{ij}^\alpha \phi_j. \quad (13.4)$$

Remark. In the following examples, for generic arguments we use only orthogonal representations, because unitary matrices in $U(N)$ have a faithful representation as orthogonal matrices in $O(2N)$.

13.1.2 Lie algebra and differential operators

We establish now a property that is particularly useful when we discuss the renormalization of symmetries in complicated situations. The variation of a differentiable function $\mathcal{S}(\phi)$ under an infinitesimal transformation (13.4) is

$$\delta\mathcal{S}(\phi) = \sum_\alpha \omega_\alpha \mathcal{D}^\alpha \mathcal{S}(\phi),$$

with

$$\mathcal{D}_\alpha = \sum_{i,j} t_{ij}^\alpha \phi_j \frac{\partial}{\partial \phi_i}. \quad (13.5)$$

In particular, an invariant function $\mathcal{S}(\phi)$ satisfies

$$\delta\mathcal{S}(\phi) = 0 \Rightarrow \mathcal{D}^\alpha \mathcal{S}(\phi) = 0, \quad (13.6)$$

because the first equation is valid for any set of parameters ω_α .

The operators \mathcal{D}^α are generators of the Lie algebra of the group G realized as differential operators acting on functions of ϕ_i . The expected commutation relations

$$[\mathcal{D}^\alpha, \mathcal{D}^\beta] = \sum_\gamma f_{\alpha\beta\gamma} \mathcal{D}^\gamma, \quad (13.7)$$

also follow directly from the commutation relations (13.3) of the generators \mathbf{t}^α .

Conversely, equation (13.6) implies that the function \mathcal{S} is invariant under all group transformations corresponding to group elements continuously connected to the identity.

Compatibility conditions. Equation (13.6), can also be viewed as a system of differential equations for $\mathcal{S}(\phi)$. Quite generally, the commutators of first order differential operators are again first order differential operators. Therefore, if a function is a solution of a system of first order partial differential equations described in terms of operators \mathcal{D}_α , it is also a solution of all equations corresponding to operators belonging to the Lie algebra generated by \mathcal{D}_α . The system (13.6) is said to be compatible if no new independent equation is obtained from the commutators $[\mathcal{D}_\alpha, \mathcal{D}_\beta]$. This condition is verified if all commutators are linear combinations of the operators \mathcal{D}_α , that is, if the \mathcal{D}_α form a basis of the Lie algebra they generate. Therefore, the Lie algebra commutation relations (13.7) are the compatibility conditions of the linear system (13.6).

13.2 Linear global symmetries and WT identities

We shall be mostly concerned with situations in which ϕ is a field depending on space coordinates x in d dimensions, and \mathcal{S} a classical action, local functional of ϕ .

Definition. In the context of QFT, we call *global symmetry* a symmetry which corresponds to a transformation of the fields whose parameters are space independent. More precisely, let $\phi(x)$ be an N -component field transforming linearly under a representation $\mathcal{R}(G)$ of a compact Lie group G ,

$$\phi_R(x) = \mathbf{R}(\mathbf{g})\phi(x). \quad (13.8)$$

The transformation (13.8) is *global* if the group element \mathbf{g} does not depend on the space coordinates x . Sometimes, the expression *rigid symmetry* is also used to avoid confusions with ‘global’ in the sense of global topological properties of the symmetry group.

From field integrals to WT identities. We consider the generating functional $\mathcal{Z}(\mathbf{J})$ of correlation functions, corresponding to a symmetric action $\mathcal{S}(\phi)$ ($\mathcal{S}(\mathbf{R}\phi) = \mathcal{S}(\phi)$),

$$\mathcal{Z}(\mathbf{J}) = \int [d\phi] \exp [-\mathcal{S}(\phi, \mathbf{J})], \quad (13.9)$$

with

$$\mathcal{S}(\phi, \mathbf{J}) = \mathcal{S}(\phi) - \int d^d x \mathbf{J}(x) \cdot \phi(x). \quad (13.10)$$

We change variables, $\phi \mapsto \mathbf{R}\phi$. The action is invariant. For each point of space, the Jacobian is $|\det(\mathbf{R})|$, which equals 1, because \mathbf{R} is an orthogonal matrix. However, as such the argument is not completely rigorous, because the product of constants on all points is not defined. A regularization by using a lattice, or by point splitting (see equation (13.22)), closes this loophole. Only the source term is affected. We conclude [101],

$$\mathcal{Z}(\mathbf{J}) = \mathcal{Z}(\mathbf{R}^T \mathbf{J}) \Rightarrow \mathcal{W}(\mathbf{J}) = \mathcal{W}(\mathbf{R}^T \mathbf{J}). \quad (13.11)$$

It is then simple to verify that the Legendre transformation implies

$$\Gamma(\boldsymbol{\varphi}) = \Gamma(\mathbf{R}\boldsymbol{\varphi}). \quad (13.12)$$

Regularization. In the case of linearly realized global symmetries, it is always possible to find a regularization that preserves the symmetry of the action. For QFTs involving only scalar boson fields, we can use momentum cut-off, lattice, or dimensional regularizations. A suitable momentum regularization amounts to adding to the action $\mathcal{S}(\phi)$ quadratic invariant terms involving higher order derivatives like (see expression (8.22)):

$$\sum_i \phi_i(x) (-\nabla^2 + m^2) \phi_i(x) \mapsto \sum_i \phi_i(x) (-\nabla^2 + m^2) \prod_{r=1}^{r_{\max}} (1 - \nabla^2 / M_r^2) \phi_i(x), \quad (13.13)$$

in which $M_r = \alpha_r \Lambda$, with $\alpha_r > 0$, and Λ is the momentum cut-off. For r_{\max} large enough ($2r_{\max} > (d-2)$, in the absence of derivative couplings), the theory becomes finite. The regularization terms are symmetric, since they are invariant under general orthogonal transformations.

Infinitesimal transformations. In the notation of equation (13.4), the variation $\delta_\omega\phi$ of ϕ under a transformation (13.8) infinitely close to the identity reads

$$\delta_\omega\phi_i(x) = \sum_{j,\alpha} \omega_\alpha t_{ij}^\alpha \phi_j(x), \quad (13.14)$$

in which ω_α are space-independent infinitesimal parameters.

A classical action $\mathcal{S}(\phi)$ invariant under such a transformation then satisfies

$$\mathcal{D}_\alpha \mathcal{S}(\phi) \equiv \int d^d x \sum_{i,j} t_{ij}^\alpha \phi_j(x) \frac{\delta \mathcal{S}}{\delta \phi_i(x)} = 0. \quad (13.15)$$

One verifies that the operators \mathcal{D}_α still satisfy the commutation relations (13.7).

13.2.1 WT identities

Since $\mathbf{R}^T = \mathbf{R}^{-1}$ is an arbitrary group element, the symmetry relations (13.11) for infinitesimal group transformations also imply

$$\int d^d x \sum_{i,j} t_{ij}^\alpha J_i(x) \frac{\delta \mathcal{Z}(J)}{\delta J_j(x)} = 0. \quad (13.16)$$

Equation (13.16) immediately implies an identical equation for the generating functional $\mathcal{W}(J) = \ln \mathcal{Z}(J)$ of connected correlation functions,

$$\int d^d x \sum_{i,j} t_{ij}^\alpha J_i(x) \frac{\delta \mathcal{W}(J)}{\delta J_j(x)} = 0. \quad (13.17)$$

Expanding equation (13.17) in a power series of the source $J(x)$, one obtains identities between the connected correlation functions that express the physical implications of the symmetry of the action.

For example, for the connected two-point function, one obtains

$$\sum_j \left[t_{ij}^\alpha W_{jk}^{(2)}(x) - W_{ij}^{(2)}(x) t_{jk}^\alpha \right] = 0.$$

However, for renormalization purpose, it is also useful to derive an equation for the generating functional $\Gamma(\varphi)$ of vertex functions. Equation (13.12) directly implies

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \varphi_i(x) \frac{\delta \Gamma}{\delta \varphi_j(x)} = 0, \quad (13.18)$$

which, expanded in powers of φ , yields WT identities for vertex functions. The equation implies that the regularized functional $\Gamma(\varphi)$ is invariant under the transformation (13.14).

Renormalization. We now perform a loop expansion,

$$\Gamma(\varphi) = \sum_{l=0}^{\infty} \Gamma_l(\varphi) g^l, \quad (13.19)$$

where the parameter g is a coupling constant that plays the role of \hbar and orders the loop expansion. Since equation (13.18) is linear in $\Gamma(\varphi)$, and independent of g , all functionals $\Gamma_l(\varphi)$ also satisfy equation (13.18).

The functional $\Gamma_0(\varphi)$ is just the action $\mathcal{S}(\varphi)$ and satisfies, by assumption, the equation. The regularized one-loop functional $\Gamma_1(\varphi)$ satisfies equation (13.18):

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \varphi_i(x) \frac{\delta \Gamma_1(\varphi)}{\delta \varphi_j(x)} = 0. \quad (13.20)$$

We now consider the asymptotic expansion of $\Gamma_1(\varphi)$ in terms of the regularizing parameter (large cut-off expansion, or $1/\varepsilon$ in dimensional regularization, for example). Because equation (13.20) is valid for any value of the regularizing parameter, it is valid for each term in the expansion and thus for the sum of the divergent contributions $\Gamma_1^{\text{div.}}(\varphi)$:

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \varphi_i(x) \frac{\delta \Gamma_1^{\text{div.}}(\varphi)}{\delta \varphi_j(x)} = 0. \quad (13.21)$$

General renormalization theory implies that $\Gamma_1^{\text{div.}}(\varphi)$ is a general local functional of the fields only restricted by power counting; equation (13.21) implies, in addition, that it is symmetric. Adding $-\Gamma_1^{\text{div.}}(\varphi)$ to the action renders the theory one-loop finite. The one-loop renormalized action is still symmetric and, therefore, the new two-loop functional $\Gamma_2(\varphi)$ still satisfies equation (13.18). After one-loop renormalization, $\Gamma_2(\varphi)$ has only local divergences, which also satisfy equation (13.21) and all arguments can be repeated. It is clear that the arguments extend to all orders.

We conclude that the renormalized action \mathcal{S}_r is the most general local functional of the field ϕ_i compatible with power counting, and invariant under the transformation (13.14).

Although this is a sophisticated derivation of a straightforward result, the same strategy, suitably adapted, makes it possible to discuss much more general situations. Therefore, we believe that it has been useful to explain it first in a case in which it can be easily understood.

Finally, we have renormalized using a minimal subtraction scheme. Additional finite renormalizations which are consistent with the symmetry can still be performed.

The problem of the Jacobian: Regularization by point splitting. To solve the problem of the Jacobian, we consider the field integral

$$X(J) = \int [d\phi] \sum_{i,j} \int d^d x d^d y D(x-y) \phi_i(x) t_{ij}^\alpha \frac{\delta}{\delta \phi_j(y)} \exp[-\mathcal{S}(\phi, J)], \quad (13.22)$$

where $D(x)$ is a function localized around $x = 0$. Differentiating explicitly with respect to $\phi_j(y)$, we find

$$\begin{aligned} X(J) &= \int [d\phi] \sum_{i,j} \int d^d x d^d y D(x-y) \phi_i(x) t_{ij}^\alpha \left[-\frac{\delta \mathcal{S}}{\delta \phi_j(y)} + J_j(y) \right] \\ &\quad \times \exp[-\mathcal{S}(\phi, J)]. \end{aligned} \quad (13.23)$$

Alternatively, integrating by parts with $\delta/\delta\phi_j(x)$ acting on the left-hand side, we obtain

$$X(J) = - \int [d\phi] \int d^d x D(0) \text{tr } \mathbf{t}^\alpha \exp[-\mathcal{S}(\phi, J)] = 0.$$

Then, in the limit $D(x) \rightarrow \delta^{(d)}(x)$, equation (13.15) implies that the contribution of the action in expression (13.23) cancels. We recover equation (13.16) since

$$0 = \int [d\phi] \int d^d x \sum_{i,j} J_j(x) t_{ij}^\alpha \phi_i(x) \exp[-\mathcal{S}(\phi, J)] = \int d^d x \sum_{i,j} J_j(x) t_{ij}^\alpha \frac{\delta}{\delta J_i(x)} \mathcal{Z}(\mathbf{J}).$$

13.3 Linear symmetry breaking

For some applications (see, *e.g.* Sections 13.5 and 16.5), it is useful to consider the following situation: the action $\mathcal{S}(\phi)$ is the sum of a symmetric part $\mathcal{S}_{\text{sym.}}(\phi)$, that is, invariant under the transformation (13.8), and a small term breaking the symmetry, linear in the field $\phi(x)$,

$$\mathcal{S}(\phi, \mathbf{c}) = \mathcal{S}_{\text{sym.}}(\phi) - \mathbf{c} \cdot \int \phi(x) d^d x, \quad (13.24)$$

in which \mathbf{c} is a ‘small’ constant vector.

13.3.1 The $O(N)$ symmetric example with linear breaking

We first specialize to an $O(N)$ symmetric classical action for an N -component vector $\phi(x)$, of the form,

$$\mathcal{S}_{\text{sym.}}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + U(\phi^2(x)) \right], \quad (13.25)$$

with

$$U(\rho) = \frac{1}{2} r \rho + \frac{1}{4!} g \rho^2, \quad \text{with } g > 0.$$

The perturbative expansion corresponding to the action (13.24) in the example (13.25) is generated by the steepest descent method. The saddle point, minimum of the classical action, corresponds to constant fields. The minimum yields the expectation value (vacuum expectation value in the particle physics terminology) \mathbf{v} of the field ϕ . It is the solution of

$$\frac{\partial U(\phi^2)}{\partial \phi_i} = 2\phi_i U'(\phi^2) = c_i \Rightarrow \left(r + \frac{g}{6} \mathbf{v}^2 \right) \mathbf{v} = \mathbf{c}. \quad (13.26)$$

The condition that \mathbf{v} is a local minimum implies that the matrix of second derivatives,

$$\mathcal{M}_{ij} = \left. \frac{\partial^2 U}{\partial \phi_i \partial \phi_j} \right|_{\phi=\mathbf{v}} = \left(r + \frac{1}{6} g \mathbf{v}^2 \right) \delta_{ij} + \frac{1}{3} g v_i v_j,$$

must have positive eigenvalues. This condition is consistent with the property that the eigenvalues of \mathcal{M} are the field masses squared in the classical approximation. The matrix has one transverse eigenvalue m_T , ($N-1$) times degenerate, corresponding to eigenvectors orthogonal to \mathbf{v} , and one longitudinal eigenvalue m_L , corresponding to an eigenvector along \mathbf{v} . They are given by

$$m_T^2 = r + \frac{1}{6} g \mathbf{v}^2 = \frac{|\mathbf{c}|}{|\mathbf{v}|}, \quad m_L^2 = r + \frac{1}{2} g \mathbf{v}^2. \quad (13.27)$$

If the action has several minima, in general, one is instructed to choose the absolute minimum of the potential, although this is irrelevant from the point of view of formal perturbation theory.

One then translates the field ϕ , setting

$$\phi(x) = \mathbf{v} + \chi(x). \quad (13.28)$$

The translation cancels the linear term, and the perturbative calculation then proceeds in the standard manner. However, after translation, the mass term is no longer symmetric, as equations (13.27) show, and a non-symmetric $(\mathbf{v} \cdot \chi)\chi^2$ interaction is generated. Correlation functions are no longer symmetric. The perturbative expansion has then to be renormalized, but the form of the UV divergences from the point of view of the symmetry is a priori unknown. It is thus important to understand how much the structure of the renormalized action reflects the structure of the action (13.24).

13.3.2 Renormalized action and symmetry to all orders

The answer follows from a simple argument. With obvious notation,

$$\mathcal{Z}(\mathbf{J}) = \mathcal{Z}_{\text{sym.}}(\mathbf{J} + \mathbf{c}) \Rightarrow \mathcal{W}(\mathbf{J}) = \mathcal{W}_{\text{sym.}}(\mathbf{J} + \mathbf{c}), \quad (13.29)$$

where $\mathcal{Z}_{\text{sym.}}(\mathbf{J})$ is the generating functional of correlation functions corresponding to the symmetric action $\mathbf{c} = 0$. Equation (13.17) then implies

$$\int d^d x \sum_{i,j} t_{ij}^\alpha [J_i(x) + c_i] \frac{\delta \mathcal{W}(\mathbf{J})}{\delta J_j(x)} = 0. \quad (13.30)$$

Expanding in powers of $J_i(x)$, one obtains a set of relations (WT identities) between connected correlation functions, which can be most conveniently expressed, in the Fourier representation, as

$$\sum_{i,j} c_i t_{ij}^\alpha \widetilde{W}_{jk_1 \dots k_n}^{(n+1)}(0, p_1, \dots, p_n) + \sum_{r=1}^n \sum_j t_{k_r j}^\alpha \widetilde{W}_{k_1 \dots k_{r-1} j k_{r+1} \dots k_n}^{(n)}(p_1, \dots, p_n) = 0. \quad (13.31)$$

The generating functional Γ of vertex functions is given by the Legendre transformation,

$$\begin{aligned} \Gamma(\boldsymbol{\varphi}) + \mathcal{W}(\mathbf{J}) &= \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\varphi}(x), \\ \varphi_i(x) &= \frac{\delta \mathcal{W}}{\delta J_i(x)} = \frac{\delta \mathcal{W}_{\text{sym.}}(\mathbf{J} + \mathbf{c})}{\delta J_i(x)}. \end{aligned} \quad (13.32)$$

By contrast, in the symmetric situation these relations read,

$$\begin{aligned} \Gamma_{\text{sym.}}(\boldsymbol{\xi}) + \mathcal{W}_{\text{sym.}}(\mathbf{J}) &= \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\xi}(x), \\ \xi_i(x) &= \frac{\delta \mathcal{W}_{\text{sym.}}(\mathbf{J})}{\delta J_i(x)}. \end{aligned} \quad (13.33)$$

Replacing $J_i(x)$ by $J_i(x) + c_i$ in the relations (13.33), one obtains

$$\begin{aligned} \Gamma_{\text{sym.}}(\boldsymbol{\varphi}) + \mathcal{W}_{\text{sym.}}(\mathbf{J} + \mathbf{c}) &= \int d^d x (\mathbf{J}(x) + \mathbf{c}) \cdot \boldsymbol{\varphi}(x), \\ \varphi_i(x) &= \frac{\delta \mathcal{W}_{\text{sym.}}(\mathbf{J} + \mathbf{c})}{\delta J_i(x)} \end{aligned} \quad (13.34)$$

and, therefore, comparing equation (13.32) with equation (13.34),

$$\Gamma(\boldsymbol{\varphi}) = \Gamma_{\text{sym.}}(\boldsymbol{\varphi}) - \mathbf{c} \cdot \int d^d x \boldsymbol{\varphi}(x). \quad (13.35)$$

This identity proves that the divergences of the functionals $\Gamma(\boldsymbol{\varphi})$ and $\Gamma_{\text{sym.}}(\boldsymbol{\varphi})$ are identical. Therefore, after replacement of the regularized symmetric action by the renormalized symmetric action, the theory is finite for any value of \mathbf{c} . This is informally expressed by stating that the *linear breaking term is not renormalized*.

To calculate the vertex functions of ϕ , one must then translate φ by the ϕ field expectation value, setting (see Section 7.7)

$$\varphi(x) = \mathbf{v} + \chi(x), \quad (13.36)$$

with

$$\frac{\delta\Gamma}{\delta\varphi_i(x)} \Big|_{\varphi_i(x)=v_i} = 0 \Rightarrow \frac{\delta\Gamma_{\text{sym.}}}{\delta\varphi_i(x)} \Big|_{\varphi_i(x)=v_i} = c_i, \quad (13.37)$$

where $\delta^2\Gamma(\mathbf{v})/\delta\phi_i\delta\phi_j$ is a positive matrix.

The vertex functions are the coefficients of the expansion of $\Gamma(\varphi)$ in powers of χ . In the tree approximation, \mathbf{v} is determined by the location of the minimum of the action (13.24).

The WT identities for $\Gamma(\varphi)$ can be inferred from the identity (13.18) for $\Gamma_{\text{sym.}}$:

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \left[\frac{\delta\Gamma}{\delta\varphi_i(x)} + c_i \right] \varphi_j(x) = 0, \quad (13.38)$$

which, after the translation (13.36), becomes

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \left[\frac{\delta\Gamma}{\delta\chi_i} (\chi + \mathbf{v}) + c_i \right] (\chi_j + v_j) = 0. \quad (13.39)$$

WT identities for vertex functions. The identity (13.39) implies non-trivial relations between vertex functions. Setting $\chi = 0$, one obtains

$$\sum_{i,j} t_{ij}^\alpha c_i v_j = 0, \quad (13.40)$$

which implies that the breaking vector \mathbf{c} and the expectation value \mathbf{v} are left invariant by the same subgroup of G . In the example of the $O(N)$ symmetry, equation (13.40) implies that the vector \mathbf{v} is proportional to the vector \mathbf{c} .

Differentiating once with respect to $\chi_k(y)$ and then setting $\chi = 0$, one finds the relation between the one- and two-point functions,

$$\int d^d x \left[\sum_{i,j} v_j t_{ij}^\alpha \Gamma_{ik}^{(2)}(x, y) + \sum_i t_{ik}^\alpha c_i \delta^{(d)}(x - y) \right] = 0, \quad (13.41)$$

with

$$\Gamma_{ij}^{(2)}(x, y) = \frac{\delta^2\Gamma(\chi + v)}{\delta\chi_i(x)\delta\chi_j(y)} \Big|_{\chi=0}.$$

In terms of the Fourier transform $\tilde{\Gamma}_{ij}^{(2)}(p)$ of the two-point function,

$$\Gamma_{ij}^{(2)}(x, y) = \int \frac{d^d p}{(2\pi)^d} e^{-ip(x-y)} \tilde{\Gamma}_{ij}^{(2)}(p), \quad (13.42)$$

equation (13.41) becomes

$$\sum_{i,j} v_j t_{ji}^\alpha \tilde{\Gamma}_{ik}^{(2)}(0) + \sum_i t_{ki}^\alpha c_i = 0. \quad (13.43)$$

The equation determines the geometric structure of the inverse propagator, in the presence of the linear symmetry breaking term, at zero momentum.

In the example (13.25), the identity (13.43) yields the value of the propagator of the components of the field orthogonal to the vector \mathbf{c} at zero momentum,

$$\tilde{\Gamma}_T^{(2)}(0) = c/v .$$

Equation (13.43) is the last equation that involves \mathbf{c} explicitly. The terms of higher degree in χ involve only the field expectation value \mathbf{v} . By identifying the coefficient of degree $(n+1)$ in χ , one obtains a relation between the Fourier transform of the $(n+1)$ -point function $\tilde{\Gamma}^{(n+1)}$ with one momentum set to 0 and the n -point function $\tilde{\Gamma}^{(n)}$:

$$\sum_{j,k} v_j t_{jk}^\alpha \tilde{\Gamma}_{ki_1 \dots i_n}^{(n+1)}(0, p_1, \dots, p_n) + \sum_{r=1}^n \sum_k t_{i_r k}^\alpha \tilde{\Gamma}_{i_1 \dots i_{r-1} k i_{r+1} \dots i_n}^{(n)}(p_1, \dots, p_n) = 0 . \quad (13.44)$$

For example, for $n = 2$ the equation becomes

$$\sum_{i,j} v_j t_{ji}^\alpha \tilde{\Gamma}_{ikl}^{(3)}(0, p, -p) + \sum_i t_{ii}^\alpha \tilde{\Gamma}_{ik}^{(2)}(p) + \sum_i t_{ki}^\alpha \tilde{\Gamma}_{il}^{(2)}(p) = 0 .$$

In a renormalization scheme based on imposing renormalization conditions to the primitively divergent vertex functions, the set of WT identities implies relations between the different parameters. Apart from the vector \mathbf{v} , the non-symmetric QFT depends on the same number of independent parameters as the symmetric QFT. In the example of the $O(N)$ symmetric $(\phi^2)^2$ QFT, in four dimensions, it is possible to impose one arbitrary renormalization condition on $\tilde{\Gamma}_{1111}^{(4)}(p_i)$, and two conditions on $\tilde{\Gamma}_{11}^{(2)}(p)$. All other conditions are determined by the WT identities (13.44) used for $n = 1$ to 4.

Non-renormalizable interactions. Note that WT identities only depend on the form (13.24) of the action. They remain valid even if non-renormalizable symmetric interactions are added (with a suitable symmetric regularization). Actually, non-renormalizable terms may be invariant under only a subgroup G' of G , provided the quadratic and quartic invariants of G' are automatically G -invariant (case of emergent symmetries).

13.4 Spontaneous symmetry breaking

Spontaneous symmetry breaking (SSB) is a possible limit of linear symmetry breaking when the breaking parameter goes to zero. In this limit, the action becomes symmetric, but depending on the values of other parameters (in our examples the coefficient of the ϕ^2 term in the action), physical observables may be explicitly symmetric or not. Many physical models in particle physics are based on the concept of SSB. Indeed, SSB makes it possible to construct models with broken symmetries that depend on no more parameters than the symmetric models. The appearance of (Nambu–Goldstone) massless particles [102] (Goldstone modes in statistical physics) is, in general, the most characteristic feature of such models (except in gauge theories).

One important issue from the point of view of perturbation theory is that SSB, in the classical limit, is associated with degenerate classical minima. Each minimum is the starting point of a perturbative expansion. A question then arises, should one sum over the contributions coming from all minima or consider only one of them? For group invariant correlation functions the question is irrelevant, since all minima give the same contribution: summing over all minima just yields a factor that disappears in the normalization of the field integral.

However, for non-invariant correlation functions a summation over all minima projects onto group-invariant functions: as a consequence, all non-vanishing correlation functions are invariant and the expectation value of the field vanishes.

In the case of degenerate classical minima, the correct procedure depends on the true physical situation, beyond perturbation theory. In the absence of phase transitions, one must sum over the contributions of all minima: quantum (or statistical) fluctuations restore the symmetry broken in the classical approximation, and the true ground state is unique. By contrast, when a phase transition occurs, there is a *breaking of ergodicity* in the ordered phase, and one must choose one specific minimum. The quantum ground state is degenerate.

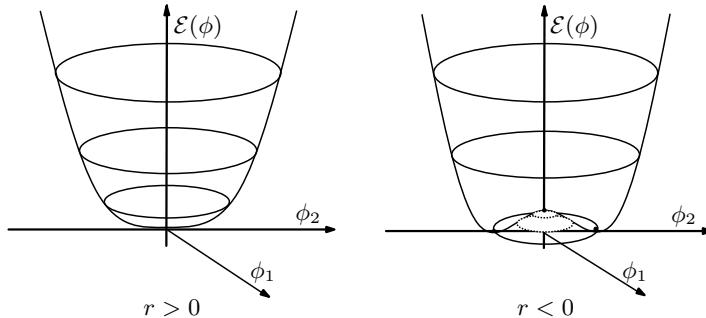


Fig. 13.1 The density of the classical ϕ action: $O(2)$ symmetry

13.4.1 Classical analysis: The $O(N)$ example

We consider the $O(N)$ symmetric example (13.24) in the classical approximation. Equation (13.26) for the field expectation value shows that as long as \mathbf{c} does not vanish, it is possible to pass continuously from a situation in which the parameter r is positive to a situation in which r is negative, without encountering any singularity. In particular, the expectation value \mathbf{v} is, at \mathbf{c} fixed, a regular function of r at $r = 0$. By contrast, if \mathbf{c} vanishes, the expectation value \mathbf{v} vanishes identically for $r > 0$, and takes a non-vanishing value for $r < 0$ such that

$$|\mathbf{v}| = \sqrt{-r/g}, \quad (13.45)$$

as one can easily understand by displaying the action density $\mathcal{E}(\phi) = U(\phi^2)$ in both situations (see Fig. 13.1). In the latter situation, the classical minimum of the action density is degenerate. Starting from a given minimum, it is possible to describe all other minima by acting on the vector \mathbf{v} with the symmetry group. In the $O(N)$ example, the surface of minima is a sphere with a radius given by equation (13.45).

Assuming a situation of SSB, we construct a perturbation theory around one minimum \mathbf{v} which is, at leading order, the field expectation value. We thus shift the field,

$$\phi(x) = \mathbf{v} + \chi(x).$$

The eigenvalues of the χ -field mass matrix are given by the limit of equations (13.27). Using equation (13.45), one finds

$$m_T^2 = 0, \quad m_L^2 = \frac{1}{3}g\mathbf{v}^2. \quad (13.46)$$

The mass matrix has $(N - 1)$ eigenvalues 0, corresponding to eigenvectors orthogonal to \mathbf{v} , and to the flatness of the potential along a group orbit. The physical consequence is that the spontaneous breaking of a continuous symmetry generates *Goldstone modes*: in the framework of particle physics, they correspond to $(N - 1)$ massless scalar particles called *Goldstone (or Nambu–Goldstone) bosons*.

13.4.2 General continuous symmetry group

We now examine a general situation. We assume that a G -symmetric action has degenerate minima. We call \mathbf{v} the minimum chosen to expand perturbation theory, and thus the field expectation value at leading order. We introduce the subgroup H of G , little group (stabilizer) of the vector \mathbf{v} , that is, the subgroup of G that leaves the vector \mathbf{v} invariant. By definition, the p generators \mathbf{t}^α of the Lie algebra $\mathcal{L}(H)$ of H satisfy

$$\mathcal{L}(H) : \quad 1 \leq \alpha \leq p \Rightarrow \sum_j t_{ij}^\alpha v_j = 0.$$

We denote by $\mathcal{L}(G/H)$ the vector space (it is not an algebra!) generated by the complementary set in the Lie algebra $\mathcal{L}(G)$ of G . It is characterized by

$$\mathcal{L}(G/H) : \quad \sum_{\alpha > p} t_{ij}^\alpha v_j \omega_\alpha = 0 \Rightarrow \omega_\alpha = 0 \text{ for all } \alpha.$$

For $\alpha > p$, the vectors $(v^\alpha)_i = \sum_j t_{ij}^\alpha v_j$ are thus linearly independent. We then parametrize the field ϕ in the form of a group element acting on a vector:

$$\phi(x) = \exp\left(\sum_{\alpha > p} \mathbf{t}^\alpha \xi^\alpha(x)\right)(\mathbf{v} + \boldsymbol{\rho}(x)) = \mathbf{v} + \sum_{\alpha > p} \xi^\alpha(x) \mathbf{t}^\alpha \mathbf{v} + \boldsymbol{\rho}(x) + \dots,$$

in which $\boldsymbol{\rho}(x)$ has components only in the subspace orthogonal to all vectors $\mathbf{t}^\alpha \mathbf{v}$. In the $O(N)$ example, $\boldsymbol{\rho}$ has only one component along \mathbf{v} . The parametrization is such that the mapping of fields $\{\boldsymbol{\rho}(x), \xi^\alpha(x)\} \mapsto \phi(x) - \mathbf{v}$ can be inverted for small fields. This property ensures that if the fluctuations of the field ϕ around its expectation value are in some sense small, perturbation theory is at least qualitatively meaningful.

Inserting the parametrization into the action, we note the following: the contributions to the action that are derivative free depend only on $\boldsymbol{\rho}(x)$ because they are G -invariant. The dependence in the fields $\xi^\alpha(x)$ is entirely contained in the terms with derivatives, therefore, these fields are massless. We conclude that, in the spontaneous breaking of the symmetry group G to a subgroup H (the group that leaves the field expectation value invariant), a number of massless Goldstone modes (bosons) equal to the number of generators of G that do not belong to H are generated. This result is valid in the classical approximation. We now generalize it to the full QFT.

13.4.3 WT identities and SSB

To connect continuously the two phases, symmetric and with SSB, without encountering a singularity, we start from the symmetric (and disordered) phase, $\mathbf{c} = 0$; we give a non-vanishing value to \mathbf{c} , perform a continuation in the coefficient of ϕ^2 from values corresponding to the symmetric phase to those corresponding to the broken phase, and again take the vanishing \mathbf{c} limit. We then assume the existence of non-trivial solutions to the equation

$$\frac{\delta \Gamma}{\delta \varphi_i(x)} \Big|_{\varphi(x)=\mathbf{v}} = 0. \tag{13.47}$$

In Section 7.11, we have explained how the existence of solutions to this equation is consistent with the convexity of the function $\Gamma(\mathbf{v})$ (equation (7.73)).

Since the WT identities (13.39) hold for any value of the parameters, and we have proceeded by analytic continuation, we still have in the $\mathbf{c} = 0$ limit,

$$\int d^d x \sum_{i,j} t_{ij}^\alpha \frac{\delta \Gamma(\chi + v)}{\delta \chi_i(x)} (\chi_j + v_j) = 0, \quad (13.48)$$

the direction of v_i being fixed by equation (13.40).

Goldstone modes. One important consequence of WT identities is obtained by taking the $\mathbf{c} = 0$ limit in equation (13.43):

$$\sum_{i,j} v_j t_{ji}^\alpha \tilde{\Gamma}_{ik}^{(2)}(0) = 0. \quad (13.49)$$

The equation has the following interpretation: as in the classical analysis, we introduce the subgroup H of G , little group (stabilizer) of the vector \mathbf{v} . Since for $\alpha > p$, the vectors $(v^\alpha)_i = \sum_j t_{ij}^\alpha v_j$ are linearly independent, equation (13.49) implies that the real symmetric matrix $\tilde{\Gamma}_{ij}(0)$ has as many eigenvectors with eigenvalue 0 as there are generators in $\mathcal{L}(G/H)$, confirming the classical analysis. The corresponding components of the field are Goldstone modes associated with the spontaneous breaking of the G -symmetry, associated with massless particles.

The physics of SSB is again extensively discussed, in the context of macroscopic phase transitions, in Chapters 14–19 devoted to critical phenomena.

13.5 Chiral symmetry breaking in strong interactions: Effective theory

One striking feature of strong interactions in low energy particle (hadron) physics is the observation of approximate spontaneously broken $SU(N) \times SU(N)$ chiral symmetries, especially for N small, which manifests itself, in particular, in the small masses of the pseudoscalar mesons. The π -meson is especially light, an indication that the explicit breaking of the $SU(2) \times SU(2)$ symmetry is quite small. With our present understanding, this property is a direct consequence of the small masses of the **u** and **d** quarks (see *e.g.* Sections 12.3.3) and the vector-like coupling of quarks to gluons (see Section 23.4.1). The mass of the **s** quark (Section 23.3), and thus, the explicit breaking of the $SU(3) \times SU(3)$ symmetry is larger, as can be inferred from the masses of the K and η pseudoscalar mesons.

Note that, according to the remark at the end of Section 13.8, since a fermion mass operator in a renormalizable QFT in four dimensions has dimension 3, the concept of a symmetry broken by fermion mass terms is indeed meaningful. However, the search for analytic methods to derive low energy properties of hadrons from a fundamental theory of quarks and gluons has up to now proved elusive. Most direct results are thus obtained from computer intensive studies of discretized lattice versions (see Chapter 25). Progress has been slow for some time but, finally, a number of precise results have been obtained. The most serious numerical difficulties are related to the dynamics of quarks, in particular, from the point of view of chiral properties.

Here, by contrast, we explain how one can construct renormalizable *effective low energy theories* based on observed hadrons like protons, neutrons, π -mesons, and so on. In such theories the chiral symmetry is explicitly broken by linear terms in some scalar fields, which have the transformation properties of fermion mass terms, and which together with the pseudoscalars transform under representations of the chiral group. Therefore, we face the situation we have discussed at some length in Section 13.3.

13.5.1 The chiral symmetry: General structure

The action for N free massless Dirac fermions in even dimensions d reads,

$$\mathcal{S}(\psi, \bar{\psi}) = - \int d^d x \sum_i \bar{\psi}_i(x) \not{\partial} \psi_i(x). \quad (13.50)$$

It has a chiral $U(N) \times U(N)$ symmetry corresponding to the transformations,

$$\psi'(x) = [\frac{1}{2}(\mathbf{1} + \gamma_S)\mathbf{U}_+ + \frac{1}{2}(\mathbf{1} - \gamma_S)\mathbf{U}_-] \psi(x), \quad (13.51)$$

$$\bar{\psi}'(x) = \bar{\psi}(x) \left[\frac{1}{2}(\mathbf{1} + \gamma_S)\mathbf{U}_-^\dagger + \frac{1}{2}(\mathbf{1} - \gamma_S)\mathbf{U}_+^\dagger \right], \quad (13.52)$$

where \mathbf{U}_\pm are two $N \times N$ unitary matrices belonging to the two $U(N)$ groups.

One couples the fermions to scalar bosons forming a complex $N \times N$ matrix $\mathbf{M}(x)$. One verifies that the interaction term,

$$-g \int d^d x \sum_{i,j} \bar{\psi}_i(x) \left[\frac{1}{2}(\mathbf{1} + \gamma_S)M_{ij}(x) + \frac{1}{2}(\mathbf{1} - \gamma_S)M_{ij}^\dagger(x) \right] \psi_j(x), \quad (13.53)$$

is invariant under the transformations (13.51) and (13.52) provided the matrix \mathbf{M} transforms as

$$\mathbf{M}'(x) = \mathbf{U}_- \mathbf{M}(x) \mathbf{U}_+^\dagger. \quad (13.54)$$

The total action also satisfies reflection Hermiticity as defined in Section 12.3.1. It can be made invariant under a space reflection P (see Section 12.3.2) if \mathbf{M} transforms as

$$\mathbf{M}_P(x) = \mathbf{M}^\dagger(\tilde{x}), \quad (13.55)$$

in which \tilde{x} is obtained from x by changing the sign of one component. Therefore, the matrix $\Sigma = (\mathbf{M} + \mathbf{M}^\dagger)/\sqrt{2}$ represents a set of scalar fields, and the matrix $\Pi = (\mathbf{M} - \mathbf{M}^\dagger)/\sqrt{2}$ a set of pseudoscalar fields. Finally, under a charge conjugation C , \mathbf{M} transforms like

$$\mathbf{M}_C = \mathbf{M}^*, \quad \text{if } d/2 \text{ is odd, and} \quad \mathbf{M}_C = {}^T \mathbf{M}, \quad \text{if } d/2 \text{ is even.} \quad (13.56)$$

A possible action for the boson fields symmetric under $U(N) \times U(N)$ transformations is then

$$\mathcal{S}(\mathbf{M}) = \int d^d x \operatorname{tr} [\nabla \mathbf{M}(x) \nabla \mathbf{M}^\dagger(x) + V(\mathbf{M}(x) \mathbf{M}^\dagger(x))], \quad (13.57)$$

where $V(\varphi)$ is a polynomial of φ . If in addition one adds a term proportional to $\det \mathbf{M} + \det \mathbf{M}^\dagger$, one reduces the symmetry to $SU(N) \times SU(N) \times U(1)$ (the factor $U(1)$ corresponds to the baryonic charge).

Finally, the most general symmetry breaking term linear in the boson fields, consistent with the discrete symmetries (13.55) and (13.56), is

$$\mathcal{S}_B(\mathbf{M}) = -\frac{1}{\sqrt{2}} \int d^d x \operatorname{tr} \mathbf{C} [\mathbf{M}(x) + \mathbf{M}^\dagger(x)], \quad (13.58)$$

in which \mathbf{C} is a Hermitian matrix:

$$\mathbf{C} = \mathbf{C}^\dagger.$$

To the transformations (13.51), (13.52), and (13.54) correspond two currents (for details, see Section A13.1). It is convenient to consider the vector current $\mathbf{V}_\mu(x)$, which is associated with the Lie algebra of the diagonal subgroup $U(N)$ of $U(N) \times U(N)$ ($\mathbf{U}_+ = \mathbf{U}_-$) that conserves parity:

$$-iV_\mu^\alpha(x) = -\bar{\psi}(x)\mathbf{t}^\alpha\gamma_\mu\psi(x) + \text{tr } \mathbf{t}^\alpha \left\{ [\partial_\mu \mathbf{M}^\dagger(x), \mathbf{M}(x)] + [\partial_\mu \mathbf{M}(x), \mathbf{M}^\dagger(x)] \right\}, \quad (13.59)$$

and the axial current $\mathbf{A}_\mu(x)$ associated with the complementary set of generators in the Lie algebra, that is, $\mathcal{L}(U(N) \times U(N)/U(N))$:

$$-iA_\mu^\alpha(x) = -\bar{\psi}(x)\mathbf{t}^\alpha\gamma_S\gamma_\mu\psi(x) + \text{tr } \mathbf{t}^\alpha \left\{ [\partial_\mu \mathbf{M}^\dagger(x), \mathbf{M}(x)]_+ + [\partial_\mu \mathbf{M}(x), \mathbf{M}^\dagger(x)]_+ \right\}, \quad (13.60)$$

where the $+$ index means that the expression between brackets is an anticommutator.

If the matrix \mathbf{C} is proportional to the identity, the chiral symmetry is broken, but the diagonal symmetry remains, and the vector current is conserved. The axial current is conserved only if \mathbf{C} vanishes:

$$\sum_\mu \partial_\mu V_\mu^\alpha(x) = -i \text{tr} \{ [\mathbf{t}^\alpha, \mathbf{C}] \Sigma(x) \}, \quad \sum_\mu \partial_\mu A_\mu^\alpha(x) = \text{tr} \{ [\mathbf{t}^\alpha, \mathbf{C}]_+ \Pi(x) \}. \quad (13.61)$$

13.6 The linear σ -model

The case $N = 2$, $d = 4$, is of particular physical interest, because the pion mass is remarkably small (the pion is almost a Goldstone boson) and, thus, the explicit breaking of chiral symmetry is small [103].

The even much smaller diagonal $SU(2)$ isospin breaking is here neglected.

13.6.1 The scalar boson sector

The preceding analysis leads to a theory with eight real scalar fields. However, the group $SU(2)$ (but not the group $U(2)$) has the property that a representation and its complex conjugate are equivalent:

$$\mathbf{U} = \tau_2 \mathbf{U}^* \tau_2, \quad \forall \mathbf{U} \in SU(2),$$

in which τ_2 is the usual Pauli matrix (we denote in this section the Pauli matrices by τ_i rather than σ_i , as in Section A12.1.4, to eliminate possible confusion with the traditional notation for fields). Therefore, \mathbf{M} and $\tau_2 \mathbf{M}^* \tau_2$ have the same transformation properties. The representation can be reduced, and the matrix \mathbf{M} can be parametrized in terms of two fields $\sigma(x)$ and $\boldsymbol{\pi}(x)$ in the form

$$\mathbf{M} = \tau_2 \mathbf{M}^* \tau_2 \equiv \frac{1}{\sqrt{2}} (\sigma + i\boldsymbol{\tau} \cdot \boldsymbol{\pi}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \sigma + i\pi_0 & \pi_2 + i\pi_1 \\ -\pi_2 + i\pi_1 & \sigma - i\pi_0 \end{bmatrix}. \quad (13.62)$$

The group $SU(2) \times SU(2)$ is the covering group of $O(4)$ which is also the symmetry group of the boson part of the action. A breaking of the $O(4)$ symmetry by a term linear in the boson fields singles out one direction in the four-dimensional space and, therefore, reduces the $O(4)$ symmetry to a residual $O(3)$ symmetry. We assume without loss of generality that the linear breaking term is proportional to $\sigma(x)$.

It follows from equation (13.55) that $\sigma(x)$ is a scalar field and $\boldsymbol{\pi}(x)$ a pseudoscalar field (associated to the pi-meson or pion). In $d = 4$ dimensions, σ and π_0 correspond to neutral mesons, while the combinations

$$\pi_{\pm} = (\pi_1 \pm i\pi_2)/\sqrt{2},$$

correspond to charged mesons, as charge conjugation shows. The classical scalar action reduces to

$$\mathcal{S}(\sigma, \boldsymbol{\pi}) = \int d^4x \left[\frac{1}{2} (\nabla\sigma(x))^2 + \frac{1}{2} (\nabla\boldsymbol{\pi}(x))^2 + V(\sigma^2(x) + \boldsymbol{\pi}^2(x)) - c\sigma(x) \right], \quad (13.63)$$

with

$$V(\rho) = \frac{1}{2}r\rho + \frac{1}{4!}\lambda\rho^2, \quad c > 0, \lambda > 0. \quad (13.64)$$

The infinitesimal transformations that correspond to the generators of $SO(4)$ that do not belong to the conserved $SO(3)$ parametrized by a vector ω , read

$$\delta_{\omega}\boldsymbol{\pi}(x) = -\omega\sigma(x), \quad \delta_{\omega}\sigma(x) = \omega \cdot \boldsymbol{\pi}(x). \quad (13.65)$$

Currents. The action (13.71) has an exact $SU(2) \times U(1)$ symmetry. To the generators of the exact $SU(2)$ or $SO(3)$ group corresponds the vector current

$$V_{\mu}^{ij}(x) = \pi^i(x)\partial_{\mu}\pi^j(x) - \pi^j(x)\partial_{\mu}\pi^i(x).$$

The current satisfies the conservation equation,

$$\sum_{\mu} \partial_{\mu} V_{\mu}^{ij}(x) = 0.$$

The axial current, which is associated with generators of $SO(4)$ that do not belong to the generators of $SO(3)$, is

$$A_{\mu}^i(x) = \sigma(x)\partial_{\mu}\pi^i(x) - \pi^i(x)\partial_{\mu}\sigma(x).$$

The divergence of the axial current, in the standard normalization, which differs by a factor 2 from the definition (13.60) (see equations (13.65) and (13.72)), is given by

$$\sum_{\mu} \partial_{\mu} \mathbf{A}_{\mu}(x) = c\boldsymbol{\pi}(x). \quad (13.66)$$

The actions (13.63) and (13.71) (see below) implement the idea of *partially conserved axial current* (PCAC) for $SU(2)$.

The pattern of symmetry breaking in the classical approximation. In the absence of fermions, the model reduces to the $(\phi^2)^2$ field theory with $O(4)$ symmetry, a special case of the $O(N)$ model discussed in Section 13.3.1. Equation (13.26) gives the relation between the expectation value v of the field σ and the symmetry breaking parameter c in the classical approximation,

$$v(r + \lambda v^2/6) = c, \quad (13.67)$$

and equation (13.27) the masses of the $\boldsymbol{\pi}$ and σ particles at the tree order,

$$m_{\pi}^2 = r + \lambda v^2/6, \quad m_{\sigma}^2 = r + \lambda v^2/2. \quad (13.68)$$

The expectation value v is experimentally accessible from the weak π -meson decay as a consequence of relation (13.66), and is denoted traditionally by f_π .

The assumption that accounts for the success of PCAC phenomenology is that the explicit symmetry breaking term is small, and that the physics is close to a situation of SSB. For the model (13.71), this means, in particular, that m_π is small compared to m_σ . With this assumption, it is possible to predict some general features of the low energy $\pi\pi$ scattering. After using the relations (13.68) to eliminate m and λ , and in terms of the standard invariant kinematic variables,

$$s = -(p_1 + p_2)^2, \quad t = -(p_1 + p_3)^2, \quad u = -(p_1 + p_4)^2, \quad (13.69)$$

the connected amputated π -field four-point function at this order can be expressed as

$$\begin{aligned} [W_{ijkl}^{(4)}]_{\text{amp.}}(s, t, u) &= \frac{s - m_\pi^2}{m_\sigma^2 - s} \frac{m_\sigma^2 - m_\pi^2}{f_\pi^2} \delta_{ij} \delta_{kl} + \frac{t - m_\pi^2}{m_\sigma^2 - t} \frac{m_\sigma^2 - m_\pi^2}{f_\pi^2} \delta_{ik} \delta_{jl} \\ &\quad + \frac{u - m_\pi^2}{m_\sigma^2 - u} \frac{m_\sigma^2 - m_\pi^2}{f_\pi^2} \delta_{il} \delta_{jk}. \end{aligned} \quad (13.70)$$

The physical scattering amplitude is then obtained by setting all momenta on the mass shell: $p_i^2 = -m_\pi^2$ and then $s + t + u = 4m_\pi^2$.

Since m_σ is supposed to be large compared to m_π , the expression (13.70) makes quantitative predictions for s, t, u of order m_π^2 , that is, at low energy. Values corresponding to infinite σ -mass are often quoted (see also Section 19.10.2). Although the $\pi\pi$ scattering amplitude, of course, cannot be measured directly, indirect methods provide an experimental confirmation of the resulting pattern.

13.6.2 The boson-fermion model

With the addition of fermions, the action can be written as

$$\mathcal{S}(\bar{\psi}, \psi, \sigma, \boldsymbol{\pi}) = \int d^d x \left\{ -\bar{\psi}(x) [\partial + g(\sigma(x) + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi}(x))] \psi(x) + \mathcal{S}(\sigma, \boldsymbol{\pi}) \right\}, \quad (13.71)$$

where the fermion doublet $\psi(x)$ corresponds to the two nucleons, proton and neutron.

The infinitesimal transformations of the fermion fields, corresponding to the equations (13.65) are

$$\delta_\omega \psi(x) = \frac{1}{2} i \gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\omega} \psi(x), \quad \delta_\omega \bar{\psi}(x) = \frac{1}{2} i \bar{\psi}(x) \gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\omega}. \quad (13.72)$$

Leading order. In the unbroken phase, the mass of the fermion vanishes for $c = 0$. The fermion mass m_N of nucleons is thus generated by the Yukawa coupling and the σ expectation value

$$m_N \equiv m_\psi = gv. \quad (13.73)$$

The Yukawa coupling constant g is arbitrary in the model and must be extracted from some experimental information: at this order, the parameter g can be identified with the coupling constant $g_{\pi NN}$ which governs the long range part of the N - N (nucleon–nucleon) potential due to pion exchange. This leads to the relation between physical quantities,

$$g_{\pi NN} = m_N/f_\pi, \quad (13.74)$$

the tree approximation of the so-called Goldberger–Treiman relation. It agrees semi-quantitatively with experiment since

$$g_{\pi NN} = 13.6, \quad \frac{m_N}{f_\pi} \simeq \frac{939}{93.3} = 10.. \quad (13.75)$$

Then all parameters but m_σ are fixed. The low energy π - N scattering amplitude, for example, can be calculated. A definite prediction can be made only for m_σ infinite, that is, in a low energy limit; it agrees well with experimental data.

13.6.3 Beyond the tree approximation

Since the action (13.71) is renormalizable, it is possible to calculate loop corrections without introducing new parameters [104–107]. Then, several difficulties are encountered.

(i) Loop corrections become large at moderate energies. For example, in the π - π scattering, one encounters the ρ resonance. It becomes then necessary to sum the perturbation expansion (see Section 41.2). In this specific example, the series has been summed using the method of Padé approximants [109, 108]. (Similar applications of Padé approximants to perturbative hadron physics are reviewed in Ref. [110].) In this way, loop corrections are useful to improve the tree level amplitudes from the point of view of unitarity at low energy.

(ii) Since the σ mass is larger than $2m_\pi$, the σ particle is unstable (it appears in $\pi - \pi$ scattering as is a resonance), because it can decay into two pions. In the exact π - π scattering amplitude, the resonance leads to singularities in the second sheet of the unitarity cut in the complex s -plane. However, at any finite order in perturbation theory, the singularities associated with the σ -particle are on the real axis, since the width of the particle is a non-perturbative effect. Fits of experimental data seem to impose a rather small σ mass. Therefore, loop corrections are affected by non-physical singularities even at rather low energy. This problem of the perturbative treatment of fields corresponding to unstable particles remains to a large extent unsolved. One possible idea is to make a systematic large m_σ expansion (see also Section 19.10.2), but the validity of the expansion is then limited to energies smaller than $4m_\sigma^2$, that is, rather low energies.

(iii) Finally, perturbative corrections to the nucleon mass are large, and this also adversely affects the position of singularities in scattering amplitudes involving fermions.

Therefore, although much effort has gone into the study of the model (13.71), only limited results have been obtained. Some rely on WT identities, as we will explain. Since the predictions depend only geometric considerations, models have been developed based on effective field theory considerations.

13.6.4 Renormalization group: β -functions and triviality issue

Renormalization group (RG) functions of a similar model are calculated in Section 20.1. We consider here the bare RG, formulated in terms of a cut-off Λ . Changing a few geometric factors, one obtains the RG β -functions in four dimensions of the model (13.71):

$$\beta_{g^2} = \frac{2}{\pi^2} g^4 + O(\text{two loops}), \quad \beta_\lambda = \frac{1}{8\pi^2} (2\lambda^2 + 8\lambda g^2 - 48g^4) + O(\text{two loops}).$$

The most likely implication of the solution of the RG equations is the following. When the ratio between the cut-off scale Λ and the physical (or renormalization) scale μ is large, the effective g^2 and λ coupling constants at scale μ decrease like

$$g^2(\mu/\Lambda) \sim \frac{\pi^2}{2 \ln(\Lambda/\mu)}, \quad \lambda(\mu/\Lambda) \sim \frac{\pi^2(1 + \sqrt{7})}{\ln(\Lambda/\mu)}$$

and, eventually, vanish at infinite cut-off. The renormalized QFT (13.71) with fermions, like the ϕ^4 QFT as we have shown in Section 9.12 is affected, in four dimensions, by the *triviality issue* (see Chapters 15, 18, and 24). Although the QFT is renormalizable in perturbation theory, it is impossible to send the cut-off Λ to infinity: the model makes sense at a mass scale μ only for renormalized couplings which are bounded by $\text{const.}/\ln(\Lambda/\mu)$. Therefore, the addition of loop corrections is meaningful only if the momenta and the coupling constants are small enough (in a correlated way as stated above). A Landau ‘ghost’ will typically be a manifestation of this problem.

Setting $g = g_{\pi NN} = 13.6$, one finds a ratio Λ/μ close to 1, while the quoted values of λ [106] lead to a ratio smaller than 2. The asymptotic form is thus inadequate but suggests that the energy scale at which the model becomes inconsistent is not much higher than the physical scale.

13.7 WT identities

We have pointed out some difficulties one encounters when one tries to derive consequences from the renormalizable linear σ -model using perturbation theory. However, some properties, WT identities direct consequences of the broken symmetry, are valid beyond perturbation theory. Unfortunately, equation (13.31) shows that WT identities always involve correlation functions with one π -field at vanishing four-momentum. Therefore, they would only lead to exact predictions if the π -mesons were massless, that is, if the symmetry were spontaneously broken. In reality, it is necessary to extrapolate from zero momentum to the pion mass-shell. The extrapolation is model dependent, and the results are reliable only if the predictions at zero pion mass are already in qualitative agreement with experiment.

Note that, because the scale separation between the hadron picture and QCD is not extremely large, *non-normalizable chiral symmetric interactions* have eventually also to be taken into account. This leads, in the low momentum, low quark mass regime, and in the spirit of effective field theories, to *chiral perturbation theory* [111]. However, WT identities [40] are not affected, because they depend only on the particle content and the symmetry properties of the action (13.71). A few classical results follow [112].

13.7.1 Scalar model

We first discuss again the reduced scalar model. The non-trivial part of the WT identities corresponds to the transformations (13.65). In terms of $\mathbf{J}(x)$, the source for the π -field, and $H(x)$, the source for the σ -field, the WT identity for the generating functional of connected correlation functions takes the form (equation (13.30)),

$$\int d^4x \left[J_i(x) \frac{\delta}{\delta H(x)} - (c + H(x)) \frac{\delta}{\delta J_i(x)} \right] \mathcal{W}(\mathbf{J}, H) = 0. \quad (13.76)$$

It is convenient to introduce some additional notation to take into account the residual $O(3)$ symmetry. We set, in the Fourier representation,

$$\begin{aligned} \tilde{W}_{ij}^{(2)}(p) &= \delta_{ij} B_\pi^{-1}(p), & \tilde{W}^{(2)}(p) &= B_\sigma^{-1}(p), \\ \tilde{W}_{ij}^{(3)}(p_1, p_2; p_3) &= \delta_{ij} B_\pi^{-1}(p_1) B_\pi^{-1}(p_2) B_\sigma^{-1}(p_3) C(p_1, p_2; p_3), \\ \left[\tilde{W}_{ijkl}^{(4)}(p_1, p_2, p_3, p_4) \right]_{\text{amp.}} &= \delta_{ij} \delta_{kl} A(p_1, p_2, p_3, p_4) + \delta_{ik} \delta_{jl} A(p_1, p_3, p_2, p_4) \\ &\quad + \delta_{il} \delta_{kj} A(p_1, p_4, p_3, p_2), \end{aligned}$$

with the conventions that indices correspond to π -fields, and in mixed $\pi-\sigma$ correlation functions the arguments of the π -fields are written first.

Differentiating equation (13.76) with respect to J_j and setting the sources to zero, one obtains a special example of equation (13.43):

$$v = \langle \sigma \rangle = c/B_\pi(0), \quad (13.77)$$

where $B_\pi(0)$, the value of the inverse of the π propagator at zero momentum, is now different from the pion mass squared m_π^2 , which is given by $B_\pi(p)|_{p^2=-m_\pi^2} = 0$.

Differentiating equation (13.76) once with respect to J_j and H , one obtains

$$\delta_{ij} \tilde{W}^{(2)}(p) - \tilde{W}_{ij}^{(2)}(p) = c \tilde{W}_{ij}^{(3)}(0, p; -p),$$

and thus, using equation (13.77),

$$B_\pi(p) - B_\sigma(p) = v C(0, p; -p). \quad (13.78)$$

In particular, for $p = 0$, the equation becomes

$$B_\pi(0) - B_\sigma(0) = v C(0, 0; 0). \quad (13.79)$$

Differentiating thrice with respect to J , one finds a relation between three- and four-point correlation functions,

$$c \tilde{W}_{ijkl}^{(4)}(0, p_2, p_3, p_4) = \delta_{ij} \tilde{W}_{kl}^{(3)}(p_3, p_4; p_2) + \text{2 terms}. \quad (13.80)$$

It follows that

$$v A(0, p_2, p_3, p_4) = C(p_3, p_4; p_2) B_\sigma^{-1}(p_2) B_\pi(p_2). \quad (13.81)$$

First, for $p_2^2 = -m_\pi^2$, B_π vanishes and the equation reduces to Adler's consistency condition,

$$A(0, p_2, p_3, p_4)|_{p_2^2 = -m_\pi^2} = 0. \quad (13.82)$$

Moreover, setting $p_3 = 0$ in (13.81) and eliminating the function C between (13.78) and (13.81), one finds

$$v^2 A(0, p, 0, -p) = B_\pi(p) [B_\sigma^{-1}(p) B_\pi(p) - 1]. \quad (13.83)$$

The first term in the right-hand side has a double zero at the pion mass. Therefore, taking the derivative with respect to p^2 , one recovers Weinberg's relation,

$$v^2 \frac{\partial}{\partial p^2} [A(0, p, 0, -p) + B_\pi(p)] \Big|_{p^2 = -m_\pi^2} = 0. \quad (13.84)$$

These equations yield model- and parameter-independent constraints on the π - π scattering amplitude, which unfortunately is slightly off-shell because at least one of the π momenta vanishes. One verifies immediately that the function A in the tree approximation (13.70) satisfies both conditions (13.82) and (13.84).

Another constraint on the π - π scattering amplitude is obtained, for example, by setting all momenta to zero in (13.83):

$$v^2 A(0, 0, 0, 0) = B_\pi(0) [B_\sigma^{-1}(0) B_\pi(0) - 1]. \quad (13.85)$$

However, the equation involves an independent free quantity $B_\sigma(0)$. Again, one verifies that expression (13.70) satisfies equation (13.85) in the tree approximation.

13.7.2 Full boson–fermion model

We denote by $\bar{\eta}$ and η the sources for the fermion fields. From the invariance of the action under the transformations (13.72), one derives for the generating functional $\mathcal{W}(\eta, \bar{\eta}, J, H)$ of connected correlation functions the WT identity,

$$\int d^4x \left\{ \frac{i}{2} \left[\bar{\eta}(x) \gamma_5 \tau \frac{\delta}{\delta \bar{\eta}(x)} - \eta(x) \gamma_5 \tau \frac{\delta}{\delta \eta(x)} \right] - \mathbf{J}(x) \frac{\delta}{\delta H(x)} + (H(x) + c) \frac{\delta}{\delta \mathbf{J}(x)} \right\} \mathcal{W}(\eta, \bar{\eta}, \mathbf{J}, H) = 0. \quad (13.86)$$

The relations relevant for particle physics correspond to $H = 0$. The simplest and most famous identity is obtained by differentiating with respect to η and $\bar{\eta}$ and setting all sources to zero. It is, actually, most conveniently written in terms of vertex functions:

$$v \tilde{\Gamma}_{\pi NN}^{(3)}(0; p, -p) = \frac{i\tau}{2} \left\{ \gamma_5, \tilde{\Gamma}_{NN}^{(2)}(p) \right\}_+. \quad (13.87)$$

The index $+$ in the right-hand side means anticommutator in the space of γ matrices. We have explicitly taken into account the property that the fermion propagator is proportional to the identity in the group indices. This relation between the inverse nucleon propagator and the πNN vertex generalizes the relation (13.74). It has a physical interpretation in terms of the weak current under the name of Goldberger–Treiman relation. The right-hand side is known when the nucleons are on mass-shell. The left-hand side can approximately be related to the nucleon weak β -decay, which involves the matrix element of the axial current at zero momentum between nucleon states: since the pion has the quantum numbers of the divergence of the axial current, as can be seen in equation (13.66), one contribution to this matrix element has the pion pole. In the strict chiral limit with zero mass pions, this would be the only contribution. One assumes that since the pion mass is small, the chiral limit is a good approximation. The relation then becomes in traditional notation

$$\frac{G_A}{G_V} \simeq g_{\pi NN} \frac{f_\pi}{m_N}. \quad (13.88)$$

Replacing by experimental numbers one finds 1.22 for the left-hand side and 1.36 for the right-hand side, a notable improvement over the tree approximation (13.75).

More generally, one can set H to zero, differentiate once with respect to η and $\bar{\eta}$, and an arbitrary number of times with respect to \mathbf{J} and, finally, set all momenta on mass-shell. One then obtains model-independent relations (generalizing equation (13.82)) involving amplitudes for the emission of one non-physical pion at zero momentum. However, to determine completely the low energy π – N scattering amplitude, it is necessary to introduce also the $NN\sigma$ vertex and the result then depends on the σ -mass. The predictions for the π – N scattering lengths, in the infinite σ -mass limit, agree reasonably well with experimental results.

13.8 Quadratic symmetry breaking

A symmetry may be broken by terms of higher dimensions. Here, we give a detailed discussion only of the case of the quadratic symmetry breaking [105]. We then briefly indicate how the results generalize in the case of breaking terms of higher dimensions. We consider the action

$$\mathcal{S}(\phi) = \mathcal{S}_{\text{sym.}}(\phi) + \frac{1}{2} \sum_{i,j} \mu_{ij} \int d^d x \phi_i(x) \phi_j(x), \quad (13.89)$$

in which μ_{ij} is a constant matrix, symmetric and traceless, since a term proportional to the unit matrix can always be absorbed into the symmetric action $\mathcal{S}_{\text{sym.}}(\phi)$. In the action (13.89), the interactions are symmetric, but the mass terms break the symmetry.

It is simple to verify that a straightforward generalization of the method used in the case of the linear breaking does not work.

The difficulties one encounters have several origins. One can be directly understood by formally expanding the correlation functions in a power series of the symmetry breaking term. This generates a sum of symmetric correlation functions with multiple insertions of the operator $\frac{1}{2} \sum_{i,j} \mu_{ij} \int d^d x \phi_i(x) \phi_j(x)$ (see Fig. 13.2).

These insertions, as we have already discussed in Chapter 11, may generate new divergences which have to be taken care of. This situation has to be contrasted with the linear case, in which only field correlation functions are generated by the corresponding expansion.

Another difficulty stems from the fact that an infinitesimal group transformation generates a new quadratic term linearly independent of the initial one, as can be seen in equation (13.89). To be able to write WT identities for renormalized quantities, one has to also renormalize the operator $\int d^d x \mu_{ij} t_{jk}^\alpha \phi_i(x) \phi_k(x)$ and examine how it transforms under the group. This may generate new quadratic operators and all the operations have to be repeated. Thus, a more general strategy is required.

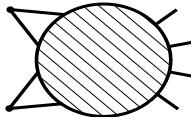


Fig. 13.2 Double insertion of quadratic symmetry breaking terms

The general method. In the example of quadratic symmetry breaking, it is clear that any infinitesimal transformation made on a linear combination of operators of the form $\phi_i(x) \phi_j(x)$ generates another linear combination of these same operators. Therefore, we immediately consider the general action

$$\mathcal{S}(\phi, K) = \mathcal{S}_{\text{sym.}}(\phi) + \frac{1}{2} \int d^d x \sum_{i,j} K_{ij}(x) \phi_i(x) \phi_j(x), \quad (13.90)$$

with

$$K_{ij}(x) = K_{ji}(x), \quad \text{tr } \mathbf{K}(x) = 0.$$

Since we need only insertions at zero momentum, we could restrict the analysis to constant sources K_{ij} . However, it is useful, and not more difficult, to work with space-dependent sources. Indeed, when the symmetric theory is massless, zero momentum insertions could lead to infrared divergences, which in this way are avoided.

We define the corresponding generating functional,

$$\mathcal{Z}(J, K) = \int [d\phi] \exp [-\mathcal{S}(\phi, K, J)], \quad (13.91)$$

with

$$\mathcal{S}(\phi, K, J) = \mathcal{S}_{\text{sym.}}(\phi) + \int d^d x \left[\frac{1}{2} \sum_{i,j} K_{ij}(x) \phi_i(x) \phi_j(x) - \mathbf{J}(x) \cdot \boldsymbol{\phi}(x) \right].$$

An infinitesimal change of variable (13.14) leads to

$$0 = \int [d\phi] \delta_\alpha \mathcal{S}(\phi, K, J) \exp [-\mathcal{S}(\phi, K, J)], \quad (13.92)$$

with

$$\delta_\alpha \mathcal{S}(\phi, K, J) = \int d^d x \sum_{i,j} t_{ij}^\alpha \phi_j(x) \left[\sum_k K_{ik}(x) \phi_k(x) - J_i(x) \right]. \quad (13.93)$$

As before, the product $\phi_j(x) \phi_k(x)$ appears, but we are now able to express it in terms of $\delta/\delta K_{kj}(x)$ instead of $\delta^2/\delta J_j(x)\delta/J_k(x)$.

Remark. We define derivatives with respect to complicated objects (here the symmetric traceless matrix K_{ij}) in the following way: let $F(\mathbf{K})$ be a function (or functional) of \mathbf{K} ; we calculate the variation of $F(\mathbf{K})$ at first order when \mathbf{K} varies by a quantity $\delta\mathbf{K}$,

$$F(\mathbf{K} + \delta\mathbf{K}) - F(\mathbf{K}) = \sum_{i,j} \int \frac{\delta F}{\delta K_{ij}(x)} \delta K_{ij}(x) d^d x + O(\|\delta\mathbf{K}\|^2). \quad (13.94)$$

The derivative is then defined in the sense of differential geometry: it is the linear operator acting on $\delta K_{ij}(x)$ in the right-hand side of equation (13.94).

In the present example, differentiation with respect to K_{ij} generates the traceless part of the product $\phi_i \phi_j$. Since equation (13.93) involves only the traceless part of the same product, the definition (13.94) makes it possible to express equation (13.92) in the form,

$$\int d^d x \left\{ 2 \sum_{i,j,k} t_{ij}^\alpha K_{ik}(x) \frac{\delta}{\delta K_{kj}(x)} - \sum_{i,j} J_i(x) t_{ij}^\alpha \frac{\delta}{\delta J_j(x)} \right\} \mathcal{Z}(J, K) = 0. \quad (13.95)$$

Because equation (13.95) is a first order functional differential equation, an identical equation holds for $\mathcal{W}(J, K)$.

The Legendre transformation is only performed with respect to the source $J_i(x)$, because the reducibility corresponds only to external ϕ lines:

$$\Gamma(\varphi, K) + \mathcal{W}(J, K) = \int dx \mathbf{J}(x) \cdot \boldsymbol{\varphi}(x), \quad \varphi_i(x) = \frac{\delta \mathcal{W}(J, K)}{\delta J_i(x)}.$$

The sources $K_{ij}(x)$ do not participate in the Legendre transformation and have to be considered as external parameters. This immediately implies (equation (7.65)):

$$\left. \frac{\delta \mathcal{W}}{\delta K_{ij}(x)} \right|_J = - \left. \frac{\delta \Gamma}{\delta K_{ij}(x)} \right|_\varphi. \quad (13.96)$$

The relation provides an additional justification for the method we have proposed. With sources for the composite operators, the effect of the Legendre transformation on the WT identity (13.95) is simple, and for $\Gamma(\varphi, K)$ leads to

$$\int d^d x \left\{ 2 \sum_{i,j,k} t_{ij}^\alpha K_{ik}(x) \frac{\delta \Gamma(\varphi, K)}{\delta K_{kj}(x)} + \sum_{i,j} \varphi_i(x) t_{ij}^\alpha \frac{\delta \Gamma(\varphi, K)}{\delta \varphi_j(x)} \right\} = 0. \quad (13.97)$$

The equation has a straightforward interpretation: $\Gamma(\varphi, K)$ is invariant under the combined, infinitesimal group transformations

$$\begin{cases} \delta_\omega \varphi_i(x) = \sum_{j,\alpha} \omega_\alpha t_{ij}^\alpha \varphi_j(x), \\ \delta_\omega K_{kj}(x) = \sum_{i,\alpha} \omega_\alpha [t_{ji}^\alpha K_{ik}(x) + t_{ki}^\alpha K_{ij}(x)]. \end{cases} \quad (13.98)$$

By performing a transformation both on ϕ and \mathbf{K} , we have rendered the breaking term $\frac{1}{2} \int d^d x \sum_{i,j} K_{ij}(x) \phi_i(x) \phi_j(x)$ group invariant. It is then almost obvious that the symmetry of the action (13.90) under the equivalent of the transformations (13.98) implies the WT identities (13.97).

Using the arguments given for the case $\mathbf{K} = 0$, one can show that, if the regularized functional $\Gamma(\varphi, K)$ satisfies equation (13.97), the renormalized functional $\Gamma(\varphi, K)$ and the renormalized action satisfy the same identity: the renormalized action $\mathcal{S}_r(\phi, K)$ is the most general local functional of ϕ and K compatible with power counting, and invariant under the group transformation (13.98).

Renormalization in dimension 4. For a ϕ^4 -like field theory in four dimensions, the action is the integral of a local function of dimension 4, the field ϕ has dimension $[\phi] = 1$, and the source \mathbf{K} has dimension $[K] = 2$, because it is coupled to an operator of dimension 2 (see Section 11.1):

$$[\phi] = 1, \quad [K] = 2, \quad [\mathcal{S}(\phi, K)] = 0.$$

The renormalized action $\mathcal{S}_r(\phi, K)$ thus has the general form

$$\begin{aligned} \mathcal{S}_r(\phi, K) &= [\mathcal{S}_{\text{sym.}}]_r(\phi) + \frac{1}{2} \int d^4 x \sum_{i,j} K_{ij}(x) A_{ij}(\phi(x)) \\ &\quad + \frac{1}{2} \sum_{i,j,k,l} b_{ij,kl} \int d^4 x K_{ij}(x) K_{kl}(x), \end{aligned} \quad (13.99)$$

in which $A_{ij}(\phi)$ is a local derivative-free polynomial of dimension 2 and $b_{ij,kl}$ is a set of constants.

Constraints on $A_{ij}(\phi)$ and $b_{ij,kl}$ are obtained by expressing the invariance of the renormalized action under transformation (13.98). In the simplest case, $A_{ij}(\phi)$ has the form

$$A_{ij}(\phi(x)) = Z_2 (\phi_i(x) \phi_j(x) - \frac{1}{N} \delta_{ij} \phi^2(x)), \quad (13.100)$$

in which N is the number of components of ϕ .

Depending on the representation content of $\phi_i(x)$, it can be a linear combination of several dimension 2 operators, and it also contains a contribution linear in ϕ .

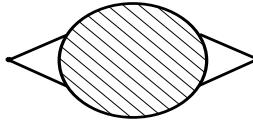


Fig. 13.3 Quadratic symmetry breaking terms: two-point function

The term quadratic in K_{ij} does not depend on ϕ , and can be factorized in front of the field integral. It gives an additive contribution to $\mathcal{W}(J, K)$. It yields an additive renormalization for the correlation function $\delta^2 \mathcal{W}(J = 0) / \delta K_{ij}(x) \delta K_{kl}(y)|_{K=0}$ which has the diagrammatic structure displayed in Fig. 13.3.

The renormalized action in the case of a specific quadratic symmetry breaking is then obtained by setting $K_{ij}(x)$ to be a constant:

$$K_{ij}(x) = \mu_{ij}.$$

The analysis of this more complex situation shows that the initial form of the action is not always completely preserved. In particular, a breaking term of a given dimension may generate new breaking terms of lower dimensions. This happens here when $A_{ij}(\phi)$ has a term linear in ϕ . Therefore, the renormalized QFT may depend on more parameters than one would have naively anticipated when adding the breaking term.

Finally, as in the case of the linear symmetry breaking, one can use the WT identities (13.97) to constrain the renormalization conditions. However, it is then necessary to consider together all the superficially divergent correlation functions, both of the field $\phi_i(x)$ and the composite operator $\phi_i(x)\phi_j(x)$, and write all relations derived from equation (13.97) by expanding in $\varphi(x)$ and $K(x)$, in which they appear.

Breaking terms of higher canonical dimensions. The preceding analysis can be generalized to breaking terms of higher canonical dimensions. One verifies that in a ϕ^4 -like theory in four dimensions, a cubic breaking term, since it is coupled to a source of canonical dimension 1, generates in general, by renormalization, breaking terms quadratic and linear in ϕ . Breaking terms which are of canonical dimension lower than the symmetric interaction are called *soft*.

Finally, it makes little sense in the context of a renormalizable ϕ_4^4 theory to speak of a breaking term of dimension 4. Indeed, such a term is coupled to a source of dimension 0. The renormalized action will contain an infinite series in the source. One verifies then that all traces of the initial symmetry are lost after renormalization.

13.8.1 Discrete symmetries

Discrete symmetries do not lead to WT identities and most of the preceding analysis does not apply. However, it is easy to prove that when the initial unrenormalized action is symmetric, the renormalized action remains symmetric. Correlation functions in the presence of an additional linear symmetry breaking term can be expanded in power series of the breaking parameter. The coefficients are symmetric correlation functions. Therefore, it remains true that the counter-terms which render the symmetric QFT finite renormalizes the QFT with linear symmetry breaking. For breaking terms of higher degree, the same strategy can be applied. However, then, symmetric correlation functions with operator insertions have to be renormalized. The question has been examined in Section 11.1. A possible strategy consists in adding to the action sources for the operators, renormalizing according to power counting, and using the discrete symmetry to constrain the polynomial in the sources and the field. One general property of the preceding analysis survives: symmetry breaking terms of a given dimension can only generate additional breaking terms of equal or lower dimensions.

A13 Currents and Noether's theorem

In the appendix, in order to discuss properties of the relativistic classical equations of motion, in the two first sections we adopt the covariant notation of real-time QFT with a metric tensor $g_{\mu\nu}$ and a metric of signature $(+ - - \dots)$. *Summation over successive upper and lower indices will be implied* (see also Chapter 28). In the two last sections, we return to Euclidean formalism and explicit summations.

A13.1 Currents in classical-field theory

If the Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$ depends only on the field $\phi(x)$ and its derivatives $\partial_\mu \phi(x)$, the classical equation of motion obtained by varying the action,

$$\mathcal{S}(\phi) = \int d^d x \mathcal{L}(\phi(x), \partial_\mu \phi(x)), \quad (A13.1)$$

is

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi(x)]} - \frac{\partial \mathcal{L}}{\partial \phi(x)} = 0. \quad (A13.2)$$

(In this notation, $\phi(x)$ and $\partial_\mu \phi(x)$ are considered as independent variables.)

If we perform on $\phi(x)$ a space-dependent group transformation parametrized by a field $\Lambda(x)$,

$$\phi \mapsto \phi_\Lambda,$$

as a consequence of the equation of motion, the action is also stationary with respect to variations of $\Lambda(x)$ at ϕ fixed:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial [\partial_\mu \Lambda(x)]} - \frac{\partial \mathcal{L}}{\partial \Lambda(x)} = 0. \quad (A13.3)$$

We define a current $J^\mu(x)$, functional of $\phi(x)$, by

$$J^\mu(x) = \left. \frac{\partial \mathcal{L}}{\partial [\partial_\mu \Lambda(x)]} \right|_{\Lambda(x)=0}, \quad (A13.4)$$

in which we have assumed that $\Lambda(x) = 0$ corresponds in the group to the identity. By construction, currents are directly associated with the generators of the Lie algebra of the symmetry group.

We can then rewrite equation (A13.3) as

$$\partial_\mu J^\mu(x) = \frac{\partial \mathcal{L}}{\partial \Lambda(x)}, \quad (A13.5)$$

which is Noether's theorem.

If a Lagrangian is invariant under space-independent group transformations, $\partial \mathcal{L}/\partial \Lambda$ vanishes and thus, the current J_μ is conserved,

$$\partial_\mu J^\mu(x) = 0. \quad (A13.6)$$

In a classical field theory, the space integral of the time component of the current,

$$\mathcal{Q}^\alpha(t \equiv x_0) = \int d^{d-1}x J_0^\alpha(x), \quad (A13.7)$$

is a *charge*. By differentiating with respect to time t , and using the current conservation equation (A13.6), one finds

$$\frac{d}{dt} \mathcal{Q}^\alpha(t) = \int d^{d-1}x \sum_{\mu=1}^{d-1} \partial_\mu J_\mu(x) = 0.$$

The charges $\mathcal{Q}^\alpha(t)$ are constants of the classical motion.

Example. If the Lagrangian density for an N -component field ϕ has the form (in real-time covariant notation)

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi(x) \cdot \partial^\mu \phi(x) - V[\phi(x)], \quad (A13.8)$$

where $V(\phi)$ is group invariant, and if the infinitesimal group transformations are

$$\delta \phi_i(x) = \sum_{i,j} t_{ij}^\alpha \Lambda_\alpha(x) \phi_j(x), \quad (A13.9)$$

(the matrices t^α are generators of the group) the current is given by

$$J^{\mu,\alpha}(x) = \sum_{i,j} t_{ij}^\alpha \partial^\mu \phi_i(x) \phi_j(x). \quad (A13.10)$$

A13.2 The energy-momentum tensor

If the action is translation invariant, the substitution $\phi(x) \mapsto \phi(x + \varepsilon)$, in which ε is a constant, leaves the action invariant. In the spirit of Section A13.1, we perform a space-dependent translation, which in fact coincides with a general change of variables (see also Section 28.1). We thus substitute in the action $\phi(x) \mapsto \phi(x + \varepsilon(x))$. If $\phi(x)$ satisfies the equation of motion, the variation of the action (A13.1) at first order in ε vanishes. In the substitution, the derivatives transform like

$$\partial_\mu \phi(x) \mapsto \partial_\mu \phi(x + \varepsilon) + \partial_\mu \varepsilon^\nu(x) \partial_\nu \phi(x + \varepsilon).$$

To calculate the variation, we then change variables, setting $x + \varepsilon = y$. Translation invariance implies that the action density depends on x only through the field ϕ . Therefore, the only new effect is to change the integration measure:

$$dy^\mu = dx^\mu + \partial_\nu \varepsilon^\mu(x) dx^\nu.$$

Comparing the new action with the initial one (A13.1), we note that the modifications come only from the derivatives and the integration measure (y is a dummy integration variable). Collecting the terms of order ε and integrating by parts, one obtains the identity

$$\partial_\mu T_\nu^\mu(x) = 0, \quad (A13.11)$$

in which the *energy-momentum tensor* $T_\nu^\mu(x)$ is defined by

$$T_\nu^\mu(x) = \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi(x)]} \partial_\nu \phi(x) - \delta_\nu^\mu \mathcal{L}[\phi(x)]. \quad (A13.12)$$

It is convenient to also introduce the tensor

$$T_{\mu\nu}(x) = g_{\mu\lambda} T^\lambda_\nu(x), \quad (A13.13)$$

in which $g_{\mu\nu}$ is the Minkowski metric tensor. In the example of the Lagrangian (A13.8),

$$T_{\mu\nu}(x) = \partial_\mu \phi(x) \partial_\nu \phi(x) - g_{\mu\nu} \left[\frac{1}{2} (\partial_\rho \phi(x)) (\partial^\rho \phi(x)) - V(\phi(x)) \right], \quad (A13.14)$$

and thus, $T_{\mu\nu}$ is a symmetric tensor.

To the tensor $T_{\mu\nu}(x)$ correspond constants of the classical motion, the energy and momentum P_μ , obtained by integrating the time components (with respect to one index) of $T_{\mu\nu}$ over space:

$$P_\mu(t \equiv x_0) = \int d^{d-1}x T_{0\mu}(x). \quad (A13.15)$$

Then,

$$\frac{d}{dt} P_\mu = 0. \quad (A13.16)$$

We note that a space-time dependent change of variables on x^μ is an arbitrary change of coordinates. This explains that the tensor $T_{\mu\nu}$ appears in the coupling of matter field to the metric tensor in Einstein's theory of *general relativity* (for details see Chapter 28 and the corresponding references).

Any current associated with an additional space-time symmetry of the action can be related to $T_{\mu\nu}$.

For instance, the $O(1, d - 1)$ pseudo-orthogonal transformations whose infinitesimal form is

$$\delta x^\mu = \Lambda_\nu^\mu(x) x^\nu, \quad (A13.17)$$

correspond to the choice

$$\varepsilon^\mu(x) = \Lambda_\nu^\mu(x) x^\nu. \quad (A13.18)$$

The corresponding currents $M^{\mu\nu\rho}$ are then

$$M^{\mu\nu\rho}(x) = T^{\mu\nu}(x) x^\rho - T^{\mu\rho}(x) x^\nu. \quad (A13.19)$$

Dilatation invariance. We again consider, as an example, the ϕ^4 field theory in four dimensions:

$$\mathcal{L}(\phi) = \frac{1}{2} \partial_\mu \phi(x) \partial^\mu \phi(x) - \frac{1}{2} m^2 \phi^2(x) - \frac{1}{4!} g \phi^4(x). \quad (A13.20)$$

For $m = 0$, the action is scale-invariant, that is, invariant in the substitution

$$\phi(x) \mapsto \phi_\lambda(x) = \lambda \phi(\lambda x). \quad (A13.21)$$

For what concerns the variation of the argument, dilatation corresponds to taking ε^μ of the form

$$\varepsilon^\mu(x) = x^\mu \lambda(x). \quad (A13.22)$$

We thus expect the dilatation current S^μ to involve $x^\nu T_\nu^\mu$. A short calculation leads to

$$S^\mu(x) = x^\nu \left[T_\nu^\mu(x) + \frac{1}{6} (\partial^2 \delta_\nu^\mu - \partial_\mu \partial^\nu) \phi^2(x) \right]. \quad (A13.23)$$

For $m \neq 0$, the current $S^\mu(x)$ is not conserved. One finds,

$$\partial_\mu S^\mu(x) = m^2 \phi^2(x). \quad (A13.24)$$

We now introduce the tensor

$$\tilde{T}_{\mu\nu}(x) = T_{\mu\nu}(x) + \frac{1}{6} (\partial^2 g_{\mu\nu} - \partial_\mu \partial_\nu) \phi^2(x). \quad (A13.25)$$

The tensor $\tilde{T}_{\mu\nu}$ can be used as an energy–momentum tensor, in place of $T_{\mu\nu}$: it is a polynomial in the field, symmetric as a tensor, and satisfies the conservation equation

$$\partial_\mu \tilde{T}_\nu^\mu = 0. \quad (A13.26)$$

In terms of $\tilde{T}_\nu^\mu(x)$, equation (A13.23) then reads

$$S^\mu(x) = x^\nu \tilde{T}_\nu^\mu(x), \quad (A13.27)$$

and the divergence of the dilatation current is

$$\partial_\mu S^\mu(x) = \tilde{T}_\mu^\mu(x). \quad (A13.28)$$

In dilatation-invariant classical theories, the trace of the ‘improved’ energy–momentum tensor \tilde{T}_ν^μ vanishes [113].

A13.3 Euclidean theory: Dilatation and conformal invariance

We now return to Euclidean formulation. We use only lower indices, and summations are explicit.

We consider a general *Euclidean* action \mathcal{S} , invariant under translation, rotation, and *dilatation*, a property of some massless field theories. We perform the infinitesimal change of variables

$$x_\mu \mapsto x_\mu + \varepsilon_\mu(x). \quad (A13.29)$$

Translation invariance implies that the variation of the action \mathcal{S} involves only the partial derivatives of $\varepsilon_\mu(x)$:

$$\delta\mathcal{S} = \int d^d x \sum_{\mu,\nu} T_{\mu\nu}(x) \partial_\mu \varepsilon_\nu(x), \quad (A13.30)$$

where $T_{\mu\nu}$ is also called the *stress tensor*. Rotation invariance implies that $\delta\mathcal{S}$ vanishes for

$$\varepsilon_\nu = \sum_\rho \Lambda_{\nu\rho} x_\rho \Rightarrow \partial_\mu \varepsilon_\nu = \Lambda_{\nu\mu}, \quad (A13.31)$$

in which $\Lambda_{\mu\nu}$ is an arbitrary antisymmetric matrix. Therefore, the integral of the stress tensor must be symmetric:

$$\int d^d x [T_{\mu\nu}(x) - T_{\nu\mu}(x)] = 0.$$

Dilatation invariance corresponds to

$$\varepsilon_\nu = \lambda x_\nu, \quad (A13.32)$$

and implies the vanishing of the integral of the covariant trace of the stress tensor:

$$\int d^d x \sum_\mu T_{\mu\mu}(x) = 0.$$

A13.3.1 The conformal group

For the simplest class of theories, like scalar field theories with an action $\mathcal{S}(\phi)$ depending only on the field $\phi(x)$ and its *first* partial derivatives, the two integral conditions imply the existence of a symmetric, traceless *stress-energy tensor*:

$$T_{\mu\nu} = T_{\nu\mu}, \quad \sum_{\mu} T_{\mu\mu} = 0.$$

It then follows that the variation of the action also vanishes for any function ε_μ that satisfies

$$\partial_\mu \varepsilon_\nu(x) + \partial_\nu \varepsilon_\mu(x) - \frac{2}{d} \delta_{\mu\nu} \nabla \cdot \varepsilon(x) = 0, \quad (A13.33)$$

where d is the dimension of Euclidean space. The group of transformations which satisfy equation (A13.33) is larger than the product of transformations that we have considered so far: it is the whole *conformal group*. Indeed, let us calculate the variation of a line element of the form

$$(ds)^2 = g(x) dx \cdot dx, \quad (A13.34)$$

which corresponds to a conformally flat metric.

We find

$$\delta [(ds)^2] = \sum_{\mu, \rho} dx_\mu dx_\mu \partial_\rho g(x) \varepsilon_\rho(x) + \sum_{\mu, \nu} dx_\mu dx_\nu (\partial_\mu \varepsilon_\nu(x) + \partial_\nu \varepsilon_\mu(x)) g(x). \quad (A13.35)$$

We see that equation (A13.33) is the necessary and sufficient condition for the line element to retain the form (A13.34). By definition, the transformations that preserve the form (A13.34) of the metric are conformal transformations.

Returning to equation (A13.33), one verifies that the solutions of degree 0 in x_μ correspond to translations, and the solutions of degree 1 correspond to rotations and dilatations. However, additional solutions of degree 2 can be found of the form,

$$\varepsilon_\mu(x) = a_\mu x^2 - 2x_\mu a \cdot x, \quad (A13.36)$$

where a_μ is a constant vector. They correspond to special conformal transformations. The integrated form of these transformations is

$$x'_\mu = \frac{x_\mu + a_\mu x^2}{1 + 2a \cdot x + a^2 x^2}. \quad (A13.37)$$

Together, all the transformations generate the *conformal group*, which is isomorphic to the group $SO(d+1, 1)$.

Two dimensions. In dimension $d = 2$, the set of equations (A13.33) reduces to

$$\partial_1 \varepsilon_2 + \partial_2 \varepsilon_1 = 0, \quad \partial_1 \varepsilon_1 - \partial_2 \varepsilon_2 = 0, \quad (A13.38)$$

which are just the Cauchy conditions and express the well-known property that all analytic transformations are conformal. The conformal group has an *infinite number of generators*. The implications of conformal invariance are much stronger than in generic dimensions, and lead to the classification of a whole class of conformal-invariant field theories [114].

Remark. The condition that the action should depend only on the field and its first derivatives can be illustrated by the simple counter-example of the quadratic action,

$$\mathcal{S}(\phi) = \int d^d x (\nabla^2 \phi(x))^2.$$

The propagator in Fourier space is $1/p^4$. The theory is obviously translation, rotation, and scale invariant. However, it is simple to verify that it is not conformal invariant.

Conformal invariance beyond the classical theory. We have seen in Chapter 9 that the scale invariance of the classical theory is broken at the quantum level. However, as we discuss in Chapter 15, at fixed points of the renormalization group, dilatation invariance and, therefore, conformal invariance are both restored [115].

Imposing conformal invariance on correlation functions determines, in particular, two- and three-point functions.

A13.4 QFT: Currents and correlation functions

Dimensional regularization is assumed here (see Section 10.1).

We have already examined the consequences of symmetries for field theories, and derived WT identities. These identities can also be derived in the operator formalism of quantum mechanics and in this case currents and charges, considered as quantum operators, play an important role. In our formulation, currents appear either in the coupling at leading order of matter to gauge fields (see Chapters 21 and 22), or as polynomials in the fields (operators in the sense of Chapter 11) satisfying identities that we now derive, and that imply special renormalization properties.

Therefore, we consider the generating functional (we set $\hbar = 1$)

$$\mathcal{Z}(J) = \int [d\phi] \exp \left[-\mathcal{S}(\phi) + \int d^d x \mathbf{J}(x) \cdot \phi(x) \right], \quad (A13.39)$$

in which the action \mathcal{S} is invariant under group transformations whose infinitesimal form is given by equation (A13.9), when $\Lambda(x)$ is a constant.

We perform a change of variables of the form of a transformation (13.14) in the integral (A13.39). We define the current $J^{\mu,\alpha}(x)$ by equation (A13.4) in terms of the action density. If $\mathcal{S}(\phi)$ is symmetric, the variation of the action reads

$$\delta\mathcal{S}(\phi) = \int d^d x \nabla \cdot \sum_{\alpha} \Lambda_{\alpha}(x) \mathbf{J}^{\alpha}(x). \quad (A13.40)$$

Identifying the coefficient of $\Lambda_{\alpha}(x)$, one obtains

$$\int [d\phi] \left[\nabla \mathbf{J}^{\alpha}(x) - \sum_{i,j} J_i(x) t_{ij}^{\alpha} \phi_j(x) \right] \exp \left[-\mathcal{S}(\phi) + \int d^d x \mathbf{J}(x) \cdot \phi(x) \right] = 0. \quad (A13.41)$$

This identity can be written as

$$\nabla_x \mathcal{Z}_{\mathbf{J}^{\alpha}}(x) = \sum_{i,j} J_i(x) t_{ij}^{\alpha} \frac{\delta \mathcal{Z}}{\delta J_j(x)}, \quad (A13.42)$$

where $\mathcal{Z}_{\mathbf{J}^{\alpha}}(x)$ is the generating functional of correlation functions with a $\mathbf{J}^{\alpha}(x)$ operator insertion.

The same equation is valid for connected correlation functions. After Legendre transformation, one finds

$$\nabla_x \Gamma_{\mathbf{J}^{\alpha}(x)} = - \sum_{i,j} \frac{\delta \Gamma}{\delta \varphi_i(x)} t_{ij}^{\alpha} \varphi_j(x). \quad (A13.43)$$

Equations (A13.42) and (A13.43) are the analogues for correlation functions of the current conservation equation (A13.6). Integrated over whole space, they yield, not surprisingly, equations (13.16–13.18), that is, the WT identities of the symmetry.

From the point of view of renormalization, equation (A13.43) implies that the insertion of $\nabla_x \mathbf{J}^\alpha(x)$ in a renormalized correlation function is finite.

In a simple renormalizable $\phi_{d=4}^4$ -like QFT, covariance then implies that the same must be true for the current $\mathbf{J}^\alpha(x)$. This result is non-trivial, since expression (A13.10) shows that $J^{\mu,\alpha}(x)$ is an operator of dimension 3. A further consequence is that the insertion of a conserved current in a correlation function does not modify the form of RG equations.

A13.5 Energy-momentum tensor and QFT

Performing the infinitesimal change of variables

$$\phi(x) = \phi'(x + \varepsilon(x)), \quad (A13.44)$$

in the field integral, one derives WT identities for the insertion of the energy-momentum tensor (also called the *stress tensor*). The variation of the action with a source is

$$\delta \left[\int d^d x J(x) \phi(x) - \mathcal{S}(\phi) \right] = \int d^d x \sum_\nu \varepsilon^\nu(x) \left[J(x) \partial_\nu \phi(x) + \sum_\mu \partial_\mu T_\nu^\mu(x) \right]. \quad (A13.45)$$

It follows that

$$\partial_\mu^x \mathcal{Z}_{T_\nu^\mu(x)} + J(x) \partial_\nu^x \frac{\delta \mathcal{Z}}{\delta J(x)} = 0. \quad (A13.46)$$

An integration of the identity over space yields

$$\int d^d x J(x) \partial_\nu \frac{\delta \mathcal{Z}}{\delta J(x)} = 0, \quad (A13.47)$$

which expresses the translation invariance of correlation functions.

After Legendre transformation, one finds

$$\sum_\mu \partial_\mu^x \Gamma_{T_\nu^\mu(x)} + \frac{\delta \Gamma}{\delta \varphi(x)} \partial_\nu \varphi(x) = 0. \quad (A13.48)$$

Again, one concludes that the insertion of the operator $\sum_\mu \partial_\mu T_\nu^\mu(x)$ in a renormalized correlation function is finite. However, this does not imply that the insertion of $T_{\mu\nu}$ itself is finite. In the $\phi_{d=4}^4$ QFT, for example, $T_{\mu\nu}$ has dimension 4. The quantity $(\delta_{\mu\nu} \nabla^2 - \partial_\mu \partial_\nu) \phi^2$ is also a symmetric tensor of dimension 4, whose divergence vanishes. Therefore, it can appear as an additive counter-term in the renormalization of T_ν^μ :

$$[T_{\mu\nu}(x)]_r = T_{\mu\nu}(x) + K(\delta_{\mu\nu} \nabla^2 - \partial_\mu \partial_\nu)[\phi^2(x)]_r, \text{ with } K \text{ constant}. \quad (A13.49)$$

Note that, automatically, the renormalized energy-momentum tensor then has a non-vanishing trace, and can no longer be improved, since the constant K is divergent. The dilatation current is not conserved, but this should have been expected, since it is impossible to regularize the theory without breaking the classical dilatation invariance, either by introducing a cut-off, or by changing the dimension. Nevertheless, it is possible to derive WT identities involving the divergence of the dilatation current. By integrating them over space, one obtains the Callan–Symanzik equations introduced in Section 9.5.

14 Critical phenomena: General considerations. Mean-field theory (MFT)

We have referred to phase transitions in the framework of quantum field theory (QFT) in Chapter 13. We want now to examine the problem of phase transitions in macroscopic systems, in particular, in statistical lattice models, although some of these lattice models also appear as lattice regularizations of Euclidean (imaginary time) QFTs.

Most of the systems we consider in this work have the following character: a large number of spins on the lattice, or of macroscopic particles in the continuum, interact through *short range forces*. For simplicity, in this work we assume that interactions *decay exponentially*. At fixed density, as long as the system is contained in a finite volume, it is *ergodic*, that is, any finite region of available phase space has a non-vanishing probability to be explored in the course of time (in the sense of Section A34.1, the system is connected). However, in the infinite volume limit, depending on the value of a control parameter, which often is the temperature, the system either remains ergodic, or experiences a breaking of ergodicity. In the latter case, the phase space decomposes into disjoint sets. When the system is initially in one of the sets, it remains at later time. For example, for Ising-like systems, the two sets correspond to the two opposite values of the spontaneous magnetization.

For simple systems, it is possible to find a *local observable* whose expectation value depends on the phase in the several phase region. One calls such an observable *order parameter*. For example, it is the spin in ferromagnetic systems.

In classical lattice models with *finite range interactions*, the partition function can be expressed in terms of a *transfer matrix*. As we have already discussed in Section 2.5, the thermodynamic limit is related to the largest eigenvalue of the matrix. In the ergodic, disordered phase the corresponding eigenvector is unique while, in a non-ergodic (ordered) phase, it is degenerate. The demand that connected correlation functions satisfy cluster properties, selects specific ground states, corresponding to pure phases. The partition function is then obtained by only summing over a subset of configurations.

In the disordered phase, the connected correlation functions decrease exponentially with distance when two non-empty sets of points are separated (corresponding to the cluster property defined in Section 7.6). The *correlation length* characterizes the slowest decay rate of correlation functions (the inverse of the smallest physical mass in the particle physics framework). Because we are interested only in large scale properties, we discuss only the class of phase transitions for which the *correlation length diverges at the critical temperature* (continuous phase transitions). For such systems, it can be shown that, near the critical (or transition) temperature, some properties of thermodynamic functions, related to large scale physics, are *universal*.

Our goal, here and in the coming chapters, is to prove universality, and to determine the behaviour of thermodynamic quantities in the neighbourhood of a phase transition, in particular, their singularities at the critical temperature.

In this chapter, we study phase transitions within the framework of *MFT* (for early references, see Ref. [116]), which predicts properties consistent with *Landau's theory of critical phenomena* [117].

We discover and determine some *super-universal properties* of thermodynamic observables, in the sense that they do not depend on the dimension of space or the symmetries of the systems. We show that these universal properties can also be reproduced by calculating correlation functions with a perturbed Gaussian measure. However, since no phase transition can occur in one dimension for systems with short-range interactions, and no phase transitions with ordering for systems with continuous symmetries in two dimensions [118], we know a priori that the MFT cannot be correct in all dimensions.

We then define a general scheme in which MFT is only the leading term in a systematic expansion, and that, in general, we thus call *mean-field approximation*. An analysis of the corrections to the mean-field approximation, confirms that the leading correction diverges at the critical temperature for dimensions smaller than or equal to 4 (e.g. see Ref. [51]). This is the first indication that if, below four dimensions, some universality survives it cannot be mean-field like [119]. The deep reason is that the microscopic scale cannot be completely eliminated (a failure of *scale decoupling*). Degrees of freedom on all scales remain coupled, and universal correlations cannot be inferred directly by replacing the Boltzmann weight by a perturbed Gaussian measure.

We are reminded of Euclidean local field theories, like the ϕ^4 field theory, which require the introduction of a large momentum cut-off (an artificial short-distance structure), which are renormalizable, or super-renormalizable, for dimensions smaller than and equal to 4, and which also have large-distance universal properties (they are, to some extent, short-distance insensitive) [120]. These properties can be studied by field theory renormalization and renormalization group (RG) methods. The derivation of universal properties of phase transitions in macroscopic systems (taking as examples ferromagnetic systems) by the field theory RG methods is the main topic of Chapters 15–17.

In the appendix, we discuss the relation between mean-field expansion and low or high temperature expansions. We generalize the mean-field techniques to a large class of spin models. We describe the *replica trick*, useful to study quenched averages.

Terminology. To describe critical phenomena, it has become customary to use the language of magnetic systems. Although such a presentation certainly helps our physical intuition, many systems to which the theory applies are non-magnetic (see Sections 15.8–15.10). Therefore, this language is, in a sense, almost as abstract as the terminology of QFT. Since the universal properties for a large class of continuous phase transitions can eventually be determined by Euclidean QFT methods, we could immediately use such a formalism. However, since we also wish to introduce QFT methods to readers with a background in statistical mechanics, we often use statistical and magnetic terminology. For example, in chapters devoted to QFT, we call correlation functions what often is called imaginary time Green's functions or Schwinger's functions; we speak about partition function, free energy, and thermodynamic potential. More confusing, perhaps, in the framework of statistical physics one calls Hamiltonian what in QFT is the Euclidean (or imaginary-time) action (*c.f.* the statistical theory of liquids).

14.1 The transfer matrix

In Section 2.5, we have introduced the transfer matrix associated with a class of one-dimensional lattice models with nearest neighbour (n.n.) interactions, relating it to the quantum statistical operator $e^{-\varepsilon H}$ for small Euclidean time steps ε .

The partition function \mathcal{Z} , with periodic boundary conditions, for a lattice of size ℓ in the Euclidean time direction, can then be expressed in terms of the transfer matrix \mathbf{T} as

$$\mathcal{Z} = \text{tr } \mathbf{T}^\ell. \quad (14.1)$$

An elementary example is provided by the Ising model, at temperature $1/\beta$ and in a magnetic field h . The partition function,

$$\mathcal{Z} = \sum_{\{S_i=\pm 1\}} \exp \left[\beta \sum_{i=1}^l (JS_i S_{i+1} + h S_i) \right], \quad (14.2)$$

where we assume $J > 0$ (ferromagnetic interaction), can be expressed in terms of the transfer matrix which is a 2×2 matrix:

$$\langle S' | \mathbf{T} | S \rangle = \exp \left[\beta \left(JSS' + \frac{1}{2}h(S + S') \right) \right], \quad (14.3)$$

where the bra-ket notation of quantum mechanics is used in this chapter to represent matrix elements (not to be confused with the symbol $\langle \bullet \rangle$ meaning expectation value).

This idea can be generalized to lattice models with finite range interactions in generic d dimensions. It is then necessary to distinguish one direction on the lattice, which plays the role of Euclidean time in QFT.

In what follows, we consider always only isotropic interactions on a hypercubic lattice and, therefore, we can always choose the transfer matrix symmetric (this corresponds to Hermitian Hamiltonians in quantum mechanics).

The inverse temperature β plays different roles in quantum and classical statistical mechanics. The parameter β of quantum mechanics is the analogue of the size of a classical system in one additional dimension. In particular, the large β (*i.e.* zero temperature) limit of quantum mechanics corresponds to the large l limit.

Since fixed temperature corresponds to fixed size in one dimension, it will become apparent later that, from the point of view of phase transitions, quantum fluctuations are often irrelevant. One exception is provided by zero-temperature quantum statistical systems in $(d - 1)$ dimensions, which often share properties with classical statistical systems in d dimensions. Moreover, from the discussion of Section 2.5, we conclude that the quantum correlation functions, in the infinite β limit, are the analogues of the infinite volume statistical correlation functions.

In the thermodynamic limit, l goes to infinity and, therefore, the partition function is related to the largest eigenvalue of the transfer matrix. The corresponding eigenvector plays the role of the ground state of a quantum Hamiltonian.

14.1.1 A few properties of the transfer matrix on a finite transverse lattice

If the size of the lattice transverse to the time axis is finite, using the positivity of the coefficients of the transfer matrix, one can derive an important property of its spectrum (see also Section A34.1). We describe the arguments in the case in which the spin distribution is discrete, and the transfer matrix is a finite matrix but the arguments generalize to continuous spin distributions.

Since \mathbf{T} is real and symmetric, its eigenvalues and eigenvectors are real. Let λ_0 be the eigenvalue with largest modulus, and $|0\rangle$ the corresponding eigenvector,

$$\mathbf{T}|0\rangle = \lambda_0|0\rangle. \quad (14.4)$$

We denote by T_{ab} the elements of the transfer matrix \mathbf{T} , and v_a^0 the components of the eigenvector $|0\rangle$, and rewrite equation (14.4) in component form as

$$\lambda_0 \sum_b T_{ab} v_b^0 = v_a^0. \quad (14.5)$$

The scalar product of the equation with the vector \mathbf{v}^0 yields

$$\sum_a v_a^0 v_a^0 \lambda_0 = \sum_{a,b} v_a^0 T_{ab} v_b^0. \quad (14.6)$$

The coefficient of λ_0 is a sum of positive terms and the matrix elements of \mathbf{T} are positive. Taking the modulus of equation (14.6), we obtain the inequality

$$|\lambda_0| \sum_a |v_a^0| |v_a^0| \leq \sum_{a,b} |v_a^0| |T_{ab}| |v_b^0|. \quad (14.7)$$

The right-hand side of the inequality is the average of \mathbf{T} taken with the vector of components $|v_a^0|$. A strict inequality would imply the existence of at least one positive eigenvalue larger than $|\lambda_0|$, in contradiction with the hypothesis. Equality, combined with the property that \mathbf{T} has *strictly* positive coefficients, implies that all components of the vector $|0\rangle$ can be chosen as non-negative. Then, according to equation (14.5), λ_0 is positive and all components of $|0\rangle$ are strictly positive. Therefore, the eigenvalue cannot be degenerate, because two vectors with strictly positive coefficients cannot be orthogonal. To summarize: because the transfer matrix is real, symmetric and has strictly positive coefficients, the eigenvalue with largest modulus is positive, the corresponding eigenvector is unique and has strictly positive components.

In the large l (*i.e.* large time) limit, the free energy \mathcal{W} is given by

$$\mathcal{W} = \ln \mathcal{Z} \sim l \ln \lambda_0. \quad (14.8)$$

Note that, in classical systems, when discussing *finite temperature phase transitions*, we generally omit the temperature factor in front of the free energy.

Since no crossing of levels can occur, λ_0 is a regular function of the inverse temperature β . For large time separation between the points \mathbf{r} and \mathbf{r}' , the connected two-point spin correlation function behaves as

$$\langle S_{\mathbf{r}} S_{\mathbf{r}'} \rangle_c = \langle S_{\mathbf{r}} S_{\mathbf{r}'} \rangle - \langle S_{\mathbf{r}} \rangle \langle S_{\mathbf{r}'} \rangle \sim (\langle 0 | \mathbf{S} | 1 \rangle)^2 e^{-|\mathbf{r}-\mathbf{r}'|/\xi}, \quad (14.9)$$

in which $|1\rangle$ is the eigenvector corresponding to the second largest eigenvalue in modulus $\lambda_1 < \lambda_0$, which we have here assumed to be positive, and ξ , *the correlation length*, is given by

$$\xi^{-1} = \ln(\lambda_0/\lambda_1). \quad (14.10)$$

This analysis shows that, on a lattice with finite transverse size, in a spin model with finite range interactions and discrete distribution, no phase transition can occur; the free energy is a regular function of the inverse temperature β , and the correlation length ξ remains finite except for β infinite. These results can be generalized to short range interactions and continuous spin distributions.

14.2 The infinite transverse size limit: Ising-like systems

When the transverse size becomes infinite, new phenomena may appear, which we examine first on the example of the d -dimensional ferromagnetic ($J > 0$) Ising model with n.n. interactions. The partition function can be written as

$$\mathcal{Z} = \sum_{\{S_{\mathbf{r}} = \pm 1\}} \exp\left(\beta J \sum_{\mathbf{r}, \mathbf{r}', \text{n.n.}, \in \mathbb{Z}^d} S_{\mathbf{r}} S_{\mathbf{r}'}\right). \quad (14.11)$$

The elements of the transfer matrix now take the form

$$\langle \{S'_{\rho}\} | \mathbf{T} | \{S_{\rho}\} \rangle = \exp\left[\beta J \left(\sum_{\rho \in \mathbb{Z}^{d-1}} S_{\rho} S'_{\rho} + \frac{1}{2} \sum_{\rho, \rho', \text{n.n.}, \in \mathbb{Z}^{d-1}} (S_{\rho} S_{\rho'} + S'_{\rho} S'_{\rho'}) \right) \right]. \quad (14.12)$$

We introduce the reflection matrix that flips all the spins:

$$\mathbf{P} |\{S_{\rho}\}\rangle = | \{-S_{\rho}\}\rangle. \quad (14.13)$$

Due to the reflection symmetry of the Ising model, the matrix \mathbf{P} commutes with the transfer matrix and, therefore, \mathbf{P} and \mathbf{T} can be diagonalized simultaneously. Since $\mathbf{P}^2 = 1$, the eigenvalues of \mathbf{P} are ± 1 .

For L finite, the general results on matrices with positive elements apply. The eigenvector $|0\rangle$, corresponding to the non-degenerate largest eigenvalue of \mathbf{T} , has positive components. It is thus an eigenvector of \mathbf{P} . Equation (14.13) shows that \mathbf{P} does not change the sign of the basis vectors. Therefore, the eigenvector $|0\rangle$ is an eigenvector of \mathbf{P} with eigenvalue $+1$,

$$\mathbf{P} |0\rangle = |0\rangle. \quad (14.14)$$

We now examine the infinite L limit at low and high temperature, (high and low β). (A proof of spontaneous symmetry breaking can be found in Ref. [121].)

High temperature. At high temperature ($\beta \rightarrow 0$), all matrix elements of \mathbf{T} become equal and \mathbf{T} becomes a projector onto the eigenvector $|0\rangle$, which has equal components on all spin configurations. All eigenvalues but one vanish, thus the correlation length vanishes. This property is independent of the volume and, therefore, previous results apply even for L infinite.

Low temperature analysis. At low temperature, the dominant configurations correspond to all spins aligned. We call $|+\rangle$ and $|-\rangle$ the two vectors corresponding to all spin up and down, respectively. At β strictly infinite, both are eigenvectors of \mathbf{T} corresponding to the largest eigenvalue, which is thus twice degenerate. The result that the eigenvalue is degenerate does not contradict the general analysis of Section 14.1, because at zero temperature, if we normalize the largest matrix elements of \mathbf{T} to 1, all matrix elements but the diagonal elements $\langle - | \mathbf{T} | - \rangle$ and $\langle + | \mathbf{T} | + \rangle$ vanish, and the general arguments of Section 14.1 no longer apply. Note that $|+\rangle$ and $|-\rangle$ are not eigenvectors of \mathbf{P} since

$$\mathbf{P} |+\rangle = |-\rangle. \quad (14.15)$$

At low but finite temperature, the eigenvectors can no longer be exactly $|+\rangle$ and $|-\rangle$ but also necessarily have components on configurations in which a finite number of spins have been reversed, as low temperature perturbation theory shows. However, this does not forbid the existence of two degenerate eigenstates close to $|+\rangle$ and $|-\rangle$, and exchanged by \mathbf{P} . What is relevant is the large L behaviour of the matrix elements of \mathbf{T} connecting $|+\rangle$ and $|-\rangle$, more precisely the ratio

$$\delta = \frac{\langle + | \mathbf{T} | - \rangle}{\langle + | \mathbf{T} | + \rangle}. \quad (14.16)$$

At low temperature, at leading order, the ratio is related to the additional minimal energy corresponding to the configurations with the two different (twisted) boundary conditions, spins up on one side and spins down on the other side: it is proportional to the area of the surface on which the spins are reversed. Therefore, δ behaves like

$$\delta \propto e^{-\beta JL^{d-1}}. \quad (14.17)$$

The eigenvectors and eigenvalues are qualitatively given by diagonalizing a 2×2 matrix τ in the $\{|+\rangle, |-\rangle\}$ subspace:

$$\tau = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}. \quad (14.18)$$

Two different situations arise, depending on the dimension d of space.

(i) $d = 1$

Then, δ is finite and the eigenvector $|0\rangle$ is the linear combination

$$|0\rangle = |+\rangle + |-\rangle,$$

which is also an eigenvector of \mathbf{P} with the eigenvalue +1.

(ii) $d > 1$

In the infinite volume limit, δ vanishes, the largest eigenvalue remains twice degenerate. The finite size correlation length diverges as

$$\xi_L \propto e^{\beta JL^{d-1}}. \quad (14.19)$$

Clearly, in the infinite volume limit, no analytic continuation is possible between the high and low temperature situations and, therefore, thermodynamic quantities must have at least one singularity in β , at some finite value β_c [121]. We argue in Section A34.3 that, at low temperature, a breaking of ergodicity occurs, and thus β_c also corresponds to a phase transition in the dynamic sense.

Remarks.

(i) This heuristic analysis of the infinite volume limit is qualitatively correct in the whole low temperature phase. At β_c , the situation is different; an infinite number of eigenvalues have the same infinite volume limit λ_0 . This situation is studied in detail in Chapter 32.

(ii) We have analysed here a lattice model, but we show in Section 39.1 that, in the case of models defined in continuum space, instantons lift the degeneracy of the ground state in the semi-classical limit. With the correspondence,

$$\beta \mapsto \hbar^{-1}, \quad J \mapsto \text{instanton action},$$

the analysis is then exactly the same.

(iii) From the preceding analysis, we derive a criterion of spontaneous symmetry breaking (SSB). We consider the ratio

$$r = \lim_{l \rightarrow \infty} \frac{\text{tr } \mathbf{P} \mathbf{T}^l}{\text{tr } \mathbf{T}^l}, \quad (14.20)$$

in which \mathbf{P} , more generally, is an element of the symmetry group of the model.

In the symmetric phase, the ground state of the transfer matrix is invariant under a group operation, therefore, the ratio r is 1.

By contrast, if the symmetry is spontaneously broken, \mathbf{P} exchanges the various ground states and, therefore, r vanishes in the infinite volume limit. In the example of the Ising model, one finds at low temperature,

$$r \propto e^{-\beta J L^{d-1}}.$$

The criterion involving the ratio (14.20) has one advantage: by specializing to $l = L$, the linear size of the lattice, it can be expressed in terms of the ratio of two partition functions on a d -dimensional lattice of linear size L , with different boundary conditions, the denominator corresponds to periodic boundary conditions in the time direction, the numerator corresponds to twisted boundary conditions, that is, the configurations on the two boundaries differ by a transformation of the symmetry group. In the case of the Ising model, twisted boundary conditions are anti-periodic boundary conditions. For $l = L$, this ratio naturally incorporates the condition that the thermodynamic limit is taken by sending the sizes in all dimensions to infinity in the same way. It represents, as we show in Section A34.3, the probability in some dynamics that the system can evolve from initial conditions in which almost all spins are up to a configuration in which most of the spins are reversed.

14.2.1 Order parameter and cluster properties

When the ground state $|0\rangle$ (vacuum state in the sense of particle physics) is degenerate, the determination of the infinite volume correlation functions becomes a subtle question, which depends explicitly on the way the infinite volume limit is reached; in particular, it may depend on boundary conditions. The sensitivity to boundary conditions is another characteristic of the several phase region.

We now examine the cluster properties of correlation functions in such a situation, for Ising-like systems.

We assume that the spontaneous magnetization is $\pm m \neq 0$, depending on the phase. The first question that arises concerns the space of configurations. We first assume that we calculate the partition function by summing over all configurations. Then, the correlation functions are symmetric. Therefore, $\langle \mathbf{S} \rangle = 0$ but this results from the sum $m - m$. The large distance behaviour of two-point function takes the form

$$\lim_{|x-y| \rightarrow \infty} \langle \mathbf{S}(x) \mathbf{S}(y) \rangle = m^2.$$

Then, the connected two-point function,

$$\lim_{|x-y| \rightarrow \infty} \langle (\mathbf{S}(x) - \langle \mathbf{S} \rangle)(\mathbf{S}(y) - \langle \mathbf{S} \rangle) \rangle = m^2,$$

does not satisfy cluster properties. In the same way, even though the ground state of the transfer matrix is degenerate and, thus, any linear combination of two different ground states is a ground state, only the ground states corresponding to the maximal magnetizations $|m|$ are consistent with cluster properties.

The pure phases can be selected by the following procedure: in a spin system, at finite volume, one adds a constant magnetic field, which breaks the symmetry. One takes the infinite volume limit, and then the zero field limit.

As we have already emphasized, from the viewpoint of dynamics, the phase transition corresponds to a breaking of ergodicity. If a system is prepared in one phase, it remains in this phase at all later times.

Therefore, in the two phase region of a spin system, expectation values should be calculated by averaging only over configurations which fluctuate either around the configurations with all spins up, or the configurations with all spins down.

14.3 Continuous symmetries

We now briefly discuss a model which has a continuous symmetry to stress one important difference with the case of discrete symmetries. We consider a classical spin system, where the spins \mathbf{S}_r are N -component vectors of unit length, interacting through a pair n.n. ferromagnetic interaction, with a continuous $O(N)$ symmetry. The partition function on a d -dimensional hypercubic lattice can be written as

$$\mathcal{Z} = \int \prod_{\mathbf{r} \in \mathbb{Z}^d} d\mathbf{S}_{\mathbf{r}} \delta(\mathbf{S}_{\mathbf{r}}^2 - 1) \exp \left[\sum_{\mathbf{r}, \mathbf{r}', \text{n.n.}, \in \mathbb{Z}^d} \beta \mathbf{S}_{\mathbf{r}} \cdot \mathbf{S}_{\mathbf{r}'} \right]. \quad (14.21)$$

Such a model can be considered as a regularization of the $O(N)$ non-linear σ -model (Chapter 19).

At high temperature ($\beta \rightarrow 0$), as in the case of the Ising model, the ground state of the transfer matrix has uniform components on all configuration vectors and is, therefore, invariant under the transformations of the $O(N)$ symmetry group.

To understand the structure of low temperature phases, we now use the generalization of the ratio (14.20). The symmetry operation here is a rotation $R(\alpha)$ of angle α . We thus define a partition function $\mathcal{Z}_L(\alpha)$ with twisted boundary conditions in the Euclidean time direction:

$$\mathbf{S}_{t=L, \rho} \cdot \mathbf{S}_{t=0, \rho} = \cos \alpha, \quad \rho \in \mathbb{Z}^{d-1}. \quad (14.22)$$

For convenience, we choose periodic boundary conditions in the remaining dimensions.

We then examine the behaviour of the ratio

$$r_L = \frac{\mathcal{Z}_L(\alpha)}{\mathcal{Z}_L(0)}, \quad (14.23)$$

for β large, in the large L limit.

At low temperature, the configurations with minimal energy correspond to take the spins aligned in $(d-1)$ dimensions and rotating by an angle α/L between two adjacent sites along the time axis. This has to be contrasted with the case of the discrete symmetry, in which the transition between the two configurations imposed by the boundary conditions occurs between two sites.

The additional energy ΔE due to the rotation is thus

$$\Delta E = L^{d-1} \times L \times [\cos(\alpha/L) - 1] \sim -\alpha^2 L^{d-2}/2 \quad (14.24)$$

and, therefore,

$$r_L \propto \exp(-\alpha^2 \beta L^{d-2}/2). \quad (14.25)$$

In the case of a continuous symmetry, it is easier to pass from one minimum of the energy to another. This property has the direct consequence that it is more difficult to break the symmetry, and that Goldstone modes appear in the phase of SSB. For $d \leq 2$, r_L has a finite limit and the symmetry is never broken (although, as we shall see, a phase transition without ordering is possible for $d = 2$, $N = 2$). This result, for which we have given a heuristic argument, can be proven rigorously (Mermin–Wagner–Coleman theorem [118]).

For $d > 2$, SSB occurs at low temperature. There exists some finite temperature T_c , at which a phase transition occurs.

Note that, if we consider a system with a longitudinal size $l \neq L$, by contrast, we obtain

$$r \propto \exp(-\alpha^2 \beta L^{d-1}/2l). \quad (14.26)$$

Hence, using the results of Sections 3.4 and 3.5, we conclude that the finite size correlation length ξ_L behaves like

$$\xi_L \propto \beta L^{d-1}. \quad (14.27)$$

To relate this result with the possibility of a phase transition, we note that if we assume that the correlation length is large but much smaller than the transverse size of the lattice then we have the estimate

$$\langle S \rangle^2 \propto L^{-2d} \sum_{\mathbf{r}, \mathbf{r}'} \langle \mathbf{S}_\mathbf{r} \cdot \mathbf{S}_{\mathbf{r}'} \rangle \propto L^{-2d} \sum_{\mathbf{r}, \mathbf{r}'} \exp(-|\mathbf{r} - \mathbf{r}'|/\xi_L)$$

and, therefore,

$$\langle S \rangle \propto (\xi_L/L)^d. \quad (14.28)$$

The assumption implies that $\langle S \rangle$ goes to 0 in the large volume limit. A phase transition with ordering is only possible if

$$\xi_L \underset{L \rightarrow \infty}{\geq} L \Rightarrow d \geq 2.$$

The limiting case $\xi_L \propto L$, that is, $d = 2$, is quite subtle and we shall show eventually that it characterizes the critical temperature of a second order phase transition with vanishing spontaneous magnetization in the low temperature phase.

Concluding remarks. The high and low temperature analysis provides information about the existence of a phase transition, and the nature of the phases. However, nothing can be inferred about the behaviour of thermodynamic quantities in the neighbourhood of the critical temperature. These problems are examined in the coming sections, within the mean-field approximation, and in Chapters 15 and 16 using RG methods.

14.4 Mean-field approximation

In the first part of the chapter, we have discussed the existence of phase transitions in ferromagnetic systems by comparing the phase structure at low and high temperatures. However, we have used methods which provide no information about the behaviour of thermodynamic quantities at the transition itself. For this purpose, the simple mean-field approximation [116], can be used, which makes it possible to study the neighbourhood of the transition temperature. We distinguish between first and second order phase transitions. In the latter case, the correlation length diverges at the critical (transition) temperature. This has important physical consequences: in the mean-field approximation, the singular behaviour of thermodynamic quantities at the critical temperature, and, more generally, in the neighbourhood of the critical temperature and in small magnetic field, is *universal*, that is, does not depend on the details of the interactions, on the dimension of space and, in the high-temperature phase at least, on the symmetry of the models.

The mean-field approximation can be derived by various methods, as a partial summation of high temperature series (see Section A14.2), as the result of variational calculations, or as the leading order in a calculation by the steepest descent method.

We adopt the latter viewpoint, because it makes a systematic calculation of corrections to the mean-field approximation possible and, therefore, a discussion of its validity. The role of space dimension 4 then emerges.

Moreover, in this framework, it becomes apparent that the mean-field approximation is a Gaussian or weakly perturbed, *quasi-Gaussian*, approximation.

14.4.1 Ising-like ferromagnetic systems

As an example, we again consider a classical spin system on a d -dimensional lattice of points with integer coordinates. The spin variable on a lattice site i of coordinates $(n_1^i, \dots, n_d^i) \in \mathbb{Z}^d$ is denoted by S_i . The Hamiltonian \mathcal{H} (the energy of a spin configuration) is the sum of a ferromagnetic, translation invariant, short range, two-spin interaction and a site-dependent magnetic field term,

$$-\beta\mathcal{H}(S) = \sum_{i,j} V_{ij} S_i S_j + \sum_i H_i S_i, \quad (14.29)$$

where the inverse temperature $\beta = 1/T$ has been absorbed into the potential V_{ij} and into the magnetic field H_i .

In this way, the partition function,

$$\mathcal{Z}(H) = \int \left(\prod_i d\rho(S_i) \right) \exp [-\beta \mathcal{H}(S)], \quad (14.30)$$

where $d\rho(S)$ is the normalized distribution of the spin configurations at each site, is also the generating functional of spin correlation functions.

We first discuss systems which, in zero magnetic field, have an Ising-like $S \mapsto -S$ symmetry. We assume that the potential V_{ij} has the hypercubic symmetry of the lattice, for simplicity. Expressions similar to expression (14.30) also appear in the construction of the Feynman–Kac representation of partition function in quantum mechanics (*e.g.* in Sections 2.2–2.4) or, in QFT, as lattice regularizations of the generating functional of Euclidean correlation functions (Section 8.7). However, as we show in Section A14.1, on the lattice it is possible to calculate the partition function by a high temperature expansion: the site measure $d\rho(S)$ is kept fixed and one expands in powers of $\sum_{i,j} V_{ij} S_i S_j$, the term which connects different lattice sites. By contrast, in the perturbative expansion of field theory, the term $\sum_{i,j} V_{ij} S_i S_j$ is the analogue of the kinetic term while the measure $d\rho(S)$ contains the interactions. These remarks will be clarified later.

The pair potential. The potential $V_{ij} = V_{ji}$ is assumed to be translation invariant, $V_{ij} \equiv V(\mathbf{n}^{ij})$, where \mathbf{n}^{ij} is the vector joining the sites i and j . It is then natural to introduce its Fourier transform,

$$\tilde{V}(\mathbf{k}) = \sum_{\mathbf{n}} V(\mathbf{n}) \exp(i\mathbf{k} \cdot \mathbf{n}), \quad (14.31)$$

where the components k_μ of the momentum \mathbf{k} can be restricted to the Brillouin zone $|k_\mu| \leq \pi$. We now define more precisely what we mean by short range potential: the potential decreases at large distance fast enough for its Fourier transform $\tilde{V}(\mathbf{k})$ to have a convergent Taylor series expansion at $\mathbf{k} = 0$.

We choose also V_{ij} non-negative to ensure that the two-spin interaction is ferromagnetic. The positivity of the pair potential implies

$$|\tilde{V}(\mathbf{k})| \leq \tilde{V}(0) = \sum_{\mathbf{n}} V(\mathbf{n}) \equiv v, \quad (14.32)$$

where the parameter v is proportional to β , the inverse temperature.

Then, the expansion of $\tilde{V}(\mathbf{k})$ for \mathbf{k} small takes the form

$$\tilde{V}(\mathbf{k}) = v\tilde{U}(\mathbf{k}), \quad \tilde{U}(\mathbf{k}) = 1 - a^2\mathbf{k}^2 + O(k^4) \quad (14.33)$$

where k is the length of the vector \mathbf{k} , a a numerical constant and the hypercubic symmetry of the potential has been used. Note that $\tilde{V}(\mathbf{k})$ has up to order k^2 an *emergent* $O(d)$ symmetry, consequence of the hypercubic symmetry.

The one-site problem. The solution of the elementary one-site problem, gives us the opportunity to introduce a few useful functions, and to analyse their properties.

The Laplace transform of the spin distribution

$$z(h) = \int d\rho(S) e^{Sh} = \int d\rho(S) \cosh(Sh), \quad (14.34)$$

is the one-site partition function. The corresponding free energy $w(h)$ is then (note that we omit a factor β in the definition of the free energy, which for finite temperature transitions plays no role)

$$w(h) = \ln z(h). \quad (14.35)$$

For the Ising model with $S_i = \pm 1$, $w(h) = \ln \cosh h$.

The even function $w(h)$ is increasing for $h > 0$. Moreover, $w''(h)$ is positive (Schwartz's inequality), and thus $w(h)$ is convex. Finally, we assume that the spin distribution decreases faster than a Gaussian for large values of the spin S . It is then easy to verify that $z(h)$ is an entire function, and this implies that $w(h)$ is regular on the real axis, and in some neighbourhood of the origin. Moreover,

$$\lim_{|h| \rightarrow \infty} w(h)/h^2 = 0. \quad (14.36)$$

The magnetization, expectation value of the spin, is

$$m = \langle S \rangle = w'(h).$$

The thermodynamic potential $g(m)$, *Legendre transform* of the free energy $w(h)$ (see Section 1.8), is defined by

$$g(m) + w(h) = hm, \quad m = w'(h). \quad (14.37)$$

In the example of the Ising model, one finds

$$g(m) = \frac{1}{2}(1+m)\ln(1+m) + \frac{1}{2}(1-m)\ln(1-m). \quad (14.38)$$

The Legendre transformation implies the relation

$$g''(m) = 1/w''(h). \quad (14.39)$$

Since $w(h)$ is a convex function, $g(m)$ is also convex. It is analytic for m small, and can thus be parametrized as

$$g(m) = \sum_{p=1} \frac{g_{2p}}{2p!} m^{2p}, \quad g_2 > 0. \quad (14.40)$$

In the Ising model, one finds

$$g_2 = 1, \quad g_4 = 2.$$

14.4.2 Mean-field approximation and steepest descent method

We now set up an algebraic formalism that makes it possible to obtain mean-field results as a leading order approximation, and to systematically calculate corrections to the mean-field approximation. It is technically convenient here to assume that V_{ii} (which is independent of the site i) vanishes (see Section A14.2).

Since the partition function can easily be calculated when all points are disconnected, a simple idea is to write the term which in the configuration energy connects spins on different sites as an integral over disconnected terms. More explicitly, for each site we insert the δ -function,

$$\delta(\sigma_i - S_i) = \frac{1}{2i\pi} \int d\lambda_i e^{\lambda_i(S_i - \sigma_i)},$$

where the integration over λ_i runs along the imaginary axis, into the representation of the partition function:

$$\mathcal{Z}(H) \propto \int \prod_i d\rho(S_i) (d\sigma_i d\lambda_i) \exp \left[-\beta \mathcal{H}(\sigma) + \sum_i \lambda_i (S_i - \sigma_i) \right]. \quad (14.41)$$

Integration over the spin variables S_i is then immediate. Introducing the function $w(H)$ defined in equation (14.35), one finds

$$\mathcal{Z}(H) = \int \prod_i (d\sigma_i d\lambda_i) \exp \left[-\beta \mathcal{H}(\sigma) + \sum_i (-\lambda_i \sigma_i + w(\lambda_i)) \right]. \quad (14.42)$$

We then evaluate the expression by the steepest descent method [51]. The saddle points are the solutions of the equations obtained by differentiating the integrand with respect to λ_i and σ_i :

$$\sigma_i = w'(\lambda_i), \quad (14.43a)$$

$$\lambda_i = 2 \sum_j V_{ij} \sigma_j + H_i. \quad (14.43b)$$

At leading order, we replace σ_i, λ_i in expression (14.42) by the solution of equations (14.43). The mean-field free energy is then

$$\mathcal{W}(H) \equiv \ln \mathcal{Z}(H) = \sum_{i,j} V_{ij} \sigma_i \sigma_j + \sum_i [H_i \sigma_i - \lambda_i \sigma_i + w(\lambda_i)] \quad (14.44)$$

$$= \sum_{i,j} V_{ij} \sigma_i \sigma_j + \sum_i [H_i \sigma_i - g(\sigma_i)]. \quad (14.45)$$

Since the expression (14.44) is stationary with respect to variations of σ_i and λ_i , the local magnetization M_i , expectation value of the spin S_i , in the mean-field approximation is given by the explicit derivative with respect to H_i :

$$M_i = \langle S_i \rangle = \frac{\partial \mathcal{W}}{\partial H_i} = \sigma_i. \quad (14.46)$$

The mean field. The saddle point value of λ_i has an interpretation as an effective magnetic field, determined by equation (14.43b) as the sum of the applied external field and a mean field representing the action of the other spins. The interaction that connects the sites has been replaced by a mean magnetic field. A more detailed analysis shows that this approximation becomes exact when the dimension d of space becomes large, so that the action of all neighbouring sites on a given site can indeed be replaced by a mean magnetic field (a property reminiscent of the central limit theorem of probability).

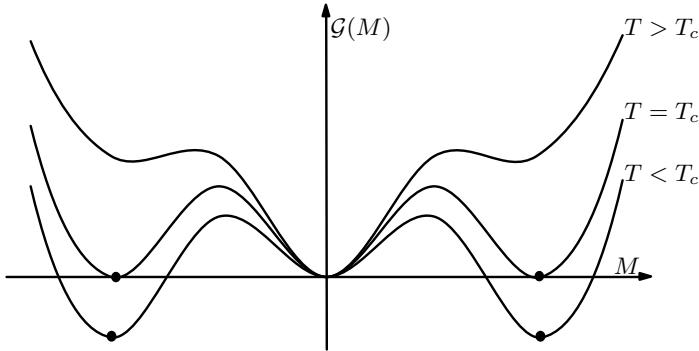


Fig. 14.1 Free energy: first order phase transition

14.4.3 Thermodynamic potential and phase transition

The thermodynamic potential $\Gamma(M)$, function of the local magnetization, is the Legendre transform of $\mathcal{W}(H)$:

$$\Gamma(M) = \sum_i M_i H_i - \mathcal{W}(H), \quad M_i = \frac{\partial \mathcal{W}(H)}{\partial H_i}. \quad (14.47)$$

From equations (14.44) and (14.46), one infers

$$\Gamma(M) = - \sum_{i,j} M_i V_{ij} M_j + \sum_i g(M_i), \quad (14.48)$$

in which $g(M)$ is the function (14.37).

In the case of translation invariant systems (and thus in a uniform magnetic field), the magnetization is uniform: $M_i = M$. Therefore,

$$\mathcal{G}(M) = \lim_{\Omega \rightarrow \infty} \Omega^{-1} \Gamma(M) = -vM^2 + g(M), \quad (14.49)$$

where v is the parameter (14.32) ($v \propto 1/T$) and Ω the volume of the lattice.

The relation between magnetic field H , magnetization M and temperature $T \propto 1/v$, called *equation of state*, becomes

$$H = \frac{\partial \mathcal{G}}{\partial M} = -2vM + g'(M). \quad (14.50)$$

In zero magnetic field, the magnetization is given by an extremum of the thermodynamic potential density $\mathcal{G}(M)$. Furthermore, since the partition function in zero field is $\exp[-\Omega \mathcal{G}(M)]$, the leading saddle points correspond to minima of $\mathcal{G}(M)$.

We now look for the minima of $\mathcal{G}(M)$ when the temperature and, as a consequence, v vary. The property (14.36) implies that, for $|M|$ large enough, $\mathcal{G}(M)$ in equation (14.49) is an increasing function. For v small (high temperature), vM^2 is negligible, and the right-hand side of equation (14.49) is convex. The minimum of $\mathcal{G}(M)$ is at $M = 0$, and the magnetization vanishes. When v increases, in general, at some value of v other local minima appear, which eventually become the absolute minima of $\mathcal{G}(M)$. When this occurs, the value of the magnetization M jumps discontinuously from 0 to a finite value corresponding to this new absolute minima. The system undergoes a *first order phase transition*. Fig. 14.1 describes the evolution.

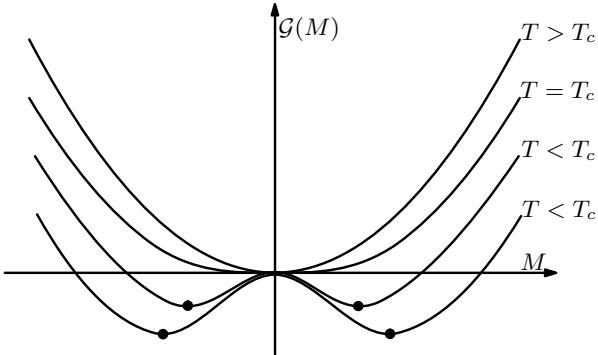


Fig. 14.2 Free energy: second order phase transition

Fluctuations around the saddle point are governed by the value of the second derivative of the potential at the minimum. Generically, in such a case, the second derivative is strictly positive and, therefore, the correlation length, which, as we know from QFT calculations and shall again see explicitly later, is proportional to $\Omega^{1/2} [\mathcal{G}''(M)]^{-1/2}$ at the minimum, remains finite.

Although first order phase transitions are common, they are not particularly interesting for us. When the correlation length remains finite, no universality emerges. Moreover, because the fluctuations are not critical, the mean-field approximation gives a satisfactory qualitative description of the physics.

However, if no absolute minimum appears at a finite distance from the origin, finally, at a critical value $v = v_c$, with

$$v_c = \frac{1}{2}g''(0), \quad (14.51)$$

corresponding to a critical temperature $T = T_c$, the origin ceases being a minimum of the potential and, below this temperature, two minima move *continuously* away from the origin (see Fig. 14.2). Since the magnetization remains continuous at v_c , the *phase transition is of second order*. Because at v_c the second derivative of $\mathcal{G}(M)$ vanishes, the *correlation length diverges*.

This is the situation we now analyse systematically, first in the framework of the mean-field approximation and then by calculating corrections coming from higher orders in the mean-field expansion.

14.5 Universality within mean-field approximation

We now examine the behaviour of various thermodynamic quantities for a temperature T close to T_c , that is, v close to v_c . Since the transition is continuous, the magnetization is small and we can expand $\Gamma(M)$, which, as $g(M)$, is a smooth even function, in a Taylor series in M . We use the parametrization (14.37):

$$\Gamma(M_i) = - \sum_{i,j} V_{ij} M_i M_j + \sum_i \left(\frac{g_2}{2!} M_i^2 + \frac{g_4}{4!} M_i^4 + \dots \right). \quad (14.52)$$

The convexity of $g(M)$ implies that g_2 is positive. The parameter g_4 is also generically positive, because we have assumed that no first order transition occurs at higher temperature. In the mean-field approximation, the critical value is then $v_c = g_2/2$.

14.5.1 Homogeneous observables

With the parametrization (14.40), the relation (14.50) between applied field, magnetization and temperature becomes

$$H = 2(v_c - v)M + \frac{1}{6}g_4 M^3 + O(M^5). \quad (14.53)$$

For $H = 0$ and $v > v_c$ (*i.e.*, $T < T_c$), a spontaneous magnetization M is predicted, which near v_c behaves like,

$$M \sim [12(v - v_c)/g_4]^{1/2}, \quad \text{for } v - v_c \ll 1. \quad (14.54)$$

Since $T \propto 1/v$, for $v - v_c \ll 1$, $v - v_c \propto T_c - T$. Quite generally, one defines,

$$M \propto (T_c - T)^\beta, \quad (14.55)$$

where β is the magnetic exponent. Here, one finds the mean-field or ‘classical’ value

$$\beta = 1/2. \quad (14.56)$$

The inverse of the magnetic susceptibility $\chi = \partial M / \partial H$ is the second derivative of $\mathcal{G}(M)$ with respect to M . In zero field,

$$\begin{aligned} \chi_+^{-1} &= 2(v_c - v), & T > T_c, \\ \chi_-^{-1} &= 4(v - v_c), & T < T_c. \end{aligned} \quad (14.57)$$

Therefore, the magnetic susceptibility diverges at T_c with susceptibility exponents γ, γ' :

$$\begin{aligned} \chi_+ &\sim C_+(T - T_c)^{-\gamma}, & \gamma = 1, \\ \chi_- &\sim C_-(T_c - T)^{-\gamma'}, & \gamma' = 1, \end{aligned} \quad (14.58)$$

and

$$C_+ / C_- = 2. \quad (14.59)$$

At T_c , for a small uniform applied magnetic field H , the equation of state (14.53) becomes

$$H \sim \frac{1}{6}g_4 M^3. \quad (14.60)$$

In general, one defines $H \propto M^\delta$, and thus the exponent δ has the mean-field value

$$\delta = 3. \quad (14.61)$$

More generally, for v close to v_c , M and H small, the equation of state (14.53) can be cast into the *universal scaling form* proposed by Widom [122]:

$$H = h_0 M^\delta f((T - T_c)M^{-1/\beta}), \quad (14.62)$$

where h_0 is a positive normalization constant. With a suitable normalization of the temperature, the function f then takes the simple form

$$f(x) = 1 + x. \quad (14.63)$$

Specific heat. In zero field, the derivative of the free energy per unit volume with respect to v , which is a measure of the temperature, is proportional to the average energy:

$$\frac{1}{\Omega} \left. \frac{\partial \mathcal{W}(H)}{\partial v} \right|_{H=0} = M^2(H=0),$$

where the stationarity of W with respect to M and λ has been used. Above T_c it vanishes, and, below T_c it is proportional to the square of the spontaneous magnetization. Deriving again with respect to v , we obtain a quantity proportional to the specific heat \mathcal{C} :

$$\mathcal{C}(T \rightarrow T_{c+}) = 0, \quad \mathcal{C}(T \rightarrow T_{c-}) = 12/g_4. \quad (14.64)$$

In the mean-field approximation, the specific heat has a non-universal jump at T_c .

14.5.2 The two-point correlation function

More generally, the relation between applied magnetic field and local magnetization can be written as

$$H_i = \frac{\partial \Gamma}{\partial M_i} = -2 \sum_j V_{ij} M_j + g'(M_i). \quad (14.65)$$

A differentiation with respect to M_i and an inversion in the matrix sense, then yields the two-point connected correlation function,

$$W_{ij}^{(2)} \equiv W^{(2)}(\mathbf{n}^{ij}) = \left. \left(\frac{\partial^2 \Gamma}{\partial M_i \partial M_j} \right)^{-1} \right|_{M_i=M} = [-2V_{ij} + g''(M)\delta_{ij}]^{-1},$$

whose Fourier transform $\tilde{W}^{(2)}(\mathbf{k})$ is

$$\tilde{W}^{(2)}(\mathbf{k}) = \sum_{\mathbf{n}} W^{(2)}(\mathbf{n}) \exp(i\mathbf{k} \cdot \mathbf{n}) = [g''(M) - 2\tilde{V}(\mathbf{k})]^{-1}. \quad (14.66)$$

($\tilde{V}(\mathbf{k})$ is defined by equation (14.31)).

In general, one assumes that, at T_c , the two-point function behaves like

$$W^{(2)}(\mathbf{n}) \underset{|\mathbf{n}| \rightarrow \infty}{\propto} 1/|\mathbf{n}|^{d-2+\eta} \Leftrightarrow \tilde{W}^{(2)}(\mathbf{k}) \underset{k \rightarrow 0}{\propto} 1/k^{2-\eta}. \quad (14.67)$$

Inserting the expansion (14.33) into expression (14.66), one finds at $v = v_c$, in zero field, for k small,

$$\tilde{W}^{(2)}(\mathbf{k}) \propto 1/k^2, \quad (14.68)$$

result which yields the classical or mean-field value of the exponent η :

$$\eta = 0. \quad (14.69)$$

More generally for $k^2 = O(v_c - v)$ small, $\tilde{W}^{(2)}(\mathbf{k})$ can be expanded as:

$$\tilde{W}^{(2)}(\mathbf{k}) = \tilde{W}^{(2)}(0) [1 + k^2 \xi^2 + O(k^4/(v - v_c))]^{-1}. \quad (14.70)$$

This shows that the two-point function has an Ornstein and Zernike or free scalar field form (with mass $1/\xi$). We have parametrized $\tilde{W}^{(2)}(\mathbf{k})$ by introducing ξ^2 , which is proportional to the second moment of $W^{(2)}(\mathbf{n})$. The length ξ characterizes, up to a numerical factor, the exponential decay of the correlation functions, and can be taken as a measure of the correlation length.

The expansions (14.33) and thus (14.70) are, up to order k^2 , $O(d)$ symmetric, and thus have an *emergent symmetry* larger than the symmetry of the complete function \tilde{V} .

For non-vanishing magnetization with $M^2 = O(v - v_c)$, from equations (14.66) and (14.70), and with the parametrization (14.40), one then derives

$$\xi^{-2} = \frac{1}{2v_c a^2} (g_2 + \frac{1}{2} g_4 M^2 - 2v). \quad (14.71)$$

In zero magnetic field, this implies

$$\begin{aligned} \xi_+^{-2} &= a^{-2} (1 - v/v_c), & \text{for } T > T_c, \\ \xi_-^{-2} &= 2a^{-2} (v/v_c - 1), & \text{for } T < T_c. \end{aligned} \quad (14.72)$$

Introducing general correlation length exponents ν, ν' defined by

$$\xi_+ = f_+(T - T_c)^{-\nu}, \quad \xi_- = f_-(T_c - T)^{-\nu'}, \quad (14.73)$$

one infers from the relations (14.72) the mean-field values of the exponents

$$\nu = \nu' = 1/2,$$

and the *universal ratio of amplitudes*,

$$f_+ / f_- = \sqrt{2}. \quad (14.74)$$

14.5.3 Continuous symmetries

If the initial spin variable \mathbf{S}_i is an N -component vector, and if both the interaction and the spin distribution have some continuous symmetry, most of the previous results clearly remain unchanged. The main difference comes from the appearance of several type of correlation functions when the magnetization does not vanish. Moreover, in zero field for any temperature below T_c , Goldstone (massless) modes are expected (Section 13.4). We illustrate these properties with the example of the $O(N)$ symmetry.

MFT. The mean-field thermodynamic potential now takes the form

$$\Gamma(\mathbf{M}) = - \sum_{i,j} V_{ij} \mathbf{M}_i \cdot \mathbf{M}_j + \sum_i g(\mathbf{M}_i), \quad (14.75)$$

where $g(\mathbf{M})$ has a small M expansion of the form,

$$g(\mathbf{M}) = \frac{1}{2} g_2 \mathbf{M}^2 + \frac{1}{24} g_4 \mathbf{M}^4 + O(M^6),$$

with $g_2 > 0$ and $g_4 > 0$. It is convenient to set

$$g(\mathbf{M}) = G(\mathbf{M}^2).$$

Differentiating with respect to $M_{\alpha,i}$, where $M_{\alpha,i}$ ($\alpha = 1, \dots, N$) are the components of the magnetization vector \mathbf{M}_i , one finds

$$\frac{\partial \Gamma(M)}{\partial M_{\alpha,i}} = -2 \sum_j V_{ij} M_{\alpha,j} + 2 M_{\alpha,i} G'(\mathbf{M}^2). \quad (14.76)$$

For uniform field and magnetization, this yields

$$H = M(-2v + G'(M^2)) \sim (-2v + g_2)M + \frac{1}{6}g_4 M^3, \quad (14.77)$$

in which H, M now are the lengths of the vectors \mathbf{M}, \mathbf{H} .

From equation (14.75), we derive the two-point vertex or inverse correlation function:

$$\begin{aligned} \Gamma_{\alpha\beta,ij}^{(2)} &= \left. \frac{\partial^2 \Gamma(M)}{\partial M_{\alpha,i} \partial M_{\beta,j}} \right|_{\mathbf{M}_i=\mathbf{M}} \\ &= (-2V_{ij} + 2\delta_{ij}G'(M^2)) \delta_{\alpha\beta} + 4\delta_{ij}M_\alpha M_\beta G''(M^2), \end{aligned} \quad (14.78)$$

in which M_α ($\alpha = 1, \dots, N$) are the components of the uniform magnetization vector.

We denote by \mathbf{u} the unit vector along the direction of spontaneous magnetization:

$$\mathbf{M} = M\mathbf{u}_\alpha.$$

The two-point function (14.78) can be decomposed into transverse and longitudinal parts with respect to \mathbf{u} and written as

$$\Gamma_{ij,\alpha\beta}^{(2)} = u_\alpha u_\beta \Gamma_{L,ij}^{(2)} + (\delta_{\alpha\beta} - u_\alpha u_\beta) \Gamma_{T,ij}^{(2)}. \quad (14.79)$$

In the Fourier representation, the two components are given by

$$\tilde{\Gamma}_L^{(2)}(\mathbf{k}) = 2G'(M^2) + 4M^2G''(M^2) - 2\tilde{V}(\mathbf{k}), \quad (14.80a)$$

$$\tilde{\Gamma}_T^{(2)}(\mathbf{k}) = 2G'(M^2) - 2\tilde{V}(\mathbf{k}). \quad (14.80b)$$

The expressions (14.80a) and (14.66) are similar. Using equation (14.77), one can rewrite $\tilde{\Gamma}_T^{(2)}(\mathbf{k})$ as

$$\tilde{\Gamma}_T^{(2)}(\mathbf{k}) = H/M + 2\left(\tilde{V}(0) - \tilde{V}(\mathbf{k})\right). \quad (14.81)$$

This equation shows that, in zero field, the inverse transverse two-point function vanishes like k^2 at zero momentum for any temperature below T_c , indicating the presence of $(N-1)$ Goldstone (massless) modes.

Goldstone modes: generalization. For a uniform magnetization, the thermodynamic potential density $\mathcal{G}(M)$ (defined in (14.49)) inferred from equation (14.75) has the most general form consistent with $O(N)$ symmetry since

$$\mathcal{G}(M) = -v\mathbf{M}^2 + G(\mathbf{M}^2).$$

Therefore, the preceding results restricted to zero momentum are not specific to the mean-field approximation. In the ordered phase, the transverse correlation length diverges and the Goldstone phenomenon is completely general, in agreement with the results proven in Sections 13.3 and 13.4 (to which we refer for an extensive discussion of the issue).

14.5.4 Landau's theory of critical phenomena

Although all preceding results have been established in the framework of a specific approximation scheme, the mean-field approximation, they also follow, as shown by Landau [117], from more general assumptions that can be formalized as follows:

- (i) The thermodynamic potential can be expanded in powers of M around $M = 0$.
- (ii) The coefficients of the expansion are regular functions of the temperature for T close to T_c , the temperature at which the coefficient of M^2 vanishes, and of external thermodynamic parameters.
- (iii) The functions $\tilde{\Gamma}^{(n)}$, which appear in the expansion of $\Gamma(M)$ in powers of $\tilde{M}(k)$, the Fourier transform of the local magnetization M_i ,

$$\Gamma(M) = \sum_n \frac{1}{n!} \int d^d k_1 \cdots d^d k_n \delta^{(d)} \left(\sum k_i \right) \tilde{M}(k_1) \cdots \tilde{M}(k_n) \tilde{\Gamma}^{(n)}(k_1, \dots, k_n), \quad (14.82)$$

(we assume translation invariance) are regular for real values of k_i . Finally, second order transition implies $\tilde{\Gamma}^{(4)}(0, 0, 0, 0) > 0$.

Note that the thermodynamic potential of Landau's theory has the properties of the effective action of Section 8.8.1. In continuum space, it has a local expansion in terms of powers of $M(x)$ and its derivatives. It will eventually become clear that the discussion that follows is very much a kind of reformulation of the analysis of Sections 8.8.1 and 8.9.

Hidden in the mean-field assumptions is the general idea that physics on different scales decouple and that critical phenomena, therefore, can be described, at leading order, in terms of a finite number of effective macroscopic variables.

It is thus somewhat puzzling that the universal mean-field predictions are in quantitative (sometimes even qualitative) disagreement with the empirical information available. The leading corrections to the saddle point approximation (14.44) give some explanation. Many arguments presented in Chapter 9, adapted to dimensions $d < 4$, are also relevant here.

A remark is, here, in order: the validity of the saddle point approximation does not rely on the assumption that there are no fluctuations around the saddle point, but that these fluctuations are approximately independent of the thermodynamic variables. For example, they do not depend on the field or magnetization. In this sense, one can say that the mean-field approximation is a *quasi-Gaussian* or perturbed Gaussian approximation, since this property is exact in a Gaussian model.

14.6 Beyond the mean-field approximation

An expansion parameter. Since from the point of view of the representation (14.83) of the partition function, the mean-field approximation appears as a saddle point approximation, it is possible to calculate systematic corrections by expanding around the saddle point, generating what we will call the mean-field expansion.

However, to organize the mean-field expansion, it is convenient to introduce a parameter which orders the expansion, and singles out the mean-field approximation as a leading order. We thus replace the spin S_i on each site by the average σ_i of ℓ independent spins,

$$\sigma_i = \frac{1}{\ell} \sum_{k=1}^{\ell} S_i^{(k)}$$

with identical distribution $d\rho(S)$ [51]. For $\ell = 1$, one recovers the initial distribution.

The σ distribution given by,

$$R_\ell(\sigma) = \int \prod_k d\rho(S^{(k)}) \delta \left(\ell\sigma - \sum S^{(k)} \right),$$

can be rewritten by introducing a Fourier representation as

$$\begin{aligned} R_\ell(\sigma) &= \frac{1}{2i\pi} \int \prod_k d\rho(S^{(k)}) \int d\lambda \exp \left[\lambda \left(\sum_{k=1}^{\ell} S^{(k)} - \ell\sigma \right) \right] \\ &= \frac{1}{2i\pi} \int d\lambda \exp [\ell(w(\lambda) - \lambda\sigma)], \end{aligned}$$

where the λ integration runs along the imaginary axis, and $(w(\lambda))$ is defined by equation (14.35)). For ℓ large, the σ distribution is close to a Gaussian distribution.

We also rescale the temperature $\beta \mapsto \ell\beta$. The partition function then becomes

$$\mathcal{Z}(H) = \int \prod_i (d\sigma_i d\lambda_i) \exp \left[-\ell\beta \mathcal{H}(\sigma) + \ell \sum_i (-\lambda_i \sigma_i + w(\lambda_i)) \right]. \quad (14.83)$$

For ℓ large, the partition function (14.83) can be calculated by the steepest descent method. To simplify the comparison with the mean-field expressions, it is convenient to define

$$\mathcal{W}(H) \equiv \ln \mathcal{Z}(H)/\ell. \quad (14.84)$$

The leading order approximation is then identical to the mean-field approximation. Corrections generate a $1/\ell$ expansion.

We set

$$\Sigma_\ell(\sigma) = -\ln R_\ell(\sigma)/\ell. \quad (14.85)$$

Then, at leading order, for $\ell \rightarrow \infty$,

$$\Sigma_\ell(\sigma) = g(\sigma) + O(1/\ell), \quad (14.86)$$

where $g(\sigma)$ is the function defined in equation (14.37).

In terms of Σ , the partition function has the representation

$$\mathcal{Z}(H) = \int \prod_i (\mathrm{d}\sigma_i \mathrm{d}\lambda_i) \exp \left[-\ell \beta \mathcal{H}(\sigma) - \ell \sum_i \Sigma_\ell(\sigma_i) \right]. \quad (14.87)$$

In expression (14.87), we recognize a lattice regularized field integral for a scalar boson field, of the type considered in Section 8.7. The *mean-field approximation* is analogous to the *classical approximation of quantum field theory*.

An interpretation of the integer ℓ . We decompose the initial lattice into identical cells each containing ℓ spins. We take ℓ large but finite. Then, from the remark at the end of Section A14.1, the distribution of the average spin σ in a cell is close to a Gaussian distribution. In expression (14.87), it is obtained by completely neglecting the interaction. The interaction between average, macroscopic spins is complicated but can be approximated by an effective pair interaction because the average spins are small. The partition function (14.87) thus can be considered as the initial partition function expressed in terms of macroscopic spins, with a perturbed Gaussian distribution.

14.6.1 Perturbative expansion: The two-point function at one-loop order

At order $1/\ell$, two types of corrections appear. First, corrections to $\Sigma_\ell(\sigma)$ are generated by expanding expression (14.85) to order $1/\ell$. However, these corrections are equivalent to a modification of the coefficients of $g(\sigma)$ in the expansion in powers of σ , and we have seen that universal properties do not depend on their actual values. To calculate the other corrections, we can approximate $\Sigma_\ell(\sigma)$ by $g(\sigma)$. Then,

$$\mathcal{Z}(H) = \int \prod_i \mathrm{d}\sigma_i \exp \left[\ell \left(\sum_{i,j} \sigma_i V_{ij} \sigma_j - \sum_i g(\sigma_i) + \sum_i H_i \sigma_i \right) \right]. \quad (14.88)$$

In the Hamiltonian, we recognize the thermodynamic potential of the mean-field approximation. This suggests that, in general, fluctuations can be calculated by considering the thermodynamic potential of Landau's theory as a Hamiltonian. In the quantum field theoretical framework, this amounts to introducing an effective action of the type considered in Section 8.8.1.

The weakly perturbed Gaussian (or quasi-Gaussian) model, which reproduces the results of the mean-field approximation, is valid if the one-loop correction resulting from the Gaussian integration around the saddle point depends smoothly on H and β .

In the one-loop approximation (equation (7.93)), $\Gamma(M)$ is given by

$$\Gamma(M) = - \sum_{i,j} M_i V_{ij} M_j + \sum_i g(M_i) + \frac{1}{2\ell} \text{tr} \ln [-2V_{ij} + g''(M_i)\delta_{ij}]. \quad (14.89)$$

In the disordered phase, expanding the expression in powers of M to order M^2 , one obtains the inverse two-point function $\Gamma^{(2)}$ to order $1/\ell$. The propagator written in matrix form as (\mathbf{V} is the matrix with elements V_{ij}),

$$\Delta = [g_2 \mathbf{1} - 2\mathbf{V}]^{-1},$$

has the Fourier representation (\mathbf{n}^{ij} is the vector joining the sites i and j),

$$\Delta_{ij} = \frac{1}{(2\pi)^d} \int d^d k e^{i\mathbf{k}\cdot\mathbf{n}^{ij}} \tilde{\Delta}(\mathbf{k}),$$

where, for k small, $\tilde{\Delta}^{-1}(\mathbf{k})$ has the expansion

$$\tilde{\Delta}^{-1}(\mathbf{k}) = g_2 - 2\tilde{V}(\mathbf{k}) = g_2 - 2v + 2va^2k^2 + O(k^4). \quad (14.90)$$

Then, in the Fourier representation, the inverse two-point correlation function reads

$$\tilde{\Gamma}^{(2)}(\mathbf{p}) = g_2 - 2\tilde{V}(\mathbf{p}) + \frac{g_4}{2\ell} \int \frac{d^d k}{(2\pi)^d} \tilde{\Delta}(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \quad (14.91)$$

In particular, the inverse of the magnetic susceptibility χ is given by

$$\chi^{-1} = \tilde{\Gamma}^{(2)}(0) = g_2 - 2v + \frac{g_4}{2\ell} \int \frac{d^d k}{(2\pi)^d} \tilde{\Delta}(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \quad (14.92)$$

The critical value v_c now is solution of

$$0 = g_2 - 2v_c + \frac{g_4}{2\ell} \int \frac{d^d k}{(2\pi)^d} \tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \quad (14.93)$$

Subtracting this equation from equation (14.92), one finds

$$\chi^{-1} = 2(v_c - v) + \frac{g_4}{\ell} \int \frac{d^d k}{(2\pi)^d} [\tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k})] \tilde{\Delta}(\mathbf{k}) \tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \quad (14.94)$$

In the $1/\ell$ contribution, one can replace v_c by its mean-field value $g_2/2$. Then, for $v_c - v \rightarrow 0_+$ and $k \rightarrow 0$ (equation (14.33)),

$$\begin{aligned} \tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k}) &= (v - v_c) \tilde{U}(\mathbf{k}) \sim (v - v_c) (1 + O(k^2)) \\ \tilde{\Delta}^{-1}(\mathbf{k}) &= 2(v_c - v + v_c a^2 k^2) + O(k^4). \end{aligned}$$

14.6.2 The role of dimension 4

From these two expressions, it is clear that one can safely take the limit $v = v_c$ inside the integrand provided the integral $\int d^d k / (k^2)^2$ converges at zero momentum, that is, for $d > 4$. Then, χ^{-1} vanishes linearly in $(v - v_c)$ or $T - T_c$, as in the mean-field approximation. The susceptibility exponent remains $\gamma = 1$.

However, for $d \leq 4$, the limit $v = v_c$ is singular. Evaluating, for v close to v_c and $2 < d < 4$, the leading contribution to the integral, which comes entirely for the small k region, one finds

$$\begin{aligned}\chi^{-1} &= 2(v_c - v) \left[1 + \frac{g_4}{8v_c^2 a^d \ell} \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} \left(\frac{v_c - v}{v_c} \right)^{d/2-2} \right] + O\left(\frac{1}{\ell^2}\right), \\ &= 2(v_c - v) \left[1 + \frac{g_4}{8v_c^2 a^d \ell} \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} (\xi/a)^{4-d} \right] + O\left(\frac{1}{\ell^2}\right),\end{aligned}\quad (14.95)$$

where equation (14.72) has been used. When v approaches v_c , the $1/\ell$ correction to χ^{-1} , irrespective of how large ℓ is, always becomes larger than the mean-field term. Therefore, the predictions of the mean-field approximation can no longer be trusted [51]:

(i) For $d > 4$, the *universal* predictions of MFT remain unchanged.

(ii) For dimension $d \leq 4$, the divergence of the correction when $v \rightarrow v_c$, due to the small momentum behaviour of the propagator, shows that the predictions of mean-field approximation, or of quasi-Gaussian behaviour, cannot be correct in general. In Section 14.7, higher order corrections will be examined.

Moreover, the singular corrections involve the ratio ξ/a of macroscopic to microscopic scale. This is an indication that these two very different scales of physics do not decouple, a deep issue, which is very similar to the problem of ultraviolet divergences discussed in Chapter 8, and which will be examined again in the coming chapters.

14.7 Power counting and the role of dimension 4

The mean-field approximation can be obtained as a leading term of a steepest descent method, which reduces integrals to Gaussian integrals, a confirmation that the mean-field approximation is a perturbed Gaussian approximation.

In Section 14.6.2, by evaluating the leading correction, we have shown that the universal predictions of MFT cannot be trusted below and at four dimensions. We also note that the first divergent corrections depend on additional details of the initial microscopic model. One may then wonder whether large distance properties are short distance sensitive, which could imply that no form of universality survives. However, an inspection of equation (14.95), for example, reveals that, at leading order, only one new parameter appears, the coefficient of the σ^4 term in the small σ expansion of the interaction [120]. We now extend the analysis of the ‘infrared’ behaviour to all orders in the mean-field expansion, to investigate whether the roles of dimension 4 and σ^4 interaction survive. Since the analysis is perturbative, we assume that, in some sense, deviations from mean-field approximation are not too large, a point that requires further discussion.

In the spirit of what has been done in Section 14.6.1, we first perform, order by order in the mean-field expansion, a magnetic susceptibility or mass (in field theoretical terminology) renormalization. In the disordered phase, we trade the parameter v for the parameter μ such that

$$\omega^2 = \chi^{-1} = \tilde{\Gamma}^{(2)}(0), \quad \omega > 0, \quad (14.96)$$

(for $d < 4$ fixed, the perturbative expansion does not exist for $\omega = 0$, see Section 10.5.6) by inverting order by order the relation between v and μ (see Section 9.1).

The σ -propagator, for what concerns the most singular contribution at low momentum, can then be replaced by

$$\tilde{\Delta}(k) = \frac{1}{\omega^2 + k^2}, \quad (14.97)$$

and we want to study the behaviour of all contributions for $\omega \ll 1$.

We now consider the individual contributions coming from all (even due to $\sigma \mapsto -\sigma$ symmetry) powers of σ in $g(\sigma)$, allowing even for the possibility of polynomial momentum dependence in the corresponding vertices (this was not the case in the examples we have considered so far). These contributions have just the form of ordinary Feynman diagrams with the propagator (14.97), integrated up to a cut-off of order unity, since the momenta k_ν , $\nu = 1, \dots, d$, are limited to a Brillouin zone: $|k_\nu| \leq \pi$.

Following the analysis of Section 8.8, we rescale momenta and the deviation from criticality as (*cf.* equation (8.52)),

$$k \mapsto k/\Lambda, \quad \omega = m/\Lambda,$$

and the momenta in diagrams are cut now at scale Λ . Therefore, we trade the infrared (IR) behaviour when $\omega \rightarrow 0$ for the ultraviolet (UV) behaviour when $\Lambda \rightarrow \infty$. We also rescale the variables σ_i as in equation (8.52) (completing the Gaussian renormalization). We are now in the situation discussed in Section 8.9.

In a generic situation, the large scale properties of the lattice model can be reproduced by an effective σ^4 field theory in the continuum, the coefficient of the σ^4 interaction being not the coefficient read in the initial Hamiltonian but a coefficient renormalized by the other interactions [120]:

$$\mathcal{H}(\sigma) = \int d^d x \left[\frac{1}{2} (\nabla \sigma(x))^2 + \frac{1}{2} r \sigma^2(x) + \frac{1}{4!} \Lambda^{d-4} g \sigma^4(x) \right], \quad (14.98)$$

where the short-distance lattice regularization can be replaced by more convenient continuum regularizations, like in equation (8.22). In Chapters 15 and 16, we investigate the large distance behaviour of correlation functions in this effective σ^4 field theory. In Chapter 17, we discuss the effect of the neglected subleading divergences to show the internal consistency of the method.

14.8 Tricritical points

So far, we have assumed that only one control parameter could be adjusted, the temperature and that, therefore, the coefficient g_4 of the term M^4 in the expansion of $\Gamma(M)$ was positive but otherwise generic, that is, a number of order unity. However, there are situations in which an additional physical parameter can be varied, and both the M^2 and M^4 terms can be cancelled, leading to a tricritical behaviour [123]. This occurs for instance in He³-He⁴ mixtures or some metamagnetic systems. In the Ising-like models, this can be achieved by introducing parameters in the spin distribution.

Let us expand $\Gamma(M)$ up to order M^6 :

$$\Gamma(M) = - \sum_{i,j} V_{ij} M_i M_j + \sum_i \left(\frac{g_2}{2!} M_i^2 + \frac{g_4}{4!} M_i^4 + \frac{g_6}{6!} M_i^6 + \dots \right); \quad (14.99)$$

If the coefficient g_6 of M^6 is positive, when g_4 decreases it is possible to follow a line of critical points until g_4 vanishes. At this point, the tricritical point, the M^6 term becomes relevant, and a new analysis has to be performed. The various critical exponents have values different (like $\beta = 1/4$, $\delta = 5$) from those found for an ordinary critical point. After the tricritical point, g_4 becomes negative and the transition becomes first order.

Corrections to the tricritical theory can be studied by the method of Section 14.7. One finds that the upper critical dimension now is 3, that is, above three dimensions the mean-field approximation correctly predicts universal quantities, whereas it is definitely not valid in three dimensions and below. Moreover, using power counting arguments, one can show that the most singular corrections are reproduced by a continuum ϕ^6 QFT.

A14 Additional considerations

A14.1 High-temperature expansion

The simplest method to calculate $\mathcal{Z}(H)$ is to expand expression (14.30) in powers of the pair potential V or equivalently in powers of β (high temperature expansion) at fixed field H_i . At infinite temperature ($V_{ij} = 0$), all sites decouple and the partition function \mathcal{Z}_0 is obtained from the solution of the one-site problem:

$$\mathcal{Z}_0(H) = \prod_i z(H_i). \quad (\text{A14.1})$$

Higher order terms in the high temperature expansion,

$$\mathcal{Z}(H)/\mathcal{Z}_0(H) = 1 + \sum_{i,j} V_{ij} \langle S_i S_j \rangle_0 + \frac{1}{2!} \sum_{i,j,k,l} V_{ij} V_{kl} \langle S_i S_j S_k S_l \rangle_0 + \dots, \quad (\text{A14.2})$$

can then all be expressed in terms of the moments of the one-site spin distribution in a magnetic field:

$$\langle S_i^n \rangle_0 = z^{-1}(H_i) \left(\frac{\partial}{\partial H_i} \right)^n z(H_i). \quad (\text{A14.3})$$

For example, the corresponding Weiss free energy $\mathcal{W}(H) = \ln \mathcal{Z}(H)$ at order β , with the assumption $V_{ii} = 0$, is given by

$$\mathcal{W}(H) = \sum_i w(H_i) + \sum_{i,j} V_{ij} \langle S_i \rangle_0 \langle S_j \rangle_0 + O(\beta^2)$$

with the definition (14.35). The local magnetization follows,

$$M_i = \langle S_i \rangle = \frac{\partial \mathcal{W}}{\partial H_i} = w'(H_i) + 2 \sum_j V_{ij} \left(\langle S_i^2 \rangle_0 - (\langle S_i \rangle_0)^2 \right) \langle S_j \rangle_0 + O(\beta^2). \quad (\text{A14.4})$$

Finally, the thermodynamic potential $\Gamma(M)$, Legendre transform of $\mathcal{W}(H)$, can be written (using the definition (14.37) and the usual tricks) as

$$\Gamma(M) = \sum_i H_i M_i - \mathcal{W}(H) = \sum_i g(M_i) - \sum_{i,j} V_{ij} M_i M_j + O(\beta^2). \quad (\text{A14.5})$$

The high temperature expansion is useful mainly if the leading term, the infinite temperature result, is already qualitatively correct: the expansion gives information about the high temperature phase and necessarily diverges at the critical temperature. By contrast, as we have seen, the mean-field approximation makes it possible to describe the behaviour of thermodynamic quantities at a phase transition. Diagrammatically, the mean-field approximation corresponds to a partial summation (loop) of the high temperature expansion.

Remark. The distribution of the average spin $\sigma = \Omega^{-1} \sum_i S_i$, where Ω is the total number of lattice sites, is given, in the large volume and thus large Ω limit, by $e^{-\Gamma(\sigma)}$, where Γ is the thermodynamic potential for uniform magnetization, as we have argued in Section 7.11.2. As long as the volume is large but remains finite, the high temperature expansion converges, and $\Gamma(\sigma)$ has still a regular expansion for σ small, as in the infinite temperature limit. Since $\Gamma(\sigma)$ is proportional to Ω , the distribution of σ is close to a Gaussian distribution, a result which is a slight extension of the central limit theorem of probabilities. This remark has been used in Section 14.6.1.

A14.2 Mean-field approximation: General formalism

We now explain how the mean-field approximation can be derived for a general spin model on a lattice [124, 51, 125]. Again, the mean-field approximation can be identified with a saddle point approximation and, therefore, is the first term in a systematic expansion [125]. Finally, for completeness, we indicate how the mean-field approximation is related with the high and low temperature expansions.

In what follows, we consider systems with only one degree of freedom per lattice site but the generalization of the formalism to several degrees of freedom per site is straightforward. The mean-field approximation can also be generalized to lattice models with fermions. The equivalent of the variables S_i are then products of the form $\bar{\psi}_i \Gamma_A \psi_i$, where ψ_i and $\bar{\psi}_i$ are the fermion fields and Γ_A some element of the algebra of γ matrices. These fields correspond to ‘composite bosons’. One also has to use a measure appropriate to Grassmann variables.

A14.2.1 Mean-field approximation

We consider a lattice model described in terms of lattice variables S_i , where i denotes a lattice site, and a configuration energy or Hamiltonian \mathcal{H} . In what follows, it will be necessary to treat differently different powers of the same lattice variable S_i so that we write \mathcal{H} as $\mathcal{H}(S, S^2, \dots, S^n)$. For example, a general pair potential has to be written as

$$\sum_{i,j} V_{ij} S_i S_j = \sum_{i \neq j} V_{ij} S_i S_j + \sum_i V_{ii} S_i^2.$$

The corresponding partition function is (the temperature factor β is incorporated in \mathcal{H})

$$\mathcal{Z} = \int \left(\prod_i d\rho(S_i) \right) \exp [-\mathcal{H}(S, S^2, \dots, S^n)], \quad (A14.6)$$

in which $d\rho(S)$ is the spin distribution.

In order to construct the mean-field approximation as the leading order in an evaluation by the steepest descent method, it is necessary to first transform expression (A14.6). We introduce two sets of lattice variables $\sigma_i^{(m)}$ and $\lambda_i^{(m)}$, the quantities $\lambda_i^{(m)}$ being Lagrange multipliers, to express the constraints

$$\sigma_i^{(m)} = S_i^m, \quad m = 1, \dots, n. \quad (A14.7)$$

The partition function can then be rewritten as

$$\begin{aligned} \mathcal{Z} &= \int \prod_i d\rho(S_i) \prod_{i,m} d\sigma_i^{(m)} d\lambda_i^{(m)} \\ &\times \exp \left[-\mathcal{H}(\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(n)}) + \sum_{i,m} \lambda_i^{(m)} (S_i^m - \sigma_i^{(m)}) \right]. \end{aligned} \quad (A14.8)$$

Our definition of the Hamiltonian \mathcal{H} is such that in each monomial contributing to \mathcal{H} , which is a product of variables $\sigma_i^{(m)}$, any given site i can appear at most once. No product of the form $\sigma_i^{(m)} \sigma_i^{(m')}$ for any value of the site position i can be found.

In expression (A14.8), the integrations over all variables S_i are decoupled. We introduce the function w defined by

$$\exp \left[w(\lambda^{(1)}, \dots, \lambda^{(n)}) \right] = \int d\rho(S) \exp \left(\sum_{m=1}^n S^m \lambda^{(m)} \right). \quad (\text{A14.9})$$

Here, we assume that the integration measure $d\rho(S)$ is either of compact support or decreasing fast enough so that the integral exists. The partition function can then be rewritten as

$$\mathcal{Z} = \int \prod_{i,m} d\sigma_i^{(m)} d\lambda_i^{(m)} \exp \left[-\mathcal{H}(\sigma) - \sum_{i,m} \lambda_i^{(m)} \sigma_i^{(m)} + \sum_i w(\lambda_i^{(1)}, \dots, \lambda_i^{(n)}) \right]. \quad (\text{A14.10})$$

The mean-field approximation is the leading order in the evaluation of this expression by the steepest descent method. The saddle point equations are

$$\frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} = -\lambda_i^{(m)}, \quad (\text{A14.11a})$$

$$\frac{\partial w}{\partial \lambda_i^{(m)}} = \sigma_i^{(m)}. \quad (\text{A14.11b})$$

Equation (A14.11b) shows that, at leading order, the original weight $\prod_i d\rho(S_i) e^{-\mathcal{H}}$ has been approximated by a product of weights for each lattice site given by the right-hand side of equation (A14.9), in which the variables $\lambda_i^{(m)}$ have been replaced by their saddle point values, the ‘mean fields’.

Remarks. It is clear from these equations that, if \mathcal{H} does not depend on a given $\sigma^{(m)}$, the corresponding field $\lambda^{(m)}$ vanishes, and both disappear from the equations. Therefore, we find the same result as if we had omitted them in the first place. Note also that if \mathcal{H} is a quadratic function of the variables $\sigma^{(m)}$, one may choose to perform the $\sigma^{(m)}$ integrations explicitly.

Discussion. Let us now explain why we have introduced the n sets of variables $\sigma^{(m)}$ and $\lambda^{(m)}$, instead of one as in the example studied in Section 14.4. In the mean-field approximation, the variables S_i are replaced by some expectation value so that the expectation value of a product is replaced by the product of expectation values:

$$\langle S_{i_1} S_{i_2} \dots S_{i_n} \rangle \mapsto \langle S_{i_1} \rangle \langle S_{i_2} \rangle \dots \langle S_{i_n} \rangle.$$

It is plausible that, in some limit, this can be a good approximation if all the sites in the product are different. However, this cannot be true if in the product appears a power of variables on the same site. For example, in the Ising model

$$S_i^2 = 1 \neq (\langle S_i \rangle)^2.$$

We want to treat the one-site problem exactly, and the new Lagrange parameters makes it possible to take into account these self-correlations. Let us add a few comments which illustrate this point.

Comments. If the Hamiltonian \mathcal{H} is a function only of S_i^2 , then it is obvious that we should consider S_i^2 as the basic dynamical variable in the mean-field approximation. The procedure explained previously does it automatically, since only the parameters $\sigma^{(2)}$ and $\lambda^{(2)}$ will appear.

This formalism also solves the following simple problem. Terms in the Hamiltonian which correspond to one-body potentials, that is, sum of functions depending only on the variable at one site such as

$$\sum_i \frac{1}{2} a S_i^2 + \frac{1}{4!} b S_i^4 + \dots , \quad (A14.12)$$

could be considered both as part of the measure $\rho(S_i)$ or as part of the Hamiltonian. Therefore, one could fear that the results of the mean-field approximation depend on the formulation. The introduction of additional variables ensures that this is not the case.

Let us, for example, make the transformation

$$\begin{aligned} \mathcal{H}(S_i, S_i^2, \dots, S_i^n) &\mapsto \mathcal{H}(S_i, \dots, S_i^n) + \frac{a}{2} \sum_i S_i^2, \\ d\rho(S) &\mapsto e^{aS^2/2} d\rho(S). \end{aligned} \quad (A14.13)$$

Obviously, the lattice theory is independent of a .

Consider now the modifications this transformation induces into equations (A14.9) and (A14.11):

$$\exp\left(\tilde{w}(\lambda^{(1)}, \dots, \lambda^{(n)})\right) = \int d\rho(S) e^{aS^2/2} \exp\left(\sum_{m=1}^n S^m \lambda^{(m)}\right). \quad (A14.14)$$

Comparing the equation with equation (A14.9), we conclude

$$\tilde{w}(\lambda^{(1)}, \dots, \lambda^{(n)}) = w(\lambda^{(1)}, \lambda^{(2)} + a/2, \lambda^{(3)}, \dots, \lambda^{(n)}), \quad (A14.15)$$

and the saddle point equations become

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} + \frac{a}{2} \delta_{l2} &= -\lambda_i^{(m)}, \\ \frac{\partial w}{\partial \lambda_i^{(m)}}(\lambda^{(1)}, \lambda^{(2)} + a/2, \lambda^{(3)}, \dots, \lambda^{(n)}) &= \sigma_i^{(m)}. \end{aligned} \quad (A14.16)$$

We note that the equations are identical up to the substitution

$$\lambda_i^{(2)} + a/2 \mapsto \lambda_i^{(2)}, \quad (A14.17)$$

which does not change mean-field results.

Finally, we now understand the technical assumption $V_{ii} = 0$ of Section 14.4: it simplifies the mean-field formalism by avoiding the introduction of variables related to S_i^2 .

A14.3 Mean-field expansion

It is possible to correct the mean-field approximation, by systematically expanding the effective Hamiltonian in (A14.8) around the saddle points (A14.11), and integrating term by term, as one does in usual perturbation theory. We call this expansion the mean-field expansion. In Section 14.4, in a particular example, we have found it convenient to associate a parameter to this expansion. This can also be done in this more general framework. If we set a parameter ℓ in front of the effective Hamiltonian (A14.10), then the mean-field expansion is just a formal expansion in powers of $1/\ell$:

$$\begin{aligned} \mathcal{Z}_\ell &= \int \prod_{i,m} d\sigma_i^{(m)} d\lambda_i^{(m)} \exp [-\ell \mathfrak{H}(\lambda, \sigma)], \\ \mathfrak{H}(\lambda, \sigma) &= \mathcal{H}(\sigma) + \sum_{i,m} \lambda_i^{(m)} \sigma_i^{(m)} - \sum_i w(\lambda_i^{(1)}, \dots, \lambda_i^{(n)}). \end{aligned} \quad (A14.18)$$

As in Section 14.4.2, an interpretation can be given to the parameter ℓ , by replacing on each lattice site the spin variable S by ℓ independent variables $S^{(\alpha)}$ having the same distribution $d\rho(S)$:

$$\int d\rho(S_i) \mapsto \int \prod_{\alpha=1}^{\ell} d\rho(S_i^{(\alpha)}). \quad (A14.19)$$

Then,

$$\exp [\ell w(\lambda^{(1)}, \dots, \lambda^{(n)})] = \int \prod_{\alpha} d\rho(S^{(\alpha)}) \exp \left(\sum_{m=1}^n \lambda^{(m)} \sum_{\alpha} (S^{(\alpha)})^m \right). \quad (A14.20)$$

We insert this expression into the partition function (A14.18). We then integrate over the Lagrange multipliers $\lambda^{(m)}$ and find

$$\sigma_i^{(m)} = \frac{1}{\ell} \sum_{\alpha} (S^{(\alpha)})^m. \quad (A14.21)$$

Therefore, as in the simpler example of Section 14.4.2, the spin variable S_i , when it appears linearly, is replaced by the average of ℓ independent spins. However, the various powers of the spin variables are replaced by the corresponding average of the powers of the ℓ spins, and not the power of the average.

Finally, again the temperature is rescaled $\beta \mapsto \ell\beta$. For ℓ large, the various variables $\sigma^{(m)}$ become Gaussian variables with a dispersion of order $1/\sqrt{\ell}$.

The parameter ℓ is just a formal parameter, since really we want to calculate for $\ell = 1$, but its introduction leads to some simple considerations. It shows that the mean-field approximation becomes exact when the number of identical independent variables on each site becomes large, provided the amplitudes of the interaction terms are scaled appropriately. Such a situation is, for instance, approximately realized when d , the dimension of space, becomes large. Then a cell of fixed linear dimension contains an increasing number of independent variables. By re-expanding terms in the mean-field expansion, one can generate a systematic $1/d$ expansion.

A14.4 High-, low-temperature, and mean-field expansions

We now exhibit some relations between the mean-field expansion and the high and low temperature expansions.

A14.4.1 High temperature

We restore an explicit factor β in front of the Hamiltonian \mathcal{H} and consider an expansion in powers of β , which is a high temperature expansion. First, it is easy to expand the saddle point equations (A14.11) at first order in β and to verify that one recovers the high temperature expansion at order β :

$$\mathcal{W} = -\beta \mathcal{H}(\sigma), \quad \sigma_i^{(m)} = \langle S^m \rangle_0 .$$

More generally, the power of ℓ counts the number of loops of a connected high temperature diagram.

Indeed, first each vertex of the Hamiltonian is multiplied by a factor $\beta\ell$. From the relation (A14.21), we infer that a diagram of order K in β involving vertices that connect v_1, v_2, \dots, v_K lattice sites, respectively, is affected by ℓ to the power $-\sum v_k$. Moreover, additional factors of ℓ are generated because each $\sigma^{(m)}$ variable is the sum of ℓ independent contributions: the spin average yields a power which counts the number s of different lattice sites present in the product of the K vertices. One thus finds

$$\beta^K \mapsto K - \sum_{k=1}^K v_k + s . \quad (\text{A14.22})$$

For a connected tree diagram, one has

$$s = \sum_{k=1}^K v_k - (K - 1) . \quad (\text{A14.23})$$

Indeed, each interaction term brings in v_k independent variables. But to construct a connected diagram, each interaction term must have a variable in common with another interaction term. This suppresses exactly $(K - 1)$ independent variables.

Now, each time one adds one loop to the diagram, one suppresses one additional independent variable. Calling B the number of loops of the diagram, one finds

$$s = \sum_{k=1}^K v_k - (K - 1) - B . \quad (\text{A14.24})$$

The corresponding power of ℓ for the diagram is then

$$K - \sum_{k=1}^K v_k + s = 1 - B .$$

Mean-field expansion is an expansion in powers of ℓ . Therefore, at a given order, it sums all high temperature or strong coupling diagrams with the same number of loops.

A14.4.2 Low temperature expansion

The mean-field expansion also contains the low temperature expansion. We rewrite equations (A14.11) with a parameter β in front of the Hamiltonian as

$$\beta \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} = -H_i^{(m)}, \quad \frac{\partial w}{\partial H_i^{(m)}} = \sigma_i^{(m)}. \quad (\text{A14.25})$$

For large β (low temperature or weak coupling), the $H^{(m)}$ variables become large. To study this limit, we have thus to evaluate w for H large. This, in general, selects for S some classical value S_c . As a direct consequence, we find

$$\frac{\partial w}{\partial H_i^{(m)}} = (S_{i,c})^m = \sigma_i^{(m)}. \quad (\text{A14.26})$$

In this limit, the variables σ and H no longer play a role and one expands around the configuration that dominates at low temperature (large β). To show this more explicitly, we introduce the Fourier representation of the measure

$$d\rho(S) = dS \frac{1}{2i\pi} \int_{-i\infty}^{+i\infty} e^{\mu S} \tilde{\rho}(\mu) d\mu. \quad (\text{A14.27})$$

A low temperature expansion is possible only if $d\rho(S)$ decrease fast enough for large $|S|$, and thus $\tilde{\rho}(\mu)$ has to be an entire function.

For $H^{(m)}$ large, w can be calculated by the steepest descent method. The saddle point equations are

$$\mu = - \sum_m m H^{(m)} S^{m-1}, \quad (\text{A14.28})$$

$$S + \frac{\partial}{\partial \mu} \ln \tilde{\rho}(\mu) = 0, \quad (\text{A14.29})$$

and

$$w(H^{(m)}) = \sum_m H^{(m)} S^m + \mu S + \ln \tilde{\rho}(\mu). \quad (\text{A14.30})$$

The mean-field saddle point equations then are

$$\sigma_i^{(m)} = \frac{\partial w}{\partial H_i^{(m)}} = S_i^m, \quad (\text{A14.31})$$

$$H_i^{(m)} = -\beta \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}}. \quad (\text{A14.32})$$

In the same notation, the relevant configuration at low temperature is given by the equations

$$\beta \frac{d\mathcal{H}}{dS_i} \equiv \sum_i m S_i^{m-1} \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} = \mu_i, \quad (\text{A14.33})$$

and

$$S_i + \frac{d}{d\mu_i} \ln \tilde{\rho}(\mu) = 0. \quad (\text{A14.34})$$

Now, by summing equation (A14.32) over m after multiplication by $m(S_i)^{m-1}$, and using equation (A14.28), one reproduces equation (A14.33), while equations (A14.29) and (A14.34) are identical.

The mean-field expansion is also a partial summation of the low temperature expansion.

A14.5 Quenched averages

In this work, we discuss only the so-called annealed averages, that is, averages over random configurations for systems which explore the whole available phase space in the course of time evolution. This leads to the concept of partition function. However, another class of problems is of quite different nature, where the disorder is related to frozen degrees of freedom, or at least degrees of freedom (for instance impurities in crystals) which do not evolve during the time of observation. When a relaxation time can be defined (see Chapter 36), the relaxation time is large compared to the time of observation.

Then, it is no longer the partition function that has to be averaged over disorder, but directly the physical observables, a procedure usually called *quenched average*.

We give here, for illustration, a simple example, a Gaussian model in a quenched random magnetic field with Gaussian distribution. For simplicity, we assume continuum space and a continuous spin distribution $\sigma(x)$ corresponding to the local free action

$$\mathcal{S}(\sigma) = \mathcal{S}_0(\sigma) - \int d^d x H(x)\sigma(x),$$

with

$$\mathcal{S}_0(\sigma) = \frac{1}{2} \int d^d x [(\nabla\sigma(x))^2 + \mu^2\sigma^2(x)].$$

For each magnetic field distribution, the partition function is given by

$$\mathcal{Z} = \int [d\sigma(x)] \exp [-\mathcal{S}(\sigma)].$$

We assume the magnetic fields at different points to be uncorrelated, with a Gaussian distribution characterized by the one- and two-point functions:

$$\overline{H(x)} = h, \quad \overline{H(x)H(y)}_c = w^2\delta^{(d)}(x-y),$$

h and w being two constants and $\overline{\bullet}$ meaning quenched average.

In this model, one can immediately calculate the expectation value of the spin variable, at fixed field distribution, at a point x ,

$$\langle \sigma(x) \rangle = \int d^d y \Delta(x-y)H(y),$$

where $\Delta(x)$ is the σ -field propagator. If one now averages $\langle \sigma(x) \rangle$ over the position x in a large volume Ω , one finds

$$\Omega^{-1} \int d^d x \langle \sigma(x) \rangle = \int d^d x \Delta(x) \Omega^{-1} \int d^d y H(y).$$

Because the propagator in the absence of disorder is invariant under translation, one finds in the right-hand side the average of the magnetic field. The result is

$$\Omega^{-1} \int d^d x \langle \sigma(x) \rangle = \frac{1}{\mu^2} \overline{H(x)} = \frac{h}{\mu^2}.$$

The average of the two-point function over the sample is

$$\begin{aligned}\Omega^{-1} \int d^d y \langle \sigma(y) \rangle \langle \sigma(x+y) \rangle &= \Omega^{-1} \int d^d y d^d z d^d t \Delta(y-z) \Delta(x+y-t) H(z) H(t) \\ &= \Omega^{-1} \int d^d y d^d z \Delta(y) \Delta(x+y+z) \int d^d t H(z+t) H(t).\end{aligned}$$

Again, translation invariance, in the absence of field, reduces the space average to a field average:

$$\Omega^{-1} \int d^d t H(z+t) H(t) = \overline{H(z+t) H(t)} = h^2 + w^2 \delta^{(d)}(z).$$

It follows that

$$\Omega^{-1} \int d^d y \langle \sigma(y) \rangle \langle \sigma(x+y) \rangle = \frac{h^2}{\mu^4} + \frac{w^2}{(2\pi)^d} \int \frac{e^{ipx} d^d p}{(p^2 + \mu^2)^2}.$$

Note on the other hand that the σ -field two-point function is independent of the magnetic field:

$$\langle \sigma(y+x) \sigma(y) \rangle = \frac{1}{(2\pi)^d} \int \frac{e^{ipx} d^d p}{p^2 + \mu^2}.$$

The replica trick [126]. In this simple example, since the observables can be calculated explicitly, it is easy to average over the quenched random variables. This is not the case in general, but an algebraic trick has been invented to overcome this difficulty. To average correlation functions is equivalent to averaging the free energy $\mathcal{W} = \ln \mathcal{Z}$ instead of the partition function. Then, the following identity is used:

$$\ln \mathcal{Z} = \lim_{n \rightarrow 0} \frac{\mathcal{Z}^n - 1}{n}.$$

One does not know how to calculate \mathcal{Z}^n for general real n , but one knows how to calculate it for integer n . It is sufficient to replicate n times the initial variables (the *replica trick*) [126]. In our example, one introduces a field σ_i with n components and obtains

$$\mathcal{Z}^n(J) = \int \prod_{i=1}^n [d\sigma_i(x)] \exp \left[- \sum_i \mathcal{S}(\sigma_i) + \sum_i \int d^d x J(x) \sigma_i(x) \right].$$

The Gaussian disorder average is easy to perform:

$$\begin{aligned}\overline{\mathcal{Z}^n(J)} &= \int [d\sigma_i(x)] \exp \left[- \sum_i \mathcal{S}_0(\sigma_i) \right. \\ &\quad \left. + \frac{1}{2} w^2 \int d^d x \sum_{i,j} \sigma_i(x) \sigma_j(x) + \int d^d x \sum_i (h + J(x)) \sigma_i(x) \right].\end{aligned}$$

The subtle question is then to take the $n = 0$ limit. For instance, to calculate the field expectation value, one minimizes the potential. One finds the replica symmetric solution:

$$\langle \sigma_i(x) \rangle = \frac{h}{\mu^2 - n\omega^2}.$$

In this simple example, one then sets $n = 0$ and obtains the correct result,

$$\overline{\langle \sigma(x) \rangle} = \lim_{n \rightarrow 0} \frac{1}{n} \sum_i \sigma_i = \frac{h}{\mu^2}.$$

However, note that even here the solution does not make sense for $n > \mu^2/\omega^2$ and, therefore, arguments which state that under certain conditions analytic functions known for all integers are uniquely defined, do not apply and other arguments have to be looked for.

Similarly, the replica propagator can be calculated as

$$\langle \tilde{\sigma}(p) \tilde{\sigma}(-p) \rangle = \frac{\delta_{ij}}{p^2 + \mu^2} + \frac{\omega^2}{(p^2 + \mu^2)(p^2 + \mu^2 - n\omega^2)}.$$

Summing over i, j , dividing by n , and taking the $n = 0$ limit yields the correct result. Notice that, for $m = 0$ small, the replica-propagator for $n = 0$ is more singular than the free field propagator at low momentum. This has implications for critical phenomena in disordered systems.

For less trivial models, the problem is quite complicated. The famous spin-glass model is difficult to solve even in the mean-field approximation, and the mean-field solution relies on a breaking of the symmetry between field replica, as shown by Parisi [127].

A final remark: to calculate correlations due to the disorder averaging, we have to evaluate instead quantities like

$$\overline{\mathcal{W}(J_1) \mathcal{W}(J_2)} = \overline{\ln \mathcal{Z}(J_1) \ln \mathcal{Z}(J_2)}.$$

The replica trick clearly generalizes with $2n$ replicas.

15 The renormalization group (RG) approach: The critical theory near four dimensions

In Chapter 14, we have studied Ising-like systems (and more generally ferromagnetic systems with $O(N)$ symmetry) with short range interactions, and determined the behaviour of thermodynamic functions near a second order phase transition, in the mean-field approximation, which is, as we have indicated, a *quasi-Gaussian* approximation. The mean-field approximation predicts a set of *universal* properties, that is, properties independent of the detailed structure of the microscopic Hamiltonian, the dimension of space, and, to a large extent, of the symmetry of systems. We have then systematically examined corrections to the mean-field approximation. We have found that, above four dimensions, these corrections do not change universal quantities. By contrast, below and at four dimensions, the corrections diverge at the critical temperature, and the universal predictions of the mean-field approximation can certainly not be correct. Such an analysis shows that different scales of physics do not decouple, and leads to the important question, how much do the large scale properties of the critical theory depend on the microscopic structure? In Chapter 9, we have answered the question in four dimensions. With some general assumptions about the relevance of perturbation theory, we have shown that the large scale properties can be described by a renormalized quantum field theory, which is largely universal, and depends only on a small number of effective parameters, masses and ϕ^4 -like coupling constants. Moreover, below four dimensions, we have identified, in an expansion around the mean field, the most singular terms near criticality, and found indications that they can be also formally recovered from a continuum, low mass ϕ^4 field theory (Section 14.6), a result we use later in the chapter.

However, to understand universality beyond the mean-field approximation, it is necessary to call upon a new idea. Flow equations for effective Hamiltonians based on the RG concept provide the essential tools one needs.

The RG idea, as applied to critical phenomena, has been suggested by Kadanoff [128] and then implemented by Wilson [129]. We first describe the basic RG ideas in a somewhat abstract and intuitive framework. The formulation lacks precision, and the arguments are largely heuristic, but the importance of fixed points in Hamiltonian space emerges. The special role of Gaussian models and their universal properties can be related to the existence of a fixed point, the Gaussian fixed point.

The abstract RG scheme is extremely suggestive, but it requires discovering proper implementations, which is not a straightforward task. Non-perturbative implementations can be found in Refs. [61] and [62]. Numerical investigations have been initiated by Wetterich [130]. However, once the role of the ϕ^4 interaction was recognized by Wilson [120], perturbative methods started to be investigated.

In the framework of the *perturbative mean-field expansion*, this problem has strong similarities with the question of the low mass, low momentum behaviour in effective quantum field theories (QFTs), as discussed in Chapter 8. Universality within perturbation theory relies first on the recognition of the role of renormalizable field theories. Universality is then related, to a large extent, to the question of *regularization independence*, a problem tackled by using *renormalization theory* (Chapter 9). Asymptotic RG equations with a small number of parameters then follow (Section 9.11).

Therefore, in Section 15.3, we use what we have learned both from the analysis of corrections to the mean-field approximation in Section 14.4, and of the relevant eigenoperators at the Gaussian fixed point near dimension 4 [131], and construct an asymptotic RG for an effective ϕ^4 QFT [72]. We show that the RG equations that appear as a consequence of the renormalizability of a class of local field theories provide a direct realization of the abstract RG equations we have introduced in Section 15.1. Universality in the theory of critical phenomena is thus directly related to the property that renormalizable local field theories are insensitive, to a large extent, to the short distance structure and physics can, therefore, be inferred, at leading order at large distance, from the properties of renormalized correlation functions of QFT. Conversely, from the viewpoint of phase transitions, QFTs are always close to criticality and their existence, beyond perturbation theory, relies, from the abstract RG point of view, on the presence of fixed points in Hamiltonian space.

The results confirm in a very direct manner the deep connection, first recognized by Wilson, between QFT and the theory of critical phenomena.

Finally, after dimensional continuation (Chapter 10), the solution of the RG equations in $d = 4 - \varepsilon$ dimensions makes it possible to calculate universal quantities in an ε -expansion, following Wilson and Fisher [75].

More generally, using the methods of perturbative QFT, it is possible to efficiently calculate universal quantities, and derive universal properties to all orders, for a large class of statistical models of physical interest [80, 132, 133] (see also the textbooks Refs. [134, 135], and the elementary introduction Ref. [64]).

Many systems to which the RG predictions, based on the $O(N)$ symmetric $(\phi^2)^2$ field theory, apply are non-magnetic. In this chapter, we consider three systems, or classes of systems, which are of special interest, because they lead to precise comparisons between RG results, experiments, and simulations: $N = 0$ describes the statistical properties of polymer chains (or self-avoiding random walks) asymptotically when their length becomes large [136]; $N = 1$ and $N = 2$ also describe the liquid–vapour and the superfluid helium transitions, respectively.

15.1 RG: The general idea

We consider a general effective (because it describes only the large scale properties of some initial microscopic model) Hamiltonian, local functional of a field $\phi(x)$ (the classical spin of the lattice models of Chapter 14) in the continuum, in d dimensions, of the form,

$$\mathcal{H}(\phi) = \sum_{n=0} \frac{1}{n!} \int d^d x_1 d^d x_2 \cdots d^d x_n \mathcal{H}_n(x_1, x_2, \dots, x_n) \phi(x_1) \cdots \phi(x_n), \quad (15.1)$$

which has the properties of the thermodynamic potential (14.82) of Landau’s theory (Section 14.5.4) as well as of the effective actions of Section 8.8.1. We assume that, in the domain of interest, it is a regular function of all parameters; that it is *translation invariant* and, therefore, the Fourier transforms of the functions \mathcal{H}_n are products of a δ -function of momentum conservation by functions analytic at low momenta (assumption of short range forces or locality). Such a Hamiltonian depends on an infinite number of parameters or coupling constants.

To the Hamiltonian $\mathcal{H}(\phi)$ (or configuration energy), correspond connected correlation functions defined by

$$W^{(n)}(x_1, x_2, \dots, x_n) = \left[\int [d\phi] \phi(x_1) \cdots \phi(x_n) e^{-\beta \mathcal{H}(\phi)} \right]_{\text{connect.}}. \quad (15.2)$$

We want to study the large distance behaviour of connected correlation functions at a second order phase transition, that is, the behaviour of $W^{(n)}(\lambda x_1, \dots, \lambda x_n)$ (with all points distinct) when the dilatation parameter λ becomes large.

15.1.1 The RG idea: Fixed points

The RG idea is to trade the initial problem, studying the behaviour of correlation functions as a function of dilatation parameter λ acting on space variables, for the study of the flow of a scale-dependent (effective) Hamiltonian $\mathcal{H}_\lambda(\phi)$, which has the same large scale correlation functions, at fixed space variables. More precisely, we want to construct a Hamiltonian $\mathcal{H}_\lambda(\phi)$ that has correlation functions $W_\lambda^{(n)}(x_i)$ satisfying

$$W_\lambda^{(n)}(x_1, \dots, x_n) - Z^{-n/2}(\lambda) W^{(n)}(\lambda x_1, \dots, \lambda x_n) = R_\lambda^{(n)}(x_1, \dots, x_n), \quad (15.3)$$

where the functions $R^{(n)}$ decrease faster than any power of λ for $\lambda \rightarrow \infty$. The mapping $\mathcal{H}(\phi) \mapsto \mathcal{H}_\lambda(\phi)$ is called an RG transformation. We define the transformation such that $\mathcal{H}_{\lambda=1}(\phi) \equiv \mathcal{H}(\phi)$.

The Hamiltonian flow can be constructed by a partial integration over short distance degrees of freedom [61, 62] (see also Ref. [64] for an elementary introduction). Then, the Hamiltonian $\mathcal{H}_\lambda(\phi)$ has a $\lambda \rightarrow \infty$ limit only for a suitable choice of the renormalization factor $Z(\lambda)$ (the simplest example is provided by the central limit theorem of probabilities [64]).

Due to the assumption of space-translation invariance, after a Fourier transformation and a factorization of the δ -function of momentum conservation, equation (15.3) becomes

$$\widetilde{W}_\lambda^{(n)}(p_1, \dots, p_n) = Z^{-n/2}(\lambda) \lambda^{(1-n)d} \widetilde{W}^{(n)}(p_1/\lambda, \dots, p_n/\lambda) + \widetilde{R}_\lambda^{(n)}. \quad (15.4)$$

Various RG transformations differ by the form of $R^{(n)}$, and the function $Z(\lambda)$. In explicit constructions, the $R^{(n)}$ are generated by the integration over the large momentum modes of $\phi(x)$. When both space and field are continuous variables, one can find RG transformations with $R^{(n)} \equiv 0$. The simplest such RG transformation corresponds to a rescaling of space and field, but this transformation has a fixed point only in exceptional cases (Gaussian models), and thus more general transformations are required [61, 62]. In the following equations we omit the terms $R^{(n)}$ and thus equalities between correlation functions will mean up to terms decreasing faster than any power.

The fixed point Hamiltonian. The coupling constants appearing in \mathcal{H}_λ are all explicit functions of λ . We assume that we have found an RG transformation such that, for $\lambda \rightarrow \infty$, the Hamiltonian $\mathcal{H}_\lambda(\phi)$ has a limit $\mathcal{H}^*(\phi)$, the (infrared (IR)) fixed point Hamiltonian. If such a fixed point exists in Hamiltonian space, then the correlation functions $W_\lambda^{(n)}$ have corresponding limits $W_*^{(n)}$, and equation (15.3) becomes

$$W^{(n)}(\lambda x_1, \dots, \lambda x_n) \underset{\lambda \rightarrow \infty}{\sim} Z^{n/2}(\lambda) W_*^{(n)}(x_1, \dots, x_n). \quad (15.5)$$

We now introduce a second scale parameter μ , and calculate $W^{(n)}(\lambda \mu x_i)$ from equation (15.5) in two different ways. We obtain a relation involving only $W_*^{(n)}$:

$$W_*^{(n)}(\mu x_1, \dots, \mu x_n) = Z_*^{n/2}(\mu) W_*^{(n)}(x_1, \dots, x_n), \quad (15.6)$$

with

$$Z_*(\mu) = \lim_{\lambda \rightarrow \infty} Z(\lambda \mu) / Z(\lambda). \quad (15.7)$$

Equation (15.6), being valid for arbitrary μ , immediately implies that Z_* forms a representation of the dilatation semi-group:

$$Z_*(\lambda_1)Z_*(\lambda_2) = Z_*(\lambda_1\lambda_2). \quad (15.8)$$

Then, under reasonable assumptions,

$$Z_*(\lambda) = \lambda^{-2d_\phi}. \quad (15.9)$$

The fixed point correlation functions have a power-law behaviour characterized by a positive number d_ϕ , which can be called the IR *dimension* of the order parameter $\phi(x)$.

Returning to equation (15.7), we conclude that $Z(\lambda)$ also has asymptotically a power-law behaviour. Equation (15.5) then shows that the correlation functions $W^{(n)}$ have the scaling behaviour at large distances,

$$W^{(n)}(\lambda x_1, \dots, \lambda x_n) \underset{\lambda \rightarrow \infty}{\sim} \lambda^{-nd_\phi} W_*^{(n)}(x_1, \dots, x_n), \quad (15.10)$$

with a power d_ϕ that is a property of the fixed point.

The right-hand side of the equation, which determines the critical behaviour of correlation functions, depends only on the fixed point Hamiltonian. Therefore, the correlation functions corresponding to all Hamiltonians that flow after RG transformations into the same fixed point, have the same critical behaviour. This property is called *universality*. The space of Hamiltonians is thus divided into *universality classes*. Universality, beyond the mean-field approximation, relies upon the existence of IR fixed points in the space of Hamiltonians.

Applied to the two-point function, this result shows, in particular, that, if $2d_\phi < d$, the correlation length ξ diverges and, therefore, the corresponding Hamiltonians are necessarily critical. Critical Hamiltonians define in Hamiltonian space the *critical surface*, which is invariant under an RG flow. In the generic case where ξ is finite, the correlation length ξ/λ corresponding to \mathcal{H}_λ goes to zero. The Fourier components of correlation functions become momentum independent, and thus correlation functions become δ -functions in space. This trivial fixed point corresponds to $2d_\phi = d$.

15.1.2 Hamiltonian flows. Scaling operators

We now assume that the dilatation parameter can be varied continuously (on the lattice this can only be implemented approximately) and perform an additional small dilatation which leads from the scale λ to the scale $\lambda(1 + d\lambda/\lambda)$.

We perform an RG transformation, the consequence of equation (15.5), for the Hamiltonian itself. We assume that the transformation corresponds to a stationary Markov process (Section 2.1), where $\ln \lambda$ plays the role of time. We thus can write differential equations of the form,

$$\lambda \frac{d}{d\lambda} \mathcal{H}_\lambda = \mathcal{T}[\mathcal{H}_\lambda], \quad (15.11)$$

$$\lambda \frac{d}{d\lambda} \ln Z(\lambda) = 2 - d - \eta[\mathcal{H}_\lambda], \quad (15.12)$$

where we assume that \mathcal{T} is a differentiable mapping (in some topology) of the space of Hamiltonians into itself, and that η is a real continuous function, defined on the space of Hamiltonians. Moreover, both equations do not depend on λ explicitly, but only through \mathcal{H}_λ .

Explicit examples of such equations are provided by the quadratic flow equations in Refs. [61, 62], which are also discussed in Chapter 16 of Ref. [64]. Note that we have written the right-hand side of equation (15.12) in an unnatural way for later convenience.

We assume that the RG transformation (15.11) has a fixed point, necessarily solution of the equation

$$\mathcal{T}[\mathcal{H}^*] = 0, \quad (15.13)$$

where \mathcal{H}^* is the fixed point Hamiltonian, and that \mathcal{H}_λ converges towards \mathcal{H}^* for $\lambda \rightarrow \infty$.

The dimension d_ϕ of the field ϕ is then

$$d_\phi = \frac{1}{2}(d - 2 + \eta[\mathcal{H}^*]). \quad (15.14)$$

Linearized flow equations. To study the Hamiltonian flow near the fixed point, and characterize the fixed point stability, we linearize the RG equation (15.11) near the fixed point [131]. Setting $\mathcal{H}_\lambda = \mathcal{H}^* + \Delta\mathcal{H}_\lambda$, we obtain

$$\lambda \frac{d}{d\lambda} \Delta\mathcal{H}_\lambda = L^*(\Delta\mathcal{H}_\lambda), \quad (15.15)$$

where L^* is a linear operator, independent of λ , acting in Hamiltonian space.

We assume that L^* has a discrete set of eigenvalues l_i corresponding to a set of local eigenoperators \mathcal{O}_i . Then, $\Delta\mathcal{H}_\lambda$ can be expanded on the \mathcal{O}_i 's as

$$\Delta\mathcal{H}_\lambda = \sum_i h_i(\lambda) \mathcal{O}_i, \quad (15.16)$$

and the transformation (15.15) becomes

$$\lambda \frac{d}{d\lambda} h_i(\lambda) = l_i h_i(\lambda). \quad (15.17)$$

The integration then yields

$$h_i(\lambda) = \lambda^{l_i} h_i(1). \quad (15.18)$$

15.1.3 Classification of eigenvectors or scaling fields

The eigenvectors \mathcal{O}_i can be classified into four families depending on the values of the eigenvalues l_i [131].

(i) Eigenvalues with a positive real part: the corresponding eigenoperators are called *relevant*. If \mathcal{H}_λ has a component on one of these operators, this component grows with λ , and \mathcal{H}_λ moves away from the neighbourhood of \mathcal{H}^* . Operators that generate a finite correlation length (and thus $T \neq T_c$) are clearly relevant since a dilatation further decreases the effective correlation length. In Section 15.6, we calculate the corresponding eigenvalue for the ϕ^4 field theory.

(ii) Eigenvalues for which $\text{Re}(l_i) = 0$. Then, two situations can arise: either $\text{Im}(l_i)$ does not vanish and the corresponding component has a periodic behaviour (no example is met in this work) or $l_i = 0$. Eigenoperators corresponding to a vanishing eigenvalue are called *marginal*. In Section 15.5, we show that in the ϕ^4 field theory the operator $\phi^4(x)$ is marginal in four dimensions. To solve the RG equation (15.11), and determine the behaviour of the corresponding component, it is necessary to expand beyond the linear approximation. Generically, one finds

$$\lambda \frac{d}{d\lambda} h_i(\lambda) \sim B h_i^2(\lambda). \quad (15.19)$$

Depending on the sign of the constant B and the initial sign of h_i , the fixed point then is marginally unstable or stable. In the latter case, the solution takes for λ large the form

$$h_i(\lambda) \sim -1/(B \ln \lambda). \quad (15.20)$$

A marginally stable operator generally leads to a logarithmic approach to a fixed point.

An exceptional example is provided by the XY model in two dimensions ($O(2)$ symmetric non-linear σ -model), which, instead of an isolated fixed point, has a line of fixed points. The operator that corresponds to a motion along the line is obviously marginal (see Chapter 31).

(iii) Eigenvalues with negative real parts. The corresponding operators are called *irrelevant*. The effective components on these operators go to zero for large dilatations.

All these eigenvalues, which are characteristic of the fixed point, may appear in the asymptotic expansion at large distances of the correlation functions corresponding to some critical or near-critical Hamiltonian.

(iv) Finally, some operators do not affect physics. An example is provided by the operator associated to a constant multiplicative renormalization of the dynamical fields $\phi(x)$. These operators are called *redundant*. In QFT, quantum equations of motion correspond to redundant operators with vanishing eigenvalue.

Classification of fixed points. Fixed points can be classified according to their local stability properties, that is, to the number of relevant operators. This number is also the number of parameters it is necessary to fix to impose to a general Hamiltonian to be on the surface which flows into the fixed point. For a non-trivial fixed point corresponding to critical Hamiltonians, it is the co-dimension of the critical surface.

The critical domain. In the mean-field approximation, we have derived universal properties not only for the critical theory, but also for temperatures close to T_c . They can be found also in this RG framework. Indeed, let us add to a critical Hamiltonian a term proportional to a relevant operator with a very small coefficient. For small dilatations, the RG flow is hardly affected. After some large dilatation, the flow will start deviating substantially from the flow of the critical Hamiltonian. But, by then, the components of the Hamiltonian on all irrelevant operators are already small. In the case in which relevant operators induce a finite correlation length, the maximal dilatation is of the order of the ratio between the correlation length and the microscopic scale.

This argument shows that the behaviour of correlation functions as a function of amplitudes of relevant operators is universal in the limit of asymptotically small amplitudes. We call the domain of parameters in which we expect universality the *critical domain*. Note that this strategy is reminiscent of renormalization theory (Chapter 9), where it corresponds to fine tuning.

15.1.4 Explicit RG equations for correlation functions

Equation (15.3) can be rewritten (neglecting R_n) as

$$Z^{n/2}(\lambda) W_\lambda^{(n)}(x_1/\lambda, \dots, x_n/\lambda) = W^{(n)}(x_1, \dots, x_n). \quad (15.21)$$

We assume that the Hamiltonian \mathcal{H}_λ has been parametrized in terms of constants $h_i(\lambda)$,

$$\mathcal{H}_\lambda = \sum_i h_i(\lambda) \mathcal{O}_i.$$

We can then change notation and set

$$W_\lambda^{(n)}(x_1, \dots, x_n) \equiv W^{(n)}(\{h(\lambda)\}; x_1, \dots, x_n),$$

where $\{h(\lambda)\}$ stands for the set of all $h_i(\lambda)$. We then differentiate equation (15.21) with respect to λ . The right-hand side does not depend on λ and, therefore,

$$\lambda \frac{d}{d\lambda} \left[Z^{n/2}(\lambda) W^{(n)}(\{h(\lambda)\}; x_1/\lambda, \dots, x_n/\lambda) \right] = 0. \quad (15.22)$$

We introduce the differential operator

$$D_{\text{RG}} \equiv - \sum_\ell x_\ell \frac{\partial}{\partial x_\ell} - \sum_i \beta_i(\{h\}) \frac{\partial}{\partial h_i} + \frac{n}{2} (2 - d - \eta(\{h\})),$$

with

$$\beta_i(\{h\}) = -\lambda \frac{d}{d\lambda} h_i(\lambda), \quad 2 - d - \eta(\{h\}) = \lambda \frac{d}{d\lambda} \ln Z(\lambda), \quad (15.23)$$

where we have used properties of equations (15.11) and (15.12), which imply that β_i et η do not depend on λ explicitly, but only through the $h_i(\lambda)$. Equation (15.22) becomes

$$D_{\text{RG}} W^{(n)}(\{h(\lambda)\}; x_1/\lambda, \dots, x_n/\lambda) = 0.$$

Finally, since λ plays no explicit role anymore, we set $\lambda = 1$. The equation becomes a partial differential equation for correlation functions, the form of RG equations we use most frequently,

$$\left[- \sum_\ell x_\ell \frac{\partial}{\partial x_\ell} - \sum_i \beta_i(\{h\}) \frac{\partial}{\partial h_i} + \frac{n}{2} (2 - d - \eta(\{h\})) \right] W^{(n)}(\{h\}; x_1, \dots, x_n) = 0. \quad (15.24)$$

With this notation, a fixed point is defined by the common solution to all equations

$$\beta_i(h^*) = 0, \quad \forall i,$$

and equation (15.24) then implies the scaling behaviour derived more directly.

15.2 The Gaussian fixed point

In a rather general framework, we have shown that universality emerges as a consequence of the existence of RG fixed points in Hamiltonian space. Therefore, to be able to quantitatively describe the critical behaviour, one has to construct RG flows explicitly and to find their fixed points. A global analysis has never been performed. In general, one is able only to exhibit a few fixed points and study their local stability.

However, the subspace of quadratic Hamiltonians can be explored completely (free field theories in the QFT terminology), which corresponds to Gaussian Boltzmann weights.

Gaussian distributions play a special role, since they appear as asymptotic distributions in the case of a large number of weakly coupled stochastic degrees of freedom. Moreover, as we have explained in Section 14.7, the weakly perturbed Gaussian (or quasi-Gaussian) model reproduces all results of the mean-field approximation, which makes the study of Gaussian models especially interesting.

We consider a general quadratic Hamiltonian in the continuum, invariant under space translations and, for simplicity, rotations. For short range interactions, the class of interactions considered in Section 14.4, in terms of the Fourier components $\tilde{\phi}(q)$ of the field $\phi(x)$, the Hamiltonian has the form

$$\mathcal{H}_G(\phi) = \frac{1}{2} \int d^d q \tilde{\phi}(q) U(q) \tilde{\phi}(-q), \quad (15.25)$$

with

$$U(q) = \sum_{r \geq 0} \tilde{u}_r^{(2)} q^{2r}, \quad (15.26)$$

where U is a positive function with a convergent q^2 expansion. This form is equivalent to the derivative expansion

$$\mathcal{H}_G(\phi) = \frac{1}{2} \int d^d x \phi(x) U(-\nabla^2) \phi(x). \quad (15.27)$$

The simple RG transformations $\phi \mapsto \phi \sqrt{Z(\lambda)}$, $x \mapsto \lambda x$, which leads to the Hamiltonian

$$\mathcal{H}_{G,\lambda}(\phi) = \frac{1}{2} Z(\lambda) \int \lambda^d d^d x U(-\nabla^2/\lambda^2) \phi(x), \quad (15.28)$$

implies the relation (15.3), with $R^{(2)} = 0$:

$$W_\lambda^{(2)}(x) = Z^{-1}(\lambda) W^{(2)}(\lambda x).$$

In the Hamiltonian, the coefficient of the term with $2r$ derivatives becomes

$$u_r^{(2)} \mapsto u_r^{(2)}(\lambda) = Z(\lambda) \lambda^{d-2r} u_r^{(2)}(1). \quad (15.29)$$

For $\lambda \rightarrow \infty$, the leading terms are those with the smallest number of derivatives. For $u_0^{(2)} \neq 0$, the Hamiltonian is non-critical. If we choose for renormalization factor $Z(\lambda) = \lambda^{-d}$, we obtain the trivial fixed point

$$\mathcal{H}_G^*(\phi) = \frac{1}{2} u_0^{(2)} \int d^d x \phi^2(x).$$

The two-point correlation function has a δ -function limit ($d_\phi = d/2$), with a vanishing correlation length.

Critical Hamiltonian. For $u_0^{(2)} = 0$, the Hamiltonian is critical. If we choose

$$Z(\lambda) = \lambda^{-(d-2)} \Rightarrow d_\phi = \frac{1}{2}(d-2), \quad (15.30)$$

and thus (see equation (8.52))

$$x \mapsto \lambda x, \quad \phi \mapsto \phi \lambda^{-(d-2)/2}, \quad (15.31)$$

$$u_r^{(2)}(\lambda) = \lambda^{2-2r} u_r^{(2)}(1), \quad (15.32)$$

all $u_r^{(2)}$ with $r > 1$ vanish in the $\lambda \rightarrow \infty$ limit.

In the limit, we find the *Gaussian fixed point*,

$$\mathcal{H}_G^*(\phi) = \frac{1}{2} u_1^{(2)} \int d^d x (\nabla_x \phi(x))^2, \quad (15.33)$$

which generates the asymptotic behaviour (14.68) of the two-point function at T_c in the mean-field approximation. Finally, if we take $u_0^{(2)} \propto T - T_c > 0$, but small enough such that when $\lambda^2 u_0^{(2)} = O(1)$, $u_r^{(2)}(\lambda)$ for all $r > 1$ is small, we recover the universal two-point function of Ornstein-Zernicke form of the mean-field approximation (14.70), which is valid in the critical domain above T_c , and implies $\nu = \frac{1}{2}$.

15.2.1 Eigenoperators

We now study the RG behaviour of perturbations to the Gaussian fixed point (15.33), within the subspace of *even* Hamiltonians, that is, having an Ising-like \mathbb{Z}_2 symmetry. We consider perturbations $\Delta\mathcal{H}$ of the form of local operators, which can thus be expanded both in powers of ϕ and derivatives:

$$\Delta\mathcal{H}(\phi) = \sum_{n=2}^{\infty} \sum_{r=0}^{\infty} \sum_{\alpha} \mathcal{O}_{\alpha,r,n}(\phi),$$

where the operators $\mathcal{O}_{\alpha,r,n}(\phi)$ are space integrals of monomials $O_{\alpha,r,n}(\phi, x)$ in $\phi(x)$ and its derivatives (see equation (8.8)) taken at the same point, of degree n (n even) in ϕ , and with exactly r derivatives (r is also even):

$$\mathcal{O}_{\alpha,r,n}(\phi) = \int d^d x O_{\alpha,r,n}(\phi, x).$$

The index α emphasizes that a pair of values n, r correspond, in general, to several independent homogeneous polynomials.

The same RG transformations then yield (see equation (8.54))

$$[\mathcal{O}_{\alpha,r,n}(\phi)]_{\lambda} = Z^{n/2}(\lambda) \lambda^{d-r} \mathcal{O}_{\alpha,r,n}(\phi) = \lambda^{d-n(d-2)/2-r} \mathcal{O}_{\alpha,r,n}(\phi). \quad (15.34)$$

We see that these operators are eigenoperators with eigenvalues $l_{n,r} = d - n(d-2)/2 - r$.

Another way to express the same property is to take a fixed basis, and expand

$$\Delta\mathcal{H}(\phi) = \sum_{n=2}^{\infty} \sum_{r=0}^{\infty} \sum_{\alpha} h_{\alpha,r,n} \mathcal{O}_{\alpha,r,n}(\phi).$$

Then,

$$h_{\alpha,r,n}(\lambda) = \lambda^{d-n(d-2)/2-r} h_{\alpha,r,n}(1)$$

or, in differential form,

$$\lambda \frac{d}{d\lambda} h_{\alpha,r,n}(\lambda) = L^* h_{\alpha,r,n}(\lambda) = l_{r,n} h_{\alpha,r,n}(\lambda), \quad (15.35)$$

which provides an example of equation (15.17), with (*cf.*, equations (8.57) or (8.15))

$$l_{r,n} = d - \frac{1}{2}n(d-2) - r. \quad (15.36)$$

We can now classify all even operators:

(i) For $n = 2, r = 0$, then $l_{0,2} = 2$, and the operator is relevant, corresponding to a deviation from the critical temperature.

(ii) For $n = 2, r = 2$, then $l_{2,2} = 0$, and the operator is redundant, corresponding to a simple renormalization of the dynamical field.

(iii) Above dimension 4, all other operators are irrelevant: on the critical surface, the Gaussian fixed point is stable. At dimension 4, $\int \phi^4(x) d^d x$ ($n = 4, r = 0$ and thus $l_{0,4} = 4 - d$) is marginal, and logarithmic corrections are expected.

Below dimension 4, $\int \phi^4(x) d^d x$ is relevant and when the dimension decreases additional operators become relevant too. The Gaussian fixed point is IR unstable.

This analysis generalizes our analysis of corrections to the mean-field approximation. It is also equivalent to power counting in QFT as discussed in Section 8.8.1. The monomials that appear in the renormalizable field theory correspond exactly to the set of relevant or marginal operators.

Finally, among the odd operators, which break Ising symmetry, one is always relevant $n = 1, r = 0$, and thus $l_{0,1} = d/2 + 1$, and corresponds to adding a linear coupling to a magnetic field.

15.2.2 Beyond the Gaussian fixed point

The Gaussian fixed point is unstable for $d < 4$. In Section 10.1, we have shown that it is possible to define a dimensional continuation of the perturbative expansion of QFT to real or complex values of the dimension of space d . In this framework, we can consider a critical Hamiltonian in the neighbourhood of dimension 4, that is, for $\varepsilon = 4 - d$ small. If the dimensions of eigenoperators are continuous functions of the dimension of space, for $\varepsilon > 0$ and small, only the operator ϕ^4 will be relevant, with a small dimension. If initially the coefficient g of ϕ^4 is small, and if ε is small enough, there will be a range of dilatations large enough to render the amplitudes of all irrelevant operators negligible, but small enough for the coefficient $g(\lambda)$ to remain small because its evolution is slow. In the leading approximation, the flow of the Hamiltonian will then be governed by the flow of $g(\lambda)$, which eventually must become positive for the transition to be second order.

With these assumptions, the flow of $g(\lambda)$ beyond the linear approximation depends only on $g(\lambda)$ itself. It is plausible that the flow equation has an expansion of the form

$$\lambda \frac{dg(\lambda)}{d\lambda} \equiv -\beta(g(\lambda)) = (4-d)g(\lambda) - \beta_2 g^2(\lambda) + O(g^3(\lambda)). \quad (15.37)$$

In what follows, we assume that the dilatation $\lambda = 1$ corresponds to a situation where the expansion already makes sense, and thus $g(1)$ is small and positive.

The coefficient β_2 depends on the dimension of space, but at leading order can be replaced by its value at $d = 4$.

The direction of the flow depends on the sign of the function $\beta(g)$. We immediately note that the sign of β_2 plays a crucial role. Let us examine the different cases:

(i) $\beta_2 < 0$. Then, for $d < 4$, the first two terms have the same sign and $g(\lambda)$ increases until the expansion becomes meaningless, and nothing more can be concluded. By contrast, for $d > 4$, one finds a non-trivial repulsive fixed point,

$$g^* = \varepsilon/\beta_2 + O(\varepsilon^2). \quad (15.38)$$

If initially $g(1) < g^*$, $g(\lambda)$ converges towards the Gaussian fixed point. If $g(1) = g^*$, $g(\lambda) = g^*$ for all λ . If $g(1) > g^*$ $g(\lambda)$ increases, and again one cannot conclude.

(ii) $\beta_2 = 0$. In this exceptional case, the RG flow is governed by the term of order g^3 .

(iii) $\beta_2 > 0$. In this case, for $d > 4$, the Gaussian fixed point is stable. By contrast, for $d < 4$, the Gaussian fixed point is unstable, as we already know, but for ε small, another fixed point of the form (15.38) appears, $g^* \sim \varepsilon/\beta_2$, which is stable. Indeed, if initially $g(1) < g^*$, then $g(\lambda)$ increases, and if $g(1) > g^*$, $g(\lambda)$ decreases. This is an especially interesting situation and, as we show in Section 15.4, it is realized in the $O(N)$ symmetric spin model, as first noticed by Wilson and Fisher [75]. Then, universality is predicted, and all universal quantities can be calculated in a power series in the deviation ε from dimension 4.

15.3 Critical behaviour: The effective ϕ^4 field theory

The main difficulty with the general RG approach is that it requires an explicit construction of RG transformations for Hamiltonians, which have a chance to lead to fixed points. Although the general idea is to integrate over large momentum modes of the dynamical variables, its practical implementation is not straightforward [61, 62]. Here, we use a method based on a perturbative RG. The limitation of the method is that it is applicable only when there exists a fixed point that, in a sense which will become slowly clearer, is close to the Gaussian fixed point, in the spirit of the discussion of Section 15.2.2. Moreover, we write RG equations directly for correlation or vertex functions.

In Section 14.7, we have argued that, for an Ising-like system with short range ferromagnetic interactions, in the critical domain and for $d \leq 4$, the sum of the most divergent contributions, order by order in a mean-field expansion, can be reproduced by an effective local field theory whose action is given by the first relevant terms of Landau–Ginzburg–Wilson’s Hamiltonian [120] (after some renormalization of the field),

$$\mathcal{H}(\phi) \equiv \beta H = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} v \phi^2(x) + \frac{1}{4!} g \phi^4(x) \right], \quad (15.39)$$

where v and $g > 0$ are *regular* functions of the temperature for T close to T_c .

Consistently, the analysis of the stability of the Gaussian fixed point has shown that, in four dimensions, the ϕ^4 interaction becomes marginal while all other interactions remain irrelevant. If the dimensions of operators are *continuous* functions of the space dimension, the Hamiltonian (15.39) should contain all relevant operators, at least in some neighbourhood of dimension 4.

The Hamiltonian (15.39) generates a perturbative expansion of QFT type, which can be described in terms of Feynman diagrams. These have to be calculated with a cut-off of order unity, a reflection of the initial microscopic structure. Since the field theory with cut-off provides itself an example of a microscopic theory, we will call it microscopic theory. In the following calculations we choose below a Pauli–Villars’s type regularization, which amounts to adding a number of irrelevant terms to the Hamiltonian, which, therefore, are not expected to modify the large scale properties. More general regularizations are possible, but they should satisfy some general conditions. For example, the propagator can be modified (as in momentum regularization), but the Fourier transform of the inverse propagator must remain a regular function of momentum (the forces are short range).

In most of the chapter, we consider a one-component field and a system with Ising-like symmetry. In Section 15.7, we generalize to the N -vector model with $O(N)$ symmetry. Several-component models with more than one ϕ^4 coupling constant are briefly discussed in Section 16.6.

In Section 14.7, we have already noted that a convenient way to study large scale properties (or low-momentum behaviour) is to rescale all space or momentum variables and measure distances in units of the correlation length or, at the critical temperature, in some arbitrary unit much larger than the lattice spacing, and corresponding to the typical distances at which correlations are measured. After such a rescaling, the momentum cut-off becomes a large momentum Λ , analogous to the cut-off used to regularize QFT.

Therefore, we perform such a rescaling here, together with the corresponding RG transformation (15.30) on the field $\phi(x)$, associated to the Gaussian fixed point:

$$x \mapsto \Lambda x, \quad \phi(x) \mapsto \Lambda^{1-d/2} \phi(x). \quad (15.40)$$

This transformation is similar to the transformation (8.52) and can be called a Gaussian renormalization. After the rescaling, all quantities have a dimension in units of Λ , the field ϕ having its Gaussian dimension $(d/2 - 1)$. The regularized Hamiltonian (15.39) can then be written as (see also Section 8.9)

$$\mathcal{H}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla_\Lambda \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \Lambda^{4-d} \phi^4(x) \right], \quad (15.41)$$

with $r = v \Lambda^2$ and

$$\nabla_\Lambda = \nabla \left(1 - \alpha_1 \nabla^2 / \Lambda^2 + \alpha_2 \nabla^4 / \Lambda^4 + \dots \right).$$

For a value $r = r_c(g, \Lambda)$, which corresponds to the critical temperature T_c , the correlation length ξ diverges. In terms of the scale Λ , the critical domain is then defined by

$$\begin{aligned} \text{physical mass } &= \xi^{-1} \ll \Lambda \Rightarrow |r - r_c| \ll \Lambda^2, \\ \text{distances } &\gg 1/\Lambda \quad \text{or momenta} \ll \Lambda, \\ \text{magnetization } M &\equiv \langle \phi(x) \rangle \ll \Lambda^{(d/2)-1}. \end{aligned} \quad (15.42)$$

These conditions are met if Λ is identified with the cut-off of a usual QFT. However, an inspection of the action (15.41) also shows that, in contrast with what is assumed in conventional QFT, the dependence on Λ of the ϕ^4 coupling constant is fixed. For $d = 4$, we have already discussed this issue in Section 9.12. For $d < 4$, in the usual formulation of QFT, the quantity $g\Lambda^{4-d}$ is fixed, a situation which, from the viewpoint of phase transitions, corresponds to a highly non-generic small coupling (and assumes some *fine tuning*).

15.4 RG equations near four dimensions

The Hamiltonian (15.41) can now be studied by field theoretical methods. By direct calculation, we have shown in Section 10.5.6 that the perturbative expansion for the critical (or massless) theory has IR divergences for any dimension smaller than 4. If we define, by dimensional continuation, a critical theory in $d = 4 - \varepsilon$ dimension, even for ε arbitrarily small but fixed, there always exists an order $\ell \sim 4/\varepsilon$ in the perturbative expansion, at which IR divergences appear. Therefore, the idea, originally due to Wilson and Fisher [75], is to perform a double series expansion in powers of the coupling constant g and ε , with g and ε considered of the same order. Order by order in the expansion, for $r = r_c$ (*i.e.*, $T = T_c$), the critical behaviour differs from the mean-field behaviour only by powers of logarithm, and one can construct a perturbative critical theory.

To study the large cut-off limit, we then use methods developed for the construction of the renormalized massless ϕ^4 field theory in Section 9.9.2. We introduce rescaled vertex functions, called renormalized, defined by renormalization conditions at a new scale $\mu \ll \Lambda$, and functions of a renormalized coupling constant g_r (the effective coupling at scale μ). In the Fourier representation, they can be defined by

$$\begin{aligned} \tilde{\Gamma}_r^{(2)}(p; g_r, \mu, \Lambda)|_{p^2=0} &= 0, \\ \frac{\partial}{\partial p^2} \tilde{\Gamma}_r^{(2)}(p; g_r, \mu, \Lambda)|_{p^2=\mu^2} &= 1, \\ \tilde{\Gamma}_r^{(4)}(p_i = \mu\theta_i; g_r, \mu, \Lambda) &= \mu^\varepsilon g_r, \end{aligned} \quad (15.43)$$

in which θ_i is a non-vanishing numerical vector. The renormalized and microscopic vertex functions are related by the equations

$$\tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu, \Lambda) = Z^{n/2}(g, \Lambda/\mu) \tilde{\Gamma}^{(n)}(p_i; g, \Lambda), \quad (15.44)$$

where the factor Z has to be understood as the ratio between the renormalization (15.3) and the Gaussian renormalization (15.30). Renormalization theory (more precisely a slightly extended version adapted to the ε -expansion) implies that, within the double g, ε expansion, the functions $\tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu, \Lambda)$ of equation (15.44) have, at p_i , g_r , and μ fixed large cut-off limits, which are the renormalized correlation functions $\tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu)$. A detailed analysis actually shows that

$$\tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu, \Lambda) = \tilde{\Gamma}_r^{(n)}(p_i; g_r, \mu) + O(\Lambda^{-2}(\ln \Lambda)^L), \quad (15.45)$$

in which the power L increases with the order in g and ε (see equation (9.96)).

Moreover, the renormalized functions $\tilde{\Gamma}_r^{(n)}$ do not depend on the specific cut-off procedure and given the normalization conditions (15.43), are, therefore, *universal*. Since the renormalized functions $\tilde{\Gamma}_r^{(n)}$ and the initial ones $\tilde{\Gamma}^{(n)}$ are asymptotically proportional, both functions have the same small momentum or large distance behaviour. The renormalized functions thus contain the whole information about the asymptotic universal critical behaviour. Therefore, we could consider only the renormalized field theory, which indeed is the only one really useful for explicit leading order calculations (see Section 16.9). However, universality is not limited to the asymptotic critical behaviour; leading corrections have also some interesting universal properties. Moreover, renormalized quantities are not directly obtained in non-perturbative calculations and may not exist for all values of g_r (see Section 9.12). For these reasons, it is useful to study the implications of equation (15.44) also, directly in the microscopic theory.

RG equations in the microscopic parameters. Since the RG equations have been derived in Section 9.11.2, we do not repeat the derivation here. In the first part of the study, we neglect corrections subleading by powers of Λ (the assumption of perturbative RG). We return to this point in Chapter 17.

The RG equation is given by equation (9.94) and reads [72]

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right] \tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = 0. \quad (15.46)$$

Equation (15.46) is satisfied, when the cut-off is large, by the microscopic vertex functions of statistical mechanics, which are also the bare vertex functions of QFT. It is a direct consequence of the existence of a renormalized theory. We also note that equation (15.46) is a simplified form of equation (15.24), when only one operator is retained.

The form of the contributions that have been neglected has been given in equation (9.96). The RG equation (15.46) is exact for the sum of the perturbative contributions that do not vanish for Λ large, as can be verified by expanding equation (15.46) in powers of g and ε .

The RG functions β and η can be calculated in perturbation theory. At one-loop order $\eta(g)$ vanishes. The four-point vertex function has the one-loop expansion for Λ large ($\varepsilon = 4 - d$) (see equation (9.104)),

$$\tilde{\Gamma}^{(4)}(p_i) = \Lambda^\varepsilon \left[g - \frac{3g^2}{16\pi^2} (\ln \Lambda + O(1)) \right] + O(g^3, g^2\varepsilon). \quad (15.47)$$

Taking into account $\eta(g) = O(g^2)$ and inserting the expansion into equation (15.46), one obtains

$$\beta(g, \varepsilon) = -\varepsilon g + \frac{3g^2}{16\pi^2} + O(g^3, g^2\varepsilon). \quad (15.48)$$

Similarly, the two-point function at two loops is (equation (9.106)),

$$\tilde{\Gamma}^{(2)}(p) = p^2 + \frac{1}{24} \frac{1}{(8\pi^2)^2} g^2 p^2 (\ln \Lambda + O(1)) + O(g^3, g^2\varepsilon).$$

It follows (see equation (9.107)),

$$\eta(g) = \frac{1}{6} \frac{1}{(4\pi)^4} g^2 + O(g^3, g^2\varepsilon). \quad (15.49)$$

15.5 Solution of the RG equations: The ε -expansion

Equation (15.46) can be solved by the method of characteristics (see Section 9.12 for details). One introduces a dilatation parameter λ and looks for functions $g(\lambda)$ and $Z(\lambda)$ such that

$$\lambda \frac{d}{d\lambda} \left[Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \lambda\Lambda) \right] = 0. \quad (15.50)$$

Differentiating explicitly with respect to λ , one finds that equation (15.50) is compatible with equation (15.46) provided that (with $g(1) = g$, $Z(1) = 1$)

$$\ln \lambda = \int_g^{g(\lambda)} \frac{dg'}{\beta(g')}, \quad \ln Z(\lambda) = \int_1^\lambda \frac{d\sigma}{\sigma} \eta(g(\sigma)) = \int_g^{g(\lambda)} \frac{dg' \eta(g')}{\beta(g')}. \quad (15.51)$$

Equation (15.50) implies

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda) = Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \lambda\Lambda).$$

It is actually convenient to rescale Λ by a factor $1/\lambda$ and write the equation as

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda/\lambda) = Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; g(\lambda), \Lambda). \quad (15.52)$$

Equations (9.98), (9.99), and (15.52) implement approximately (because terms subleading by powers of Λ have been neglected) the general RG ideas as presented in Section 15.1: $g(\lambda)$, the effective coupling constant at scale λ , characterizes the Hamiltonian \mathcal{H}_λ , equation (9.98) is the equivalent of equation (15.11) (up to the change $\lambda \mapsto 1/\lambda$); equations (15.12) and (9.99) differ by the definition of $Z(\lambda)$, because the Gaussian renormalization (15.30) has already been performed.

Equation (15.52) shows that it is equivalent to increase Λ or to decrease λ . Therefore, to investigate the large Λ limit, we study the behaviour of the effective coupling constant $g(\lambda)$ when λ goes to zero. In what follows, we explicitly assume that the RG functions $\beta(g)$ and $\eta(g)$ are regular functions of g for $g \geq 0$.

IR fixed points. The first equation (15.51) shows that $g(\lambda)$ increases if the function β is negative, or decreases in the opposite case. Fixed points correspond to zeros of the β -function which, therefore, play an essential role in the analysis of the critical behaviour. Those where the β -function has a negative slope are IR repulsive: the effective coupling moves away from such zeros, except if the initial coupling has an exact fixed point value. Conversely, those where the slope is positive are IR attractive.

We now assume that initially g is in the range where the behaviour of the β -function is still governed by the explicit expression (15.48).

Then, above four dimensions, that is $\varepsilon < 0$, $g(\lambda)$ decreases, approaching the origin $g = 0$. We recover that the Gaussian fixed point is IR stable.

Below four dimensions, if initially g is small enough, $g(\lambda)$ first increases, a behaviour reflecting the instability of the Gaussian fixed point.

However, for ε infinitesimal, $\beta(g)$ has another zero,

$$\beta(g^*) = 0, \quad \text{for } g^* = 16\pi^2\varepsilon/3 + O(\varepsilon^2), \quad (15.53)$$

with a positive slope,

$$\omega \equiv \beta'(g^*) = \varepsilon + O(\varepsilon^2) > 0. \quad (15.54)$$

This agrees with one of the possibilities envisaged in Section 15.2.2. Then, the first equation (15.51) shows that $g(\lambda)$ has g^* as an asymptotic limit. Linearizing the equation near the fixed point, one finds

$$\int_g^{g(\lambda)} \frac{dg'}{\omega(g' - g^*)} \sim \ln \lambda, \quad (15.55)$$

or in integrated form,

$$|g(\lambda) - g^*| \underset{\lambda \rightarrow 0}{\propto} \lambda^\omega. \quad (15.56)$$

Below dimension 4, at least for ε infinitesimal, this non-Gaussian fixed point is IR stable. In dimension 4, this IR fixed point merges with the Gaussian fixed point, and the eigenvalue ω vanishes, indicating the appearance of the marginal operator already identified in the analysis of the Gaussian fixed point in Section 15.2.

The behaviour of correlation functions. We now assume that $\tilde{\Gamma}^{(n)}(g^*)$ and $\eta(g^*)$ are finite, conditions that are satisfied within the framework of the ε -expansion. The second equation (15.51) determines the behaviour of $Z(\lambda)$ for λ small. The integral in the left-hand side is dominated by small values of σ . It follows that

$$\ln Z(\lambda) \underset{\lambda \rightarrow 0}{\sim} \eta \ln \lambda, \quad (15.57)$$

where we have set

$$\eta = \eta(g^*).$$

Equation (15.52) then determines the behaviour of $\tilde{\Gamma}^{(n)}(p_i; g, \Lambda)$ for Λ large:

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda/\lambda) \propto \lambda^{-n\eta/2} \tilde{\Gamma}^{(n)}(p_i; g^*, \Lambda). \quad (15.58)$$

Moreover, simple dimensional considerations (which are directly related to the Gaussian renormalization) imply that

$$\tilde{\Gamma}^{(n)}(p_i; g, \Lambda/\lambda) = \lambda^{-d+(n/2)(d-2)} \tilde{\Gamma}^{(n)}(\lambda p_i; g, \Lambda). \quad (15.59)$$

Combining this equation with equation (15.58), one obtains

$$\tilde{\Gamma}^{(n)}(\lambda p_i; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{d-(n/2)(d-2+\eta)} \tilde{\Gamma}^{(n)}(p_i; g^*, \Lambda). \quad (15.60)$$

After Legendre and Fourier transformations, taking into account the δ -function of momentum conservation, one infers for the connected correlation functions,

$$\tilde{W}^{(n)}(\lambda p_i; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{d-(n/2)(d+2-\eta)} \tilde{W}^{(n)}(p_i; g^*, \Lambda), \quad (15.61)$$

$$W^{(n)}(x_i/\lambda; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{(n/2)(d-2+\eta)} W^{(n)}(x_i; g^*, \Lambda). \quad (15.62)$$

These equations show that the critical vertex and connected correlation functions have a power-law behaviour for small momenta, independent of the initial value of the ϕ^4 coupling constant g , at least if g initially is small enough for perturbation theory to be applicable, or if the β -function has no additional zero.

For $n = 2$, equation (15.62) implies,

$$W^{(2)}(x) \underset{|x| \rightarrow \infty}{\propto} 1/|x|^{d-2+\eta} \quad (15.63)$$

Therefore, the definition (15.57) coincides with the usual definition of the critical exponent η . The spectral representation of the two-point function (Section 6.6) implies $\eta > 0$. Since in perturbation theory, the first contribution to the field renormalization $Z(g, \Lambda/\mu)$ arises at order g^2 , $\eta(g)$ is of order g^2 , and η of order ε^2 .

A two-loop calculation (equation (15.49)) yields

$$\eta = \frac{\varepsilon^2}{54} + O(\varepsilon^3). \quad (15.64)$$

A semi-quantitative prediction is that η is numerically small in three dimensions.

Finally, equation (15.60) has the following interpretation: the field $\phi(x)$, which had at the Gaussian fixed point the Gaussian dimension $(d - 2)/2$, has now acquired the ‘anomalous’ or IR dimension (equation (15.14))

$$d_\phi = \frac{1}{2}(d - 2 + \eta).$$

Universality. These results call for a few comments. Within the framework of the ε -expansion, we have shown that all correlation functions have, for $d < 4$, a large distance behaviour different from the one predicted by the mean-field approximation. In addition, the critical behaviour does not depend on the initial value of the ϕ^4 coupling constant g . At least for ε small, we can hope that the analysis of leading IR singularities of Section 14.4 remains valid, and thus it does not depend on any other coupling either (this point will be further discussed in Section 17.4). Therefore, the critical behaviour is *universal*, although less universal than in the mean-field approximation, since it depends on a small number of general properties of the system under consideration, like symmetry or dimension of space.

Moreover, the correlation functions obtained by neglecting, in perturbation theory and within the ε -expansion, power-law corrections when the cut-off is large, and which satisfy exactly RG equations (15.46), define implicitly a one parameter family of critical Hamiltonians, which correspond to an RG trajectory which goes from the neighbourhood of the Gaussian fixed point $g = 0$, which is IR unstable below four dimensions, to a non-trivial stable fixed point g^* .

Finally, the consistency of this analysis, based on the ε -expansion, relies on the following observation. The IR divergences found in the perturbative expansion at fixed dimension result from an expansion around an IR repulsive fixed point. The ε -expansion makes an exchange of limits, and thus a perturbative expansion around an attractive IR fixed point possible.

15.6 Critical correlation functions with $\phi^2(x)$ insertions

In Chapter 16, we study the whole critical domain (15.42).

Following a method explained in Section 9.10, correlation functions for $T \neq T_c$ are expanded in powers of *critical* correlation functions with $\phi^2(x)$ insertions.

Here, the monomial $\phi^2(x)$ has a direct physics interpretation: it is the most singular contribution (*i.e.* the most relevant) to the energy density (15.39).

Therefore, we now discuss the large distance properties of the critical mixed correlation functions of the order parameter ϕ and the energy density $\frac{1}{2}\phi^2$. We denote by $\tilde{\Gamma}^{(l,n)}(q_1, \dots, q_l; p_1, \dots, p_n; g, \Lambda)$ the Fourier transforms of vertex functions of $n \phi(x)$ fields and $l \frac{1}{2}\phi^2(x)$ monomials, (with $(l + n) \geq 2$).

15.6.1 RG equations

Renormalization theory implies that one can define renormalized vertex functions $\tilde{\Gamma}_r^{(l,n)}$ related to the microscopic functions by (see equations (9.72) or (11.6))

$$\begin{aligned} \tilde{\Gamma}_r^{(l,n)}(q_i; p_j; g_r, \mu) \\ = \lim_{\substack{\Lambda \rightarrow \infty \\ \text{at } \mu, g_r \text{ fixed}}} Z^{n/2-l} Z_2 \left[\tilde{\Gamma}^{(l,n)}(q_i; p_j; g, \Lambda) - \delta_{n0} \delta_{l2} \Lambda^{-\varepsilon} C(g, \Lambda/\mu) \right], \end{aligned} \quad (15.65)$$

where $Z_2(g, \Lambda/\mu)$ and $C(g, \Lambda/\mu)$ are two new renormalization constants.

They can be fixed by imposing, in addition to the conditions (15.43), the conditions

$$\tilde{\Gamma}_r^{(1,2)}(q; p_1, p_2; g_r, \mu) \Big|_{\substack{p_1^2 = p_2^2 = \mu^2, \\ p_1 \cdot p_2 = -\frac{1}{3}\mu^2}} = 1, \quad \tilde{\Gamma}_r^{(2,0)}(q, -q; g_r, \mu) \Big|_{q^2 = \frac{4}{3}\mu^2} = 0. \quad (15.66)$$

Differentiating with respect to Λ at g_r and μ fixed, as in Section 15.4, and using chain rule, one obtains the RG equations [72],

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \eta(g) - l \eta_2(g) \right] \tilde{\Gamma}^{(l,n)} = \delta_{n0} \delta_{l2} \Lambda^{-\varepsilon} B(g). \quad (15.67)$$

In addition to β and η , two new RG functions, $\eta_2(g)$ and $B(g)$, appear:

$$\eta_2(g) = -\Lambda \frac{\partial}{\partial \Lambda} \Big|_{g_r, \mu} \ln [Z_2(g, \Lambda/\mu) / Z(g, \Lambda/\mu)], \quad (15.68)$$

$$B(g) = \left[\Lambda \frac{\partial}{\partial \Lambda} \Big|_{g_r, \mu} - 2\eta_2(g) - \varepsilon \right] C(g, \Lambda/\mu). \quad (15.69)$$

Again, these two RG functions, because they are calculable in terms of correlation functions which do not depend on μ , cannot depend on the ratio Λ/μ .

Note that for $n = 0, l = 2$, the RG equation (15.67) is not homogeneous. This is a consequence of the non-multiplicative character of the renormalization in this case. Multiple insertions of operators of higher dimension like $\phi^4(x)$ lead to even more complicated RG equations. This question has been discussed in Chapter 11, and we want only to warn the reader here against a too naive application of RG ideas.

In the homogeneous case, equation (15.67) can be solved by the method of characteristics explained in Section 15.5, in exactly the same way as equation (15.46). A new scale-dependent function $\zeta_2(\lambda)$ is associated with the RG function $\eta_2(g)$. We set

$$\lambda \frac{d}{d\lambda} \left[Z^{-n/2}(\lambda) \zeta_2^{-l}(\lambda) \tilde{\Gamma}^{(l,n)}(q_i; p_j; g(\lambda), \lambda\Lambda) \right] = 0,$$

and $\zeta_2(\lambda)$ satisfies

$$\lambda \frac{d}{d\lambda} \ln \zeta_2(\lambda) = \eta_2[g(\lambda)].$$

It follows from the arguments of Section 15.5 that the critical behaviour of $\tilde{\Gamma}^{(l,n)}$ is governed by the IR fixed point g^* .

To relate the RG function $\eta_2(g)$ to standard exponents, we introduce the function

$$\nu(g) = [\eta_2(g) + 2]^{-1}. \quad (15.70)$$

For λ small, ζ_2 behaves like

$$\ln \zeta_2 \sim (1/\nu - 2) \ln \lambda, \quad \text{with} \quad \nu = \nu(g^*).$$

Combining again the solution of equation (15.67) with simple dimensional analysis, one obtains the behaviour

$$\tilde{\Gamma}^{(l,n)}(\lambda q_i; \lambda p_j; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{d-n(d-2+\eta)/2-l/\nu} \tilde{\Gamma}^{(l,n)}(q_i; p_j; g^*, \Lambda). \quad (15.71)$$

For connected correlation functions, this translates into the relations

$$\tilde{W}^{(l,n)}(\lambda q_i; \lambda p_j; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{d-(n/2)(d+2-\eta)-l/\nu} \tilde{W}^{(l,n)}(q_i; p_j; g^*, \Lambda), \quad (15.72)$$

$$W^{(l,n)}(y_i/\lambda; x_j/\lambda; g, \Lambda) \underset{\lambda \rightarrow 0}{\propto} \lambda^{(n/2)(d-2+\eta)+l(d-1/\nu)} W^{(l,n)}(y_i; x_j; g^*, \Lambda). \quad (15.73)$$

Using equations (15.68) and (15.70), it is straightforward to calculate $\nu(g)$ at one-loop order. One finds

$$2\nu(g) = 1 + \frac{g}{32\pi^2} + O(g^2), \quad (15.74)$$

and then, setting $g = g^*$ (equation (15.53)), the exponent ν at first order in ε :

$$2\nu = 1 + \frac{1}{6}\varepsilon + O(\varepsilon^2).$$

The $\langle \phi^2 \phi^2 \rangle$ correlation function. The ϕ^2 (energy density) two-point function $\tilde{\Gamma}^{(2,0)}$ satisfies the inhomogeneous RG equation (15.67). To solve it, we first look for a particular solution of the form $\Lambda^{-\varepsilon} C_2(g)$. Then,

$$\beta(g) C'_2(g) - [\varepsilon + 2\eta_2(g)] C_2(g) = B(g). \quad (15.75)$$

The solution is uniquely determined by imposing its *regularity* at $g = g^*$.

The general solution of equation (15.67) then is the sum of this particular solution and the general solution of the homogeneous equation, which has a behaviour given by equation (15.71):

$$\tilde{\Gamma}^{(2,0)}(q; g, \Lambda) - \Lambda^{-\varepsilon} C_2(g) \underset{q \rightarrow 0}{\sim} K \Lambda^{2/\nu-4} q^{d-2/\nu}. \quad (15.76)$$

Note that the regular contribution depends on g but not the constant K .

After Fourier transformation, the singular term has the scaling behaviour (15.73) and the constant gives a δ -function.

The dimension of the ϕ^2 operator. The scaling behaviour (15.73) determines the dimension of the operator ϕ^2 at the IR fixed point. Quite generally, the connected n -point correlation function with insertions of l scaling fields \mathcal{O}_i of dimensions d_i has, in real space, the scaling behaviour

$$\left\langle \prod_i \mathcal{O}_i(y_i/\lambda) \prod_j \phi(x_j/\lambda) \right\rangle_{\lambda \rightarrow 0} \propto \lambda^D, \quad \text{with } D = n d_\phi + \sum_i d_i.$$

Equation (15.73) then implies $d_{\phi^2} = d - 1/\nu$.

Remark. Real physics corresponds to integer values of ε , $\varepsilon = 1, 2$. Although one can prove the validity of RG results only within the framework of the ε -expansion, we will eventually assume that their validity extends beyond an infinitesimal neighbourhood of dimension 4. The comparison between RG results derived from the ε expansion and experimental, or numerical, data thus provides a crucial test of the theory (see Section 16.9).

The fixed point Hamiltonian. Field theory considerations indicate that the explicit determination of a fixed point Hamiltonian \mathcal{H}^* is in general not easy, except for the Gaussian fixed point. Since fixed point correlation functions have been calculated, they define implicitly a class of fixed point Hamiltonians. It has also been proposed to define \mathcal{H}^* as a limit of a sequence of Hamiltonians \mathcal{H}_k whose correlation functions $W_k^{(n)}$ converge towards fixed point correlation functions $W_*^{(n)}$ in the following way:

$$\left| W_k^{(n)}(\lambda q_i) - W_*^{(n)}(\lambda q_i) \right| = O(\lambda^{2k}).$$

In Sections 15.3–15.5, we have constructed \mathcal{H}_1 .

15.6.2 Renormalization theory and critical phenomena

Let us comment about the significance of these results for QFT in the context of renormalization theory. From the preceding analysis, we infer that the existence of a renormalizable field theory, beyond perturbation theory, relies on two properties:

- (i) The existence of a continuous phase transition.
- (ii) The existence of an IR RG fixed point that ensures that the large distance physics is short-distance insensitive, and can be described by a renormalized QFT.

However, an inspection of the action (15.39) shows that, in contrast with conventional renormalization theory, the ϕ^4 coupling constant has a dependence in Λ given *a priori*. For $d < 4$, the ϕ^4 coupling is very large in terms of the scale relevant for the critical domain. By contrast, in QFT, in the usual formulation of renormalization theory, the *microscopic* coupling constant is considered as an adjustable parameter: a super-renormalizable QFT corresponds to a Hamiltonian maintained close to the Gaussian fixed point by adjusting the coefficients of all relevant and marginal operators. The resulting theory is thus less generic.

These coefficients then introduce a new scale μ^{-1} in the theory, much larger than the microscopic scale, in such a way that it is possible to define, in addition to the universal large distance physics, a short distance or large momentum physics with $\mu \ll |p_i| \ll \Lambda$.

By contrast, if one assumes that parameters in particle physics theories are *generic*, that is, that coupling constants at the cut-off scale (which can be identified with the scale of some new physics) are numbers of order 1, we are driven back into the statistical framework. This is also the viewpoint of effective field theories (Section 8.8).

Finally, we have performed the RG analysis of the ϕ^4 QFT in four dimensions, in the context of particle physics, in Section 9.12. In the domain of attraction of the IR fixed point $g = 0$, the renormalized coupling constant of the ϕ^4 field theory goes to 0 logarithmically when the cut-off becomes infinite, leading to the so-called *triviality problem*.

15.7 The $O(N)$ -symmetric $(\phi^2)^2$ field theory

We now briefly indicate how the considerations of the chapter generalize to models with an $O(N)$ symmetry (see Section (14.75)). We consider a field ϕ with N components ϕ_i , and the $O(N)$ symmetric Hamiltonian

$$\mathcal{H}(\phi) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \Lambda^{4-d} (\phi^2(x))^2 \right] d^d x. \quad (15.77)$$

The vertex functions have now group indices, and satisfy WT identities of the kind described in Section 13.2. As a consequence, for example,

$$\begin{aligned} \tilde{\Gamma}_{ij}^{(2)}(p) &= \delta_{ij} \tilde{\Gamma}^{(2)}(p) \\ \tilde{\Gamma}_{ijkl}^{(4)}(p_1, p_2, p_3, p_4) &= \delta_{ij} \delta_{kl} \tilde{\Gamma}^{(4)}(p_1, p_2, p_3, p_4) + \delta_{ik} \delta_{jl} \tilde{\Gamma}^{(4)}(p_1, p_3, p_2, p_4) \\ &\quad + \delta_{il} \delta_{kj} \tilde{\Gamma}^{(4)}(p_1, p_4, p_3, p_2), \end{aligned}$$

where $\tilde{\Gamma}^{(4)}(p_1, p_2, p_3, p_4)$ is symmetric in the exchanges $p_1 \leftrightarrow p_2$ and $p_3 \leftrightarrow p_4$.

With these definitions, the two first renormalization conditions (15.43) are unchanged, and the third one becomes

$$\tilde{\Gamma}_r^{(4)}(p_i = \mu \theta_i; g_r, \mu, \Lambda) = \frac{1}{3} \mu^\varepsilon g_r. \quad (15.78)$$

The derivation of the RG equations is identical, and one finds,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right] \tilde{\Gamma}_{i_1, \dots, i_n}^{(n)}(p_1, \dots, p_n; g, \Lambda) = 0. \quad (15.79)$$

The RG equations with $\frac{1}{2} \phi^2$ insertions are identical to equations (15.67). The RG functions at leading order are (equations (10.89–10.91)),

$$\beta(g) = -\varepsilon g + \frac{N+8}{48\pi^2} g^2 + O(g^3), \quad (15.80)$$

$$\eta(g) = \frac{(N+2)}{72(8\pi^2)^2} g^2 + O(g^3), \quad (15.81)$$

$$\eta_2(g) = -\frac{1}{48\pi^2} (N+2)g + O(g^2). \quad (15.82)$$

One finds again an IR fixed point $g^* \sim 48\pi^2 \varepsilon / (N+8)$, with $\omega = \beta'(g^*) \sim \varepsilon > 0$. All the scaling relations follow, but with different exponents,

$$\eta = \frac{N+2}{2(N+8)^2} \varepsilon^2 + O(\varepsilon^3), \quad \eta_2 = -\frac{N+2}{N+8} \varepsilon + O(\varepsilon^2) \Rightarrow \nu = \frac{1}{2} + \frac{(N+2)}{4(N+8)} \varepsilon + O(\varepsilon^2).$$

15.8 Statistical properties of long self-repelling chains

Remarkably enough, the statistical properties of long polymer chains

can be derived from the critical behaviour of the $O(N)$ symmetric $(\phi^2)^2$ field theory in the ‘non-physical’ $N = 0$ limit [136]. (For a comprehensive experimental and theoretical discussion of polymer physics, see Ref. [137].) To describe the asymptotic statistical properties of long polymers, we consider a model of chains in the continuum with repulsive contact self-interaction. We will follow the derivation of Ref. [138] (see also Ref. [139]).

We denote by u the position along the chain, $\mathbf{r}(u)$ the position in space of the point on the chain of parameter u . We characterize the chain by the probability distribution for a chain $\mathbf{r}(u)$ of total length s ,

$$[d\rho(\mathbf{r}(u))] = [dr(u)] \exp \left\{ - \left[\frac{1}{4} \int_0^s \dot{\mathbf{r}}^2(u) du + \frac{g}{6} \int_0^s du_1 du_2 \delta^{(d)}(\mathbf{r}(u_1) - \mathbf{r}(u_2)) \right] \right\}. \quad (15.83)$$

The special limit $g = 0$ corresponds to a Brownian chain, or a Gaussian random walk (see also Section 34.5). The self-avoiding random walk on a lattice is a discretized form of the model, and provides a regularization: at a microscopic scale Λ^{-1} , the chain becomes much stiffer than what is implied by expression (15.83). Finally, note that any short-range potential would yield, at leading order for long chains, the same results as the δ -function interaction, as one can verify by repeating, in the more general case, the derivation of an equivalent ϕ^4 like field theory, and using the analysis of corrections to scaling as in the case of critical phenomena.

15.8.1 A generating function

Various characteristic properties of the chain can be derived from the two-point function

$$G^{(2)}(\mathbf{k}, s) = \langle e^{i\mathbf{k} \cdot (\mathbf{r}(s) - \mathbf{r}(0))} \rangle, \quad (15.84)$$

where brackets mean average with respect to the distribution (15.83). Indeed, the expansion of $G^{(2)}(\mathbf{k}, s)$ in powers of \mathbf{k} yields the successive moments of the distribution of the origin-to-end positions:

$$G^{(2)}(\mathbf{k}, s) = 1 - \frac{1}{2!} \frac{\mathbf{k}^2}{d} \langle (\mathbf{r}(s) - \mathbf{r}(0))^2 \rangle + \frac{(\mathbf{k}^2)^2}{4!} \frac{1}{d(d+2)} \langle (\mathbf{r}(s) - \mathbf{r}(0))^4 \rangle + \dots, \quad (15.85)$$

where the rotation symmetry of the distribution (15.83) has been used.

Actually, it is easier to calculate the Laplace transform of $G^{(2)}(\mathbf{k}, s)$ with respect to s ,

$$\tilde{Z}^{(2)}(\mathbf{k}, t) = \int_0^\infty e^{-st} G^{(2)}(\mathbf{k}, s) ds. \quad (15.86)$$

The function $G^{(2)}(\mathbf{k}, s)$ is then recovered by inverting the Laplace transformation,

$$G^{(2)}(\mathbf{k}, s) = \frac{1}{2i\pi} \oint_C e^{st} \tilde{Z}^{(2)}(\mathbf{k}, t) dt, \quad (15.87)$$

in which the contour C is parallel to the imaginary t axis at the right of all singularities of $\tilde{Z}^{(2)}$. Since only the expansion of $G^{(2)}$ in powers of \mathbf{k} (equation (15.85)) is useful, \mathbf{k} has to be considered as infinitesimal. Thus, only the location of singularities in the limit $\mathbf{k} \rightarrow 0$ plays a role.

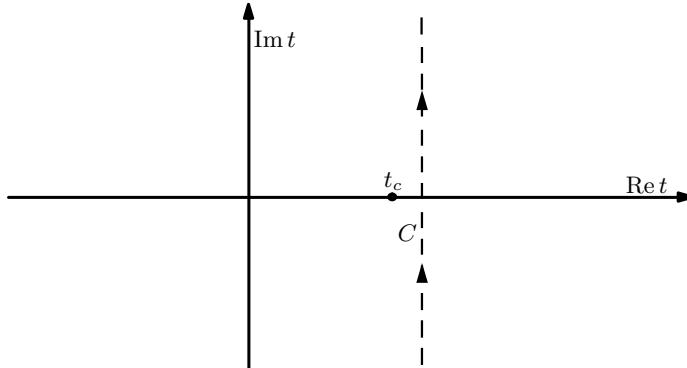


Fig. 15.1 Contour C in the t plane

We denote by t_c the first singularity met when the contour is displaced to the left (Fig. 15.1). On C , the real part $\text{Re}(t)$ of t thus satisfies

$$\text{Re}(t) > \text{Re}(t_c).$$

One verifies that the behaviour of $\tilde{Z}^{(2)}$ near the singularity governs the large positive s behaviour of the integral (15.87), and thus of the moments of $\mathbf{r}(s) - \mathbf{r}(0)$.

15.8.2 Some exact results. Flory's approximation

The Gaussian random walk. In expression (15.84), we can substitute

$$\mathbf{r}(s) - \mathbf{r}(0) = \int_0^s \dot{\mathbf{r}}(u) du.$$

For vanishing self-repulsion, one obtains

$$G^{(2)}(\mathbf{k}, s) = \int [d\mathbf{r}(u)] \exp \left\{ - \int_0^s du \left[\frac{1}{4} \dot{\mathbf{r}}^2(u) - i \mathbf{k} \cdot \dot{\mathbf{r}}(u) \right] \right\}.$$

Shifting $\dot{\mathbf{r}}$, one finds

$$G^{(2)}(\mathbf{k}, s) = e^{-s\mathbf{k}^2}, \quad (15.88)$$

which implies the well-known scaling laws of the Gaussian random walk:

$$\langle (\mathbf{r}(s) - \mathbf{r}(0))^2 \rangle \sim s^n. \quad (15.89)$$

The Laplace transform (15.86) takes the free field form,

$$\tilde{Z}^{(2)}(\mathbf{k}, t) = (t + \mathbf{k}^2)^{-1}. \quad (15.90)$$

The one-dimensional chain. If the chain self-repulsion plays a role, it must increase the average spatial extension of the chain for s large, compared to the Gaussian case. This is obvious in the one-dimensional example. The chain is then completely stretched:

$$r(s) = \pm s, \quad (15.91)$$

and, therefore,

$$G^{(2)}(k, s) = \cos ks, \quad \tilde{Z}^{(2)}(k, t) = \frac{t}{t^2 + k^2}. \quad (15.92)$$

The one-dimensional example shows that the self-repulsion can change the statistical properties of the chain. It also provides an upper bound on the moments. RG arguments will show that

$$\langle (\mathbf{r}(s) - \mathbf{r}(0))^{2n} \rangle \sim s^{2\nu n}, \quad \text{for } s \rightarrow \infty,$$

in which the exponent ν is bounded from below by $1/2$, the value for the Gaussian chain, and from above by 1 , the value for the stretched chain.

The upper-critical dimension. Note that the self-avoiding condition becomes weaker when the dimension d of embedding space increases. Actually, a simple argument suggests, as in the case of the critical behaviour, the existence of an upper-critical dimension. For a very long chain, one can consider the effect of self-avoiding as the influence of one chain onto another. A Brownian chain has Hausdorff dimension 2. Therefore, above dimension 4 ($2 + 2$) two chains no longer see each other, self-avoiding should no longer play a role, and the chain should have the statistical properties of a Brownian chain.

Beyond the Gaussian model: Flory's approximation. Before solving the problem by more systematic methods, we first describe a simple approximate solution based on an energy balance argument. For s large, the Gaussian term should scale as the interaction:

$$\left\langle \int_0^s \dot{\mathbf{r}}^2(u) du \right\rangle \sim s^{2\nu-1}, \quad (15.93)$$

$$\int du_1 du_2 \delta^{(d)}(\mathbf{r}(u_1) - \mathbf{r}(u_2)) \sim s^{2-d\nu}. \quad (15.94)$$

Equating the two exponents, one obtains the equation,

$$2\nu - 1 = 2 - d\nu \Rightarrow \nu = \frac{3}{d+2}, \quad (15.95)$$

valid for $1 \leq d \leq 4$. For $d = 1$, Flory's approximation yields the exact value $\nu = 1$. It predicts that, for $d \geq 4$, the interaction can never balance the Gaussian term, since ν is bounded by $1/2$. More surprisingly, for $d = 2$ it yields $\nu = 3/4$ which is also known, from conformal field theory arguments, to be the exact value.

Flory's approximation relies on intuitive arguments and it is difficult to see how it can be improved. Actually, the values it yields for the exponent ν are so good that one may wonder whether the result is not exact. We discuss generic dimensions in the following sections, and show that it is only an approximation. For this purpose, it is convenient to completely reformulate the model based on the distribution (15.83).

15.8.3 Equivalence with the $(\phi^2)^2$ field theory in the $N = 0$ limit

We first transform the interaction term in equation (15.83) by expressing it as a Gaussian integral over an auxiliary field $\sigma(\mathbf{r})$. For a given chain $\mathbf{r}(u)$, the following identity holds:

$$\begin{aligned} & \int [d\sigma(\mathbf{r})] \exp \left[\frac{3}{2g} \int d^d r \sigma^2(r) - \int du \sigma(\mathbf{r}(u)) \right] \\ &= \exp \left[-\frac{g}{6} \int du_1 du_2 \delta^{(d)}(\mathbf{r}(u_1) - \mathbf{r}(u_2)) \right], \end{aligned} \quad (15.96)$$

where the σ -field is imaginary, and we have used

$$\int du \sigma(\mathbf{r}(u)) = \int du \int d^d r \sigma(r) \delta^{(d)}(\mathbf{r}(u) - \mathbf{r}) = \int d^d r \sigma(r) \int du \delta^{(d)}(\mathbf{r}(u) - \mathbf{r}).$$

Substituting the field integral into expression (15.83), we identify the $\mathbf{r}(u)$ integration measure with the integrand in the path integral representation of the statistical operator in Euclidean time u of a d -dimensional quantum system with potential $\sigma(\mathbf{r})$ (see Chapter 2). The two-point function $\tilde{Z}^{(2)}$ can thus be written as, in bra–ket quantum notation,

$$\begin{aligned}\tilde{Z}^{(2)}(\mathbf{k}, t) &= \int [d\sigma(\mathbf{r})] \exp \left[\frac{3}{2g} \int d^d r \sigma^2(r) \right] \int_0^\infty e^{-ts} ds \\ &\times \int d^d r d^d r' e^{i\mathbf{k}(\mathbf{r}-\mathbf{r}')} \langle \mathbf{r}' | e^{-sH} | \mathbf{r} \rangle,\end{aligned}\quad (15.97)$$

in which H is the quantum Hamiltonian,

$$H = -\nabla_{\mathbf{r}}^2 + \sigma(\mathbf{r}).\quad (15.98)$$

After integration over s , the expression becomes

$$\begin{aligned}\tilde{Z}^{(2)}(\mathbf{k}, t) &= \int [d\sigma(\mathbf{r})] \exp \left[\frac{3}{2g} \int d^d r \sigma^2(r) \right] \int d^d r d^d r' e^{i\mathbf{k}(\mathbf{r}-\mathbf{r}')} \\ &\times \langle \mathbf{r}' | (-\nabla^2 + t + \sigma)^{-1} | \mathbf{r} \rangle.\end{aligned}\quad (15.99)$$

We now apply a variant of the so-called *replica trick* (see also Section A14.5). We introduce an N -component scalar field ϕ . Then,

$$\begin{aligned}\lim_{N \rightarrow 0} \int [d\phi(\mathbf{r})] \phi_1(\mathbf{r}) \phi_1(\mathbf{r}') \exp \left\{ -\frac{1}{2} \int d^d r \left[(\nabla \phi(r))^2 + (t + \sigma(r)) \phi^2(r) \right] \right\} \\ \propto \lim_{N \rightarrow 0} [\det(-\nabla^2 + t + \sigma)]^{-N/2} \langle \mathbf{r}' | (-\nabla^2 + t + \sigma)^{-1} | \mathbf{r} \rangle \\ = \langle \mathbf{r}' | (-\nabla^2 + t + \sigma)^{-1} | \mathbf{r} \rangle.\end{aligned}\quad (15.100)$$

Substituting the identity into expression (15.99), one obtains

$$\begin{aligned}\tilde{Z}^{(2)}(\mathbf{k}, t) &= \int [d\phi] \tilde{\phi}_1(\mathbf{k}) \tilde{\phi}_1(-\mathbf{k}) \int [d\sigma(r)] \\ &\times \exp \left[\frac{1}{2} \int d^d r \left(3\sigma^2(\mathbf{r})/g - \sigma(\mathbf{r})\phi^2(\mathbf{r}) - (\nabla \phi(\mathbf{r}))^2 - t\phi^2(\mathbf{r}) \right) \right].\end{aligned}\quad (15.101)$$

Then, the Gaussian integration over σ can be performed, and yields

$$\begin{aligned}\tilde{Z}^{(2)}(\mathbf{k}, t) &= \int [d\phi] \tilde{\phi}_1(\mathbf{k}) \tilde{\phi}_1(-\mathbf{k}) \exp [-\mathcal{S}(\phi)], \\ \mathcal{S}(\phi) &= \int d^d r \left[\frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} t\phi^2 + \frac{g}{4!} (\phi^2)^2 \right].\end{aligned}\quad (15.102)$$

This is a most remarkable result: the statistical properties of polymers are related to the properties of the $O(N)$ symmetric $(\phi^2)^2$ field theory in the non-physical $N = 0$ limit, as first pointed out by de Gennes.

15.8.4 RG approach to self-avoiding walk (SAW) and statistical properties of polymers
 The large scale statistical properties of the chain are thus related to the singularities of correlation functions of the $(\phi^2)^2$ field theory in the massless or critical limit. The short distance stiffness of the chain provides an ultraviolet (UV) cut-off Λ for the theory. Near t_c , in which t_c is the critical temperature or critical bare mass, the two-point function $\tilde{Z}^{(2)}$ has the scaling behaviour:

$$\tilde{Z}^{(2)}(\mathbf{k}, t) \sim (t - t_c)^{-\gamma} f [k(t - t_c)^{-\nu}], \quad (15.103)$$

in which ν and γ are the $N = 0$ limits of the critical exponents of the $(\phi^2)^2$ field theory.

Inserting the form (15.103) into the integral (15.87), one obtains the scaling behaviour of $G^{(2)}(s, k)$ for s large:

$$G^{(2)}(s, k) \sim e^{st_c} s^{\gamma-1} g(|\mathbf{k}| s^\nu). \quad (15.104)$$

Expanding in powers of \mathbf{k}^2 , one finds a first term

$$\langle 1 \rangle \sim e^{st_c} s^{\gamma-1} g(0),$$

instead of 1, which shows that the distribution is incorrectly normalized. The first term is actually the ratio of the number of configurations of the SAW to the number of configurations of the Gaussian random walk. Therefore, the properly normalized average $\langle \exp[i\mathbf{k}(r(s) - r(0))] \rangle$ has a scaling behaviour entirely characterized by the exponent ν . For example,

$$\langle (\mathbf{r}(s) - \mathbf{r}(0))^2 \rangle \sim s^{2\nu}. \quad (15.105)$$

The exponent can be obtained from the ε -expansion:

$$\nu = \frac{1}{2} \left(1 + \frac{1}{8}\varepsilon + \frac{15}{256}\varepsilon^2 \right) + O(\varepsilon^3). \quad (15.106)$$

By contrast, Flory's formula (15.95) predicts at the same order:

$$\nu(\text{Flory}) = \frac{1}{2} \left(1 + \frac{1}{6}\varepsilon + \frac{1}{36}\varepsilon^2 \right) + O(\varepsilon^3). \quad (15.107)$$

Comparing the two expressions, we note immediately that Flory's formula is not exact in general, though, as mentioned before, it correctly predicts the upper-critical dimension $d = 4$, and the values of ν for $d = 1$ and 2 .

For $d = 3$, RG calculations yield (see Section 41.5)

$$\nu = 0.5874 \pm 0.0003,$$

result that can be compared with $\nu_{\text{Flory}} = 0.6000$. The two values are close, but significantly different. Actually, one can estimate the difference approximately. If one simply takes into account the orders ε and ε^2 , and the property that Flory's approximation is exact for $\varepsilon = 0, 2, 3$, one can write

$$\nu_{\text{RG}}^{-1} - \nu_{\text{Flory}}^{-1} = \frac{1}{12}\varepsilon \left(1 - \frac{1}{2}\varepsilon \right) \left(1 - \frac{1}{3}\varepsilon \right) \left(1 + \frac{19}{96}\varepsilon \right)^{-1} + O(\varepsilon^3).$$

For $\varepsilon = 1$, one finds successively $\nu = 0.590$ at order ε , and $\nu = 0.592$ at order ε^2 , results consistent with the RG value obtained by summing all terms (Tables 41.3 and 41.4). Experimental results, $\nu = 0.586 \pm 0.004$, and Monte Carlo simulations, $\nu = 0.58760$ (Table 41.5), agree with the RG prediction, and also exclude the value predicted by Flory's approximation.

Moreover, with the use of RG arguments, one can derive a number of universal properties and calculate many other universal quantities, which cannot be inferred from Flory's argument.

15.9 Liquid–vapour phase transition and ϕ^4 field theory

A priori, it is not obvious that the liquid–vapour phase transition in classical fluids belongs to the same universality class as the Ising model, with its \mathbb{Z}_2 symmetry, and can be described by a ϕ^4 field theory. In particular, no apparent discrete symmetry is broken. However, there exist several methods to relate the liquid–vapour transition and the $N = 1$ ϕ^4 theory: one can either start from a model which is physically not very realistic, the lattice gas model, but rigorously equivalent to the Ising model, or from a true gas model in the continuum, and by a number of approximations for which it is difficult to provide a rigorous justification, directly obtaining a ϕ^4 field theory.

The lattice gas model [140]. One considers particles living on a lattice, such that at each site i of a d -dimensional lattice (i is the position on the lattice) the particle occupation number n_i can only be 0 or 1. The lattice is a schematic way to represent a hard-core type interaction. A longer range attractive potential is provided by a nearest neighbour (n.n.) interaction which favours the occupation of neighbour sites. Finally, one adds a chemical potential μ in the form of a linear coupling to the number of particles. The configuration energy is given by

$$-\beta\mathcal{H}(n_i) = J \sum_{\text{n.n.}} n_i n_j - \mu \sum_i n_i. \quad (15.108)$$

This model is directly related to the Ising model by the transformation

$$n_i \mapsto S_i = 2n_i - 1, \quad (15.109)$$

where S_i is an Ising spin. In d dimensions, one finds

$$-\beta\mathcal{H}(S_i) = \frac{1}{4}J \sum_{\text{n.n.}} S_i S_j + \frac{1}{2}(dJ - \mu) \sum_i S_i + \text{const..} \quad (15.110)$$

One recognizes the energy of the Ising model with a nearest-neighbour interaction and a magnetic field. Since we have shown that the Ising model leads to the ϕ^4 field theory, the same applies to the lattice gas model, and the relation between the two models makes it possible to translate the RG results into the language of classical fluids.

15.9.1 The classical gas in the continuum: Field integrals

Since the introduction of a lattice is somewhat artificial, we now consider a classical gas in the continuum. For a class of models, using a sequel of exact identities one can derive a field integral representation of the partition function (in the grand canonical formalism) [141]. Then, a few additional steps, which are quite instructive but harder to justify rigorously, lead to the ϕ^4 field theory.

We consider a real classical gas in the continuum, and assume that the potential between particles is the sum of a hard-core (h.c) potential, with a core of size a , and a short-range attractive two-body potential V such that,

$$V(\mathbf{r}) = 0 \text{ for } |\mathbf{r}| < a, \text{ and } V(\mathbf{r}) = V(r = |\mathbf{r}|) \leq 0, \quad \forall \mathbf{r},$$

but with a range much larger than a . The canonical partition function for n particles of mass m in a d -dimensional volume Ω at temperature $1/\beta$ then has the form

$$\mathcal{Z}(n, \beta, \Omega) = \left(\frac{2\pi m}{\beta h^2} \right)^{nd/2} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^n d^d r_i \exp \left[-\beta \sum_{i < j} V(\mathbf{r}_i - \mathbf{r}_j) \right]. \quad (15.111)$$

Since the potential depends only on the distance r and is short range, its Fourier transform

$$\tilde{V}(\mathbf{k}) = \int d^d r e^{-i\mathbf{k}\cdot\mathbf{r}} V(\mathbf{r}),$$

is an analytic function of k^2 in a strip, and can be expanded for k^2 small as

$$\tilde{V}(\mathbf{k}) = -v(1 - \sigma^2 \mathbf{k}^2) + O(k^4), \quad v > 0, \quad (15.112)$$

where $\sigma^2 = \langle \mathbf{r}^2 \rangle \gg a^2$ is the second moment of the distribution $-V$. We assume that the longer range of V reflects into the condition $\sigma \gg a$.

We introduce the local gas density

$$\rho(\mathbf{r}) = \sum_{i=1}^n \delta^{(d)}(\mathbf{r} - \mathbf{r}_i), \quad (15.113)$$

where \mathbf{r}_i is the position of a particle in the gas, and $\int d^d r \rho(\mathbf{r}) = n$ is the number of particles. The gas potential energy can then be written as

$$\sum_{i < j} V(\mathbf{r}_i - \mathbf{r}_j) = \frac{1}{2} \int d^d r d^d r' V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}) \rho(\mathbf{r}').$$

We now impose the relation (15.113) by a field integral in the spirit of the method of Section 14.4.2, and insert the identity

$$\int [d\rho(\mathbf{r})] \int [d\phi(\mathbf{r})] \exp \left[\int d^d r \phi(\mathbf{r}) \rho(\mathbf{r}) - \sum_{i=1}^n \phi(\mathbf{r}_i) \right] = \text{const.}$$

into the expression (15.111) (note that the integral over ϕ represents a functional δ function, and thus the ϕ -integration contour is parallel to the imaginary axis). After these transformations, the partition function becomes

$$\begin{aligned} \mathcal{Z}(n, \beta, \Omega) &= e^{n\beta V(0)/2} \int [d\phi(r) d\rho(r)] \exp \left[-\frac{1}{2} \beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') \right. \\ &\quad \left. + \int d^d r \rho(\mathbf{r}) \phi(\mathbf{r}) \right] \mathcal{Z}_{\text{h.c.}}(n, \beta, \Omega, \phi), \end{aligned} \quad (15.114)$$

where $\mathcal{Z}_{\text{h.c.}}(n, \beta, \Omega, \phi)$ is the canonical partition function of particles interacting through a h.c. two-body potential in a random external (imaginary) potential $\phi(\mathbf{r})$:

$$\mathcal{Z}_{\text{h.c.}}(n, \beta, \Omega, \phi) = \left(\frac{2\pi m}{\beta h^2} \right)^{nd/2} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^n d^d r_i \exp \left(- \sum_i \phi(\mathbf{r}_i) \right). \quad (15.115)$$

We can transform equation (15.114) into an identity between grand canonical partition functions setting

$$\mathcal{Z}(\mu, \beta, \Omega) = \sum_n \frac{z^n}{n!} \mathcal{Z}(n, \beta, \Omega), \quad (15.116)$$

where z is the gas fugacity. We also set

$$z = e^{\beta\mu}, \quad \rho_0 = (2\pi)^{d/2}/\lambda^d, \quad (15.117)$$

where μ is a chemical potential and λ the thermal wavelength,

$$\lambda = \hbar \sqrt{\beta/m}. \quad (15.118)$$

The normalization density ρ_0 , which can be changed by shifting the chemical potential, thus increases with the temperature.

Expressing n as the space integral of the local density ρ , and using the integral representation (15.114), we obtain the field integral

$$\mathcal{Z}(\mu, \beta, \Omega) = \int [d\phi(r)d\rho(r)] \exp [-\mathcal{S}(\rho, \phi)], \quad (15.119)$$

with the action

$$\begin{aligned} \mathcal{S}(\rho, \phi) = & \frac{1}{2}\beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') \\ & - \int d^d r \rho(\mathbf{r}) (\phi(\mathbf{r}) + \beta\mu) - \ln \mathcal{Z}_{\text{h.c.}}(\Omega, \phi), \end{aligned} \quad (15.120)$$

and

$$\mathcal{Z}_{\text{h.c.}}(\Omega, \phi) = \sum_n \frac{\rho_0^n}{n!} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^n d^d r_i \exp \left[- \sum_i \phi(\mathbf{r}_i) \right]. \quad (15.121)$$

Note that, if we replace the chemical potential μ by an external potential $\mu(\mathbf{r})$, we transform the partition function (15.121) into the generating functional of density correlation functions.

The integral over the density field ρ is Gaussian and can thus be calculated explicitly. The integration is particularly useful in the case of long range potentials, like the Coulomb potential, which have a Fourier transform \tilde{V} singular at $\mathbf{k} = 0$, a situation we face in Chapter 31. Here, by contrast, we consider only short range forces and the integration then slightly obscures the physical meaning of the dynamic variables, ϕ having the nature of an imaginary external potential. Instead, we introduce the integral over ϕ , even though we cannot calculate it exactly.

Finally, we note that the expression (15.119) can easily be generalized to many-body forces, leading to an action which is a polynomial of higher degree in the local density ρ .

15.9.2 Phase transition

The partition function of the hard-core potential in presence of a given external one-body potential cannot be calculated exactly. If the hard core is completely neglected, the partition function is given by the field integral,

$$\begin{aligned} \mathcal{Z}(\mu, \beta, \Omega) = & \int [d\phi(r)d\rho(r)] \exp [-\mathcal{S}(\rho, \phi)] \\ \mathcal{S}(\rho, \phi) = & \frac{1}{2}\beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') \\ & - \int d^d r (\rho(\mathbf{r}) (\phi(\mathbf{r}) + \beta\mu) + \rho_0 e^{-\phi(\mathbf{r})}). \end{aligned} \quad (15.122)$$

In a calculation by the steepest descent method, we can, at leading order, replace the field ϕ by the solution of the corresponding saddle point equation

$$\rho(\mathbf{r}) = \rho_0 e^{-\phi(\mathbf{r})} \Rightarrow \phi(\mathbf{r}) = -\ln(\rho(\mathbf{r})/\rho_0).$$

The action becomes

$$\mathcal{S}(\rho) = \frac{1}{2}\beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') + \int d^d r \rho(\mathbf{r}) [\ln(\rho(\mathbf{r})/\rho_0) - 1 - \beta\mu]. \quad (15.123)$$

We now look for a homogeneous saddle point ρ . With the notation (15.112), we obtain

$$-\beta v\rho + \ln(\rho/\rho_0) = \beta\mu. \quad (15.124)$$

Using the equation to eliminate the chemical potential, we obtain the action density

$$\mathcal{S}/\Omega = -\rho + \frac{1}{2}\beta v\rho^2 = -\beta p,$$

where Ω is the volume and p the pressure. This yields a reasonable equation of state for small densities but, for large densities, the pressure is negative and unbounded. This reflects the collapse of the system in the case of a purely attractive potential.

In the presence of the hard core, the action has to be understood with an implicit short distance cut-off a . Moreover, the hard core for large densities, $\rho = O(a^{-d})$, becomes important, its contributions to $\mathcal{S}(\rho)$ can no longer be neglected and ensure the stability of the system. More precisely, the effect of the hard core can be studied by calculating the ϕ potential. Since we are interested in phenomena associated with distances much larger than the size a of the hard core, we can make a local expansion of $\ln \mathcal{Z}_{h.c.}(\phi)$ governed by the length scale a , in the successive terms of the sum (15.121). Using the leading order equation (15.124), we see that the effect is an addition of local ρ contributions to the action, which stabilize the ρ potential. For dimensional reasons, since ρ is an inverse volume, the combinations that appear are $a^d\rho$ and $a\nabla$ in the case of derivatives, justifying the idea that the hard core is important for large or fast-varying densities. Neglecting derivatives, one obtains an expression of the form

$$\mathcal{S}(\rho) = \frac{1}{2}\beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') - \int d^d r [\rho(\mathbf{r})(\phi(\mathbf{r}) + \beta\mu) + w(\phi(\mathbf{r}))], \quad (15.125)$$

where $w(\phi)$ is the hard-core free energy in the thermodynamic limit, as a function of a chemical potential $-\phi/\beta$:

$$\exp[\Omega w(\phi)] = 1 + \sum_{n=1} \frac{\rho_0^n e^{-n\phi}}{n!} \int_{|\mathbf{r}_i - \mathbf{r}_j| > 2a} \prod_{i=1}^n d^d r_i.$$

Again, we integrate over ϕ by the steepest descent method and replace ϕ by the solution of the saddle point equation:

$$\rho(\mathbf{r}) + w'(\phi(\mathbf{r})) = 0. \quad (15.126)$$

We introduce $U(\rho)$, the Legendre transform of $w(\phi)$ defined by equation (15.126) and

$$U(\rho) = -w(\phi) - \rho\phi, \quad (15.127)$$

in such a way that the approximated action now reads

$$\mathcal{S}(\rho) = \frac{\beta}{2} \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') + \int d^d r [U(\rho(\mathbf{r})) - \beta\mu\rho(\mathbf{r})]. \quad (15.128)$$

We then need some information about the hard-core free energy in the thermodynamic limit. The sum cannot be calculated for generic dimensions, but is easy in one dimension, because one can order positions on the line. Summation relies on introducing the Fourier transform of the θ functions. Setting

$$e^{\Omega F(z)} \underset{\Omega \rightarrow \infty}{=} 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!} \int_{\substack{r_i \in \Omega \\ |r_i - r_j| > 2a}} \prod_{i=1}^n dr_i,$$

one finds that F is solution of the transcendental equation,

$$F(z) = z e^{-aF(z)}.$$

Of particular interest is the behaviour of the function for $z \rightarrow +\infty$. One finds

$$aF = \ln z - \ln \ln z + \dots.$$

The leading term is generic and reflects the situation of close packing and thus maximal density. Indeed, for Ω large but finite, the partition sum is really a polynomial of degree of the order Ω/a^d . For z large, the polynomial is dominated by the terms of highest degree. Therefore,

$$F(z) \underset{z \rightarrow \infty}{\propto} a^{-d} \ln z.$$

This implies that the function $U(\rho)$ diverges for the close packing density, as expected (we disregard here possible phenomena which could occur for densities very close to close packing of the nature of a liquid-solid transition). It is straightforward to show that for $d = 1$ $U(\rho)$ is convex, a property which generalizes to generic dimensions. Indeed, $U''(\rho) = 1/w''(\phi)$ and $w''(\phi)$ is positive, because it has the form $\langle n^2 \rangle - (\langle n \rangle)^2$.

With these properties, we can now discuss the action (15.128) in the tree approximation, following the lines of Section 14.4.3, where the mean-field approximation was considered. We look for a homogeneous saddle point ρ solution of

$$\mathcal{S}'(\rho)/\Omega = -\beta v \rho + U'(\rho) - \beta \mu = 0. \quad (15.129)$$

Depending on the number of solutions of the equation, one finds one phase, or two phases separated by a transition. This is determined by the second derivative

$$\mathcal{S}''(\rho)/\Omega = U''(\rho) - \beta v.$$

The function $U''(\rho)$ is positive, diverges as $1/\rho$ for $\rho \rightarrow 0$, and is expected to diverge at the close packing density. Therefore, it has an absolute minimum at an intermediate density. For βv smaller than the minimum of U'' , the system has only one phase, the high temperature phase. For β larger than the minimum the equation has one or three solutions depending on the chemical potential. However, in general, the corresponding transition is first order because the third derivative of the action does not vanish. A second order transition is characterized by the vanishing of both the first and third derivative of the thermodynamic potential. The vanishing of the third derivative yields the condition

$$U'''(\rho) = 0.$$

This equation determines a critical value ρ_c of order a^{-d} of the density and therefore, from equation (15.129), a relation between chemical potential and temperature. The vanishing of the second derivative then determines the critical temperature $T_c = 1/\beta_c = v/U''(\rho_c)$.

Near β_c , the action for constant fields has the typical form of the ϕ^4 field theory, up to irrelevant higher order corrections,

$$\mathcal{S}(\rho)/\Omega = \frac{1}{2}s_2(\beta)(\rho - \rho_c)^2 + \frac{1}{4!}s_4(\rho - \rho_c)^4 + O((\rho - \rho_c)^5),$$

with

$$s_2(\beta) = U''(\rho_c) - v\beta = v(\beta_c - \beta).$$

Finally, for what concerns long distance properties the potential energy can be approximated by its two first terms in the derivative expansion (equation (15.112)):

$$\frac{1}{2}\beta \int d^d r d^d r' \rho(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') = \frac{1}{2}\beta v \int d^d r \left[-\rho^2(\mathbf{r}) + \sigma^2(\nabla \rho(\mathbf{r}))^2 \right].$$

The universal properties of the liquid–vapour transition near the special second-order transition point can thus be described by the one-component ϕ^4 field theory, and the transition belongs to the universality class of the Ising model although, here, the action has no reflection symmetry.

By varying both the temperature and a second thermodynamic variable like the chemical potential, one has been able to cancel both the $(\rho - \rho_c)^2$ and the $(\rho - \rho_c)^3$ term in the expansion of the action for $\rho - \rho_c$ small, and reach a situation of second order phase transition. Then, the odd term in $(\rho - \rho_c)$ of lowest dimension is $(\rho - \rho_c)^5$ which has dimension 5 for $d = 4$ and is, therefore, irrelevant at the phase transition. This explains that the ϕ^4 field theory can correctly describe the liquid–vapour phase transition. Let us, nevertheless, mention a few specific features of this transition with respect to magnetic transitions:

(i) The complicated, although regular, relations between the thermodynamic variables to the more natural, from the symmetry point of view, magnetic field, and temperature variables of the magnetic systems. This introduces additional parameters in the description of experimental results.

(ii) The appearance of corrections to scaling due to operators of odd canonical dimensions like ϕ^5 .

15.10 Superfluid transition

In this section, we briefly justify that the $(\phi^2)^2$ field theory with $O(2)$ symmetry describes the superfluid transition. Our discussion is based on the formalism introduced in Section 4.4.5.

15.10.1 The Bose partition function

In Section 4.4.5, we have shown that the partition function for a Bose fluid at temperature $1/\beta$, for particles of mass m , with two-body interactions, in the grand canonical formulation, can be expressed as a field integral of the form (equations (4.61, 4.62)),

$$\mathcal{Z}(\beta) = \int [d\bar{\varphi}(t, x)d\varphi(t, x)] \exp[-\mathcal{S}(\varphi, \bar{\varphi})],$$

with the periodic boundary conditions $\bar{\varphi}(-\beta/2, x) = \bar{\varphi}(\beta/2, x)$, $\varphi(-\beta/2, x) = \varphi(\beta/2, x)$, and the non-relativistic Euclidean action

$$\begin{aligned} \mathcal{S}(\bar{\varphi}, \varphi) &= \int dt d^d x \bar{\varphi}(x, t) \left(\frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla^2 - \mu \right) \varphi(x, t) \\ &\quad + \frac{1}{2} \int dt d^d x d^d y \bar{\varphi}(x, t) \varphi(x, t) V_2(x, y) \bar{\varphi}(y, t) \varphi(y, t), \end{aligned} \quad (15.130)$$

where μ is the chemical potential and V_2 the pair potential.

In the case of short range forces, because we are interested in large distance phenomena, we can approximate the potential by a δ -function pseudo-potential, and the action becomes local (Section 4.4.7):

$$\mathcal{S}(\bar{\varphi}, \varphi) = \int dt d^d x \left[\bar{\varphi}(x, t) \left(\frac{\partial}{\partial t} - \frac{\hbar^2}{2m} \nabla^2 - \mu \right) \varphi(x, t) + \frac{G}{2} (\bar{\varphi}(x, t) \varphi(x, t))^2 \right]. \quad (15.131)$$

We recall that the strength G of the interaction, which can be parametrized by ma^{d-2}/\hbar^2 , where a is a length proportional to the scattering length in $d = 3$, must be positive for stability, and this corresponds to a repulsive interaction between bosons.

The model has an obvious $U(1) \equiv O(2)$ symmetry,

$$\varphi \mapsto \varphi e^{i\theta}, \quad \bar{\varphi} \mapsto \bar{\varphi} e^{-i\theta},$$

corresponding to particle number conservation.

The partition function is obtained by integrating over the complex field φ with periodic boundary conditions in the Euclidean time direction.

The action density for constant fields is

$$\mathcal{E}(\bar{\varphi}, \varphi) = -\mu\varphi\bar{\varphi} + \frac{1}{2}G(\varphi\bar{\varphi})^2.$$

It predicts, in the tree approximation, a second-order phase transition at vanishing chemical potential μ . For $\mu > 0$, the field acquires a non-vanishing expectation value,

$$|\varphi|^2 = \mu/G.$$

In the Gaussian approximation, in the disordered phase, the non-relativistic two-point function $\tilde{\Delta}(\omega, \mathbf{k})$ in the Fourier representation is

$$\tilde{\Delta}(\omega, \mathbf{k}) = \frac{1}{i\omega + k^2/2m - \mu}.$$

As a consequence of the periodic boundary conditions, the field can be expanded on a basis of periodic functions of period β as

$$\varphi(x, t) = \sum_{\nu \in \mathbb{Z}} e^{2i\pi\nu t/\beta} \varphi_{\nu}(x), \quad \bar{\varphi}(x, t) = \sum_{\nu \in \mathbb{Z}} e^{-2i\pi\nu t/\beta} \bar{\varphi}_{\nu}(x),$$

and thus ω is quantized: $\omega = \omega_{\nu} \equiv 2\pi\nu/\beta$. Near the transition point $\mu = 0$, the large distance behaviour of the equal-time two-point function,

$$\Delta(x, 0) = \frac{1}{(2\pi)^d} \frac{1}{\beta} \sum_{\nu \in \mathbb{Z}} \int \frac{d^d k e^{ikx}}{i\omega_{\nu} + k^2/2m}, \quad (15.132)$$

is dominated by the zero mode $\nu = 0$, the contributions of other modes decreasing exponentially,

$$\Delta(x, 0) \underset{|x| \rightarrow \infty}{\sim} \frac{1}{(2\pi)^d} \frac{1}{\beta} \int \frac{d^d k e^{ikx}}{k^2/2m} \propto |x|^{2-d}.$$

One finds the usual behaviour of the propagator in a massless field theory in d isotropic dimensions.

15.10.2 Critical properties: Beyond the Gaussian approximation

The critical behaviour of correlation functions, in the Gaussian approximation, is entirely obtained from the contribution of the zero mode. In a perturbative expansion, at each order, the leading singularities are again generated by the zero-mode contributions. This argument indicates that the critical behaviour, beyond the Gaussian approximation, can be derived from a simplified field theory. At leading order, we simply replace in the action the field $\varphi(x, t)$ by its zero mode $\varphi_0(x)$. A more detailed discussion in Chapter 32 confirms the analysis. In fact, the quantity $\sqrt{m/\beta}$ acts as an effective momentum cut-off, or the thermal wavelength

$$\lambda = \hbar\sqrt{\beta/m}, \quad (15.133)$$

as a short-distance cut-off. Note that the δ -function approximation for the potential makes sense if λ is much larger than the range of the potential, that is, if the temperature is low enough.

At leading order, the effective Hamiltonian for the zero mode then reads

$$\mathcal{H}(\bar{\varphi}_0, \varphi_0) = \beta \int d^d x \left[\bar{\varphi}_0(x) \left(-\frac{\hbar^2}{2m} \nabla^2 - \mu \right) \varphi_0(x) + \frac{G}{2} (\bar{\varphi}_0(x) \varphi_0(x))^2 \right], \quad (15.134)$$

which is also the classical approximation. An integration over the non-zero, non-critical modes simply generates a renormalization of the parameters of the effective long distance action. We recognize the Euclidean action of the $O(2)$ symmetric $(\phi^2)^2$ field theory,

$$\mathcal{S}(\phi) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g \Lambda^{4-d} (\phi^2(x))^2 \right] d^d x,$$

where ϕ is a two-component vector, in a different normalization and parametrization, the parameters and dynamical variables having also a different interpretation. To identify the two actions, one can set

$$\phi_1(x) + i\phi_2(x) = \zeta \varphi_0(\theta x),$$

with

$$\theta = \lambda \Lambda, \quad \zeta^2 = \Lambda^{d-2} \lambda^d.$$

Comparing the two expressions, one obtains the relations between parameters (a is the scattering length)

$$g = 12\lambda^{-d}\beta G \propto (a/\lambda)^{d-2}, \quad r = -2\mu\beta\Lambda^2.$$

This field theory can then be studied by RG methods as in Section 15.7.

Note that from the derivation of the field integral representation in Section 4.4.5, it follows that the field ϕ is not a direct observable. Only correlations of ϕ^2 , which is the density field, are physical. Two quantities which can be measured directly are the specific heat and the density near criticality. For instance, the density ρ is given by the derivative of the free energy per unit volume with respect to $\beta\mu$:

$$\rho = \Omega^{-1} \frac{\partial \mathcal{W}}{\partial(\beta\mu)} = -2\Lambda^{2-d} (m/\beta\hbar^2)^{d/2} \Omega'^{-1} \frac{\partial \mathcal{W}}{\partial r},$$

where Ω is the volume in the initial variables and Ω' in those of the $(\phi^2)^2$ field theory. They depend on the two critical exponents ν and α , which are related. A measurement of both quantities yields two different determinations of the exponent ν [143].

In this system, two parameters can be tuned to reach criticality, temperature, and chemical potential in the theoretical model, in actual experiments temperature and pressure. By varying the pressure, one can check universality and get some handle on corrections to scaling.

The zero-temperature limit. A special situation corresponds to zero temperature. Then, the Fourier modes are continuous, and the complete non-relativistic theory becomes relevant. Power counting of the type encountered in critical dynamics, with time identified to the square of a space variable, shows that the theory is exactly renormalizable in $d = 2$ space dimensions, where logarithmic corrections to the mean-field approximation are expected.

Quantum and classical fluids: The density field. We have shown that critical properties can be obtained from the classical approximation. However, we have studied a region of parameter where the zero mode is critical. By contrast, in the case of classical fluids, the field φ_0 is not critical; the thermal wave length is small, and the δ -function approximation is no longer justified. It is the density $\rho(x) = \bar{\varphi}_0(x)\varphi_0(x)$ that is the order parameter and has critical fluctuations. The field φ_0 can be integrated out to generate an effective field theory for the density field ρ of the form discussed in Section 15.9, and which eventually leads to a $N = 1$, ϕ^4 field theory in the critical domain.

15.10.3 Superfluid transition and Bose–Einstein condensation

In the case of a Bose gas, the approximation of two-body interaction is natural because the system is dilute. In the case of a Bose fluid, this is no longer the case. However, additional higher powers of $\bar{\varphi}\varphi$ in the action are irrelevant, and thus do not modify the critical behaviour. Only in the exceptional case where the renormalized $(\bar{\varphi}\varphi)^2$ interaction is small, three-body interactions, which generate a $(\bar{\varphi}\varphi)^3$ term, may become important and lead to a tricritical behaviour [144].

By contrast, in dilute Bose gases, of the kind studied in Bose–Einstein condensation experiments, three-body interactions are totally negligible compared to two-body interactions, even though the latter are also weak. In such systems, at the transition temperature, correlation functions depend only on the RG invariant length scale ξ , which satisfies [145]

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} \right) \xi(\Lambda, g) = 0,$$

and thus is proportional to

$$\xi \propto \Lambda^{-1} g^{1/(d-4)} \exp \left[\int_0^g dg' \left(\frac{1}{\beta(g')} - \frac{1}{(d-4)g'} \right) \right].$$

We note that, for generic values of the ϕ^4 coupling, that is, $g = O(1)$, typical for a fluid, the length scale ξ is of the order of the microscopic scale, and long distance physics corresponds to distances much larger than $1/\Lambda$ and thus than ξ . This physics is entirely described by the QFT of the IR fixed point.

However, if as in the case of the dilute Bose gas g is very small, then ξ is much larger than the microscopic length scale. It becomes a *crossover length* between two universal behaviours. For distances much larger than ξ , physics can still be derived from the IR fixed point, and is analogous to the superfluid transition. By contrast, for distances much larger than $1/\Lambda$ but much smaller than ξ , physics is dominated by the Gaussian (UV) fixed point, and, therefore, one observes a behaviour close to the Bose–Einstein condensation of the free Bose gas.

16 Critical domain: Universality, ε -expansion

In Chapter 15, we have derived the scaling behaviour of correlation functions at criticality, that is, $T = T_c$. We now study the *critical domain*, which is defined by the property that the correlation length is large with respect to the microscopic scale, but finite.

By a simple generalization of the formalism of Chapter 9 to dimensions $d < 4$, we first demonstrate strong scaling above T_c : in the critical domain above T_c , all correlation functions, after rescaling, can be expressed in terms of universal correlation functions, in which the scale of distance is provided by the correlation length.

However, because the correlation length is singular at T_c , in this formalism, the critical temperature cannot be crossed, making a consistent description of the whole critical domain impossible. Therefore, we then expand correlation functions, with some care, in formal power series of the deviation $(T - T_c)$ from the critical temperature. The coefficients are critical correlation functions involving $\phi^2(x)$ insertions at zero momentum (as one can infer from the analysis of Section 9.10), whose scaling behaviour has been derived in Section 15.6. Summing the expansion, one obtains renormalization group (RG) equations valid for $T \neq T_c$. In addition, to crossing the critical temperature, while avoiding critical singularities, it is necessary to break the symmetry of the Hamiltonian explicitly. We thus add a linear coupling to a small external magnetic field to the interactions. We then derive RG equations in a magnetic field, or at fixed magnetization (the field expectation value $\langle \phi(x) \rangle$). In this way, we are able to continuously connect correlation functions above and below T_c , and derive scaling properties in the whole critical domain.

In the first part of the chapter, we restrict the discussion to Ising-like systems. In Section 16.6, we generalize the results to an N -component order parameter $\phi(x)$. In Section 16.8, we show how to expand the universal two-point function when T approaches T_c , using the short distance expansion (SDE), discussed in Section 11.3.

In Section 16.9, we report a few terms of the ε expansion ($\varepsilon = 4 - d$ is the deviation from dimension 4) of various universal quantities. Calculations at fixed dimension and summation of perturbative expansions are described in Sections 41.3–41.5.

Finally, in Section 16.10, we briefly describe the application of the SDE to the determination of critical exponents.

Background material and additional references can be found, for example, in Ref. [50].

The temperature. The temperature is coupled to the total Hamiltonian or configuration energy. Therefore, a variation of the temperature generates a variation of all terms contributing to the effective Hamiltonian. However, as we have shown in Section 15.2, the most relevant contribution (the most infrared (IR) singular) corresponds to the $\phi^2(x)$ operator. Therefore, we take the difference $\tau = r - r_c$ between the coefficient of ϕ^2 in the Hamiltonian (15.41) and its critical value r_c as a linear measure of the deviation from the critical temperature, and parametrize the effective Hamiltonian (15.41) as

$$\mathcal{H}(\phi) = \int \left[\frac{1}{2} (\nabla_\Lambda \phi(x))^2 + \frac{1}{2} (r_c + \tau) \phi^2(x) + \frac{1}{4!} g \Lambda^\varepsilon \phi^4(x) \right] d^d x, \quad (16.1)$$

with $|\tau| \ll \Lambda^2$. Dimensional analysis, which results from the Gaussian renormalization, then yields for vertex (one-particle irreducible or 1PI) functions the relation

$$\tilde{\Gamma}^{(n)}(p_i; \tau, g, \Lambda) = \Lambda^{d-n(d-2)/2} \tilde{\Gamma}^{(n)}(p_i \Lambda^{-1}; \tau \Lambda^{-2}, g, 1). \quad (16.2)$$

16.1 Strong scaling above T_c : The renormalized theory

We first describe a few properties of the critical behaviour above T_c . For $T > T_c$, because the theory is massive, it is possible to introduce a special renormalized quantum field theory (QFT) with renormalization conditions imposed at zero momentum. Generalizing to any dimension $d \leq 4$ the formalism of Sections 9.2 and 9.3, we define renormalized vertex functions (see equations (9.30)) by the renormalization conditions at zero momentum,

$$\tilde{\Gamma}_r^{(2)}(p; m_r, g_r) = m_r^2 + p^2 + O(p^4), \quad (16.3)$$

$$\tilde{\Gamma}_r^{(4)}(p_i = 0; m_r, g_r) = m_r^\varepsilon g_r. \quad (16.4)$$

The condition (16.4) is written in such a way that g_r is dimensionless. The renormalized theory is derived from the initial microscopic theory by taking the large cut-off limit at g_r and m_r fixed:

$$\tilde{\Gamma}_r^{(n)}(p_i; m_r, g_r) = \lim_{\Lambda \rightarrow \infty} Z^{n/2}(m_r/\Lambda, g_r) \tilde{\Gamma}_r^{(n)}(p_i; \tau, g, \Lambda). \quad (16.5)$$

A simple dimensional analysis then yields for the renormalized vertex functions, the relation

$$\tilde{\Gamma}_r^{(n)}(p_i; m_r, g_r) = m_r^{d-(n/2)(d-2)} \tilde{\Gamma}_r^{(n)}(p_i/m_r; 1, g_r). \quad (16.6)$$

The correlation length ξ , which characterizes the decay of the connected two-point function, is proportional to m_r^{-1} : it fixes the scale in the theory with renormalization conditions at zero momentum.

Note that equation (16.5) holds for any dimension $d \leq 4$ and, therefore, some consequences of the equation are valid beyond the ε -expansion. This is one reason why this formalism is especially useful.

In Section 9.5, we have introduced the Callan–Symanzik equations, and in the following sections proved that are satisfied by the renormalized vertex functions in four dimensions. Below four dimensions, the proof is much simpler, because only the mass is divergent. With the renormalization conditions (16.3, 16.4), one thus derives

$$\begin{aligned} & \left[m_r \frac{\partial}{\partial m_r} + \beta(g_r) \frac{\partial}{\partial g_r} - \frac{n}{2} \eta(g_r) \right] \tilde{\Gamma}_r^{(n)}(p_i; m_r, g_r) \\ &= m_r^2 [2 - \eta(g_r)] \tilde{\Gamma}_r^{(1,n)}(0; p_i; m_r, g_r), \end{aligned} \quad (16.7)$$

in which the renormalized ϕ^2 insertion is specified by the condition

$$\tilde{\Gamma}_r^{(1,2)}(0; 0, 0; m_r, g_r) = 1. \quad (16.8)$$

As usual, we have given to different RG functions the same name $\{\beta, \eta\}$ because they play the same role in the equations. We examine some consequences of equation (16.7) in Section 16.1.3.

16.1.1 Microscopic and renormalized coupling constants

We first discuss the relation between microscopic and renormalized coupling constants. Several remarks are in order here:

(i) Since, in dimension $d < 4$, the ϕ^4 QFT is super-renormalizable, all correlation functions have a large cut-off limit after a simple mass renormalization, that is, after one has taken the correlation length (or the physical mass) as a parameter. The coupling constant and the field renormalizations are finite and, therefore are not required, in general.

If one renormalizes only the mass, one obtains a finite theory which is a function of the bare coupling constant $g\Lambda^\varepsilon$, the coefficient of $\phi^4(x)$ in the Hamiltonian. Therefore, perturbation theory is really an expansion in powers of $g(\Lambda/m_r)^\varepsilon$. It only makes sense if this ratio is kept fixed, which implies that the coupling constant $g\Lambda^\varepsilon$ (which characterizes the deviation from the Gaussian fixed point) goes to zero with the inverse correlation length as m_r^ε . This is indeed what is implicitly assumed in the conventional QFT renormalization theory. By contrast, in a generic situation, as envisaged in the theory of the critical phenomena, the coupling constant, which is related to microscopic parameters of the theory, is fixed. This means, as we already stressed in Chapter 15, that after introduction by rescaling of the cut-off Λ , $g \mapsto g\Lambda^\varepsilon$, g remains finite (see action (15.41)) when $\Lambda \rightarrow \infty$. Therefore, in the critical domain, that is, in the large cut-off limit, the expansion parameter becomes large for $d < 4$, and the mass renormalized perturbation theory becomes useless. This is the reason why it is necessary to introduce a field renormalization and a new expansion parameter g_r which, as we will verify, remains finite in this limit.

(ii) We note that by this method (introduction of the renormalized theory), we have taken the large cut-off limit in the following order: first $m \ll \Lambda g^{1/\varepsilon} \ll \Lambda$, and then $\Lambda g^{1/\varepsilon} \rightarrow \infty$, instead of $m \ll \Lambda \propto \Lambda g^{1/\varepsilon}$. For an application to critical phenomena, we have to assume that the result is the same at leading order.

The renormalized coupling constant. The RG β -function in equation (16.7) is given by equation (9.37),

$$\beta(g_r) = m_r \left. \frac{d}{dm_r} \right|_{\Lambda, g} g_r(g, m_r/\Lambda). \quad (16.9)$$

We now introduce the variable $\lambda = m_r/\Lambda$. Considering g_r as a function of λ at g fixed, we can rewrite equation (16.9) as

$$\lambda \frac{d}{d\lambda} g_r(\lambda) = \beta(g_r(\lambda)). \quad (16.10)$$

This equation, which is similar to equation (9.98), is a flow equation for $g_r(\lambda)$. When the correlation length increases, the ratio m_r/Λ decreases, and thus λ goes to zero. The renormalized coupling constant is driven towards an IR stable zero of $\beta(g_r)$ if such a zero exists. Assuming the existence of such an IR fixed point,

$$\beta(g_r^*) = 0, \quad \text{with } \omega = \beta'(g_r^*) > 0, \quad (16.11)$$

one can integrate equation (16.10) in the neighbourhood of g_r^* , and estimate $g_r - g_r^*$:

$$|g_r - g_r^*| \propto (m_r/\Lambda)^\omega. \quad (16.12)$$

Therefore, because the bare coupling constant, which is associated with microscopic physics, is fixed in the critical domain, the renormalized, or effective coupling constant at correlation scale, is close to the IR fixed point value.

16.1.2 Strong scaling

We first evaluate the behaviour of the renormalization constant Z defined by equation (9.38),

$$m_r \left. \frac{d}{dm_r} \right|_{g, \Lambda} \ln Z(g_r, m_r/\Lambda) = \eta(g_r). \quad (16.13)$$

Using the parameter $\lambda = m_r/\Lambda$ and the function $g_r(\lambda)$ of equation (16.10), we rewrite equation (16.13) as

$$\lambda \frac{d}{d\lambda} \ln Z(\lambda) = \eta(g_r(\lambda)). \quad (16.14)$$

In the case of an IR fixed point, for $\lambda = m_r/\Lambda$ small, at leading order one can then replace g_r by g_r^* . One infers

$$Z \propto (m_r/\Lambda)^\eta, \quad (16.15)$$

where η is the value (assumed finite) of the function $\eta(g)$ at the fixed point:

$$\eta \equiv \eta(g_r^*). \quad (16.16)$$

We also assume that the renormalized vertex functions $\tilde{\Gamma}_r^{(n)}(p_i; m_r, g_r^*)$ are finite.

The parameter Λ provides a (momentum) scale for all quantities. Once the form of the critical behaviour has been determined, all dimensional parameters (here momenta, deviation from T_c , and correlation length) can be rescaled to eliminate Λ (see equation (16.2)).

Assuming the existence of an IR fixed point, combining equations (16.5), (16.6), and (16.15), one then finds, in the critical domain, the scaling relation [80, 132, 51]

$$\tilde{\Gamma}^{(n)}(p_i, \tau, g, \Lambda = 1) \underset{m_r \ll 1, |p_i| \ll 1}{\sim} m_r^{d-(n/2)(d-2+n)} \tilde{\Gamma}_r^{(n)}(p_i/m_r; 1, g_r^*). \quad (16.17)$$

Equation (16.17) provides a proof of *strong scaling* and *universality* in the whole critical domain above T_c . The initial effective local QFT may depend on an infinite number of parameters (see Chapter 8). The ϕ^4 QFT, viewed as an effective microscopic theory where all irrelevant contributions are omitted, depends only on the way the cut-off is introduced and explicitly on p_i , g , and τ . The right-hand side of equation (16.17) only involves the renormalized correlation functions at $g_r = g_r^*$, that is, functions of ratios p_i/m_r . Moreover, the result (16.17) holds for any fixed dimension $d \leq 4$. Resorting to the ε -expansion has been avoided. It has only been necessary to assume the existence of an IR fixed point. Of course, within the ε -expansion, using the results of Chapter 10 and the remark of Section 9.10.1, we immediately obtain $\beta(g_r)$ at leading order and thus g_r^* and ω ,

$$g_r^* = \frac{16\pi^2}{3}\varepsilon + O(\varepsilon^2), \quad \omega = \varepsilon + O(\varepsilon^2). \quad (16.18)$$

However, as first suggested by Parisi [146], it is also possible to numerically analyse the perturbative expansion in powers of g_r at fixed dimension 3 or 2 (see Chapter 41). Such an analysis convincingly demonstrates the existence of an IR fixed point, and makes a precise determination of critical exponents possible, as the numerical results reported in Section 41.5 illustrate.

Finally, equations (9.39–9.42) can be used to characterize the divergence of m_r^{-1} , and thus the correlation length, as a function of the temperature or bare mass. This behaviour is derived in Section 16.3 by another, simpler, method. It confirms that the relation between m_r and τ is singular. Therefore, within the formalism of this section, based on the introduction of a zero-momentum renormalized theory, crossing the critical temperature is impossible. Indeed, all correlation functions are parametrized in terms of the correlation length, which is singular at T_c . To avoid this problem, in the next section we introduce a different formalism which is a natural extension of the formalism of Chapter 15.

A few remarks.

(i) We have shown that the renormalized coupling constant g_r has a finite limit $g_r = g_r^*$, although the expansion parameter $g\Lambda^\varepsilon$ becomes infinite. In addition, precisely at g_r^* , the field renormalization diverges as equation (16.15) shows. The conclusion is that the QFT, at the IR fixed point, behaves even below four dimensions like a renormalizable QFT, and a complete renormalization is indeed required.

(ii) A second somewhat related remark is that when $g\Lambda^\varepsilon$ varies from 0 to infinity, g_r varies from 0 to g_r^* . This property suggests that no renormalized theory exists for $g_r > g_r^*$. Since $g_r^* = O(\varepsilon)$, this argument again confirms that it is unlikely that a non-trivial ϕ^4 QFT consistent on all scales exists in four dimensions. On the other hand, to construct the renormalized theory, one takes the large cut-off limit at $g\Lambda^\varepsilon$ as fixed. This procedure is only legitimate if the bare RG has only one IR fixed point. Otherwise, other non-trivial theories might be obtained by sending the cut-off to infinity at g fixed. This point is further discussed in Section 24.1.2.

16.1.3 Large momentum behaviour

We return to equation (16.7) in order to investigate the behaviour of correlation functions at large momenta, in the critical domain $m_r \ll |p_i| \ll \Lambda$. At leading order, we can replace, in equation (16.7), g_r by its IR fixed point value g_r^* :

$$\left(m_r \frac{\partial}{\partial m_r} - \frac{n}{2} \eta \right) \tilde{\Gamma}_r^{(n)}(p_i; m_r, g_r^*) = m_r^2 (2 - \eta) \tilde{\Gamma}_r^{(1,n)}(0; p_i; m_r, g_r^*). \quad (16.19)$$

It follows from the scaling relation (16.17) that the large momentum behaviour is directly related to the approach to the critical theory $m_r = 0$. Therefore, in order to extract some information from equation (16.19), it is necessary to return to the framework of the ε -expansion. An extension of Weinberg's theorem indicates that, order by order in g_r and ε , the right-hand side of equation (16.19) is then negligible at large (non-exceptional) momenta, or small masses, as one might naively guess from the factor m_r^2 that has been factorized. By contrast, below four dimensions at a fixed dimension, there always exists an order at which the right-hand side ceases to be negligible. This is interpreted as being a consequence of expanding perturbation theory around the wrong fixed point, the IR unstable Gaussian fixed point. The correct asymptotic behaviour of correlation functions is different, because it is governed by the non-trivial point $g_r = g_r^*$. Therefore, asymptotically the renormalized vertex functions satisfy

$$\left(m_r \frac{\partial}{\partial m_r} - \frac{n}{2} \eta \right) \tilde{\Gamma}_r^{(n)}(p_i, m_r) \underset{|p_i| \gg m_r}{\approx} 0. \quad (16.20)$$

Combined with the relation (16.6), this leads to [80, 132, 51]:

$$\tilde{\Gamma}_r^{(n)}(\lambda p_i; m_r) \underset{|p_i| \gg m_r}{\sim} \lambda^{d-(n/2)(d-2+\eta)} \tilde{\Gamma}_r^{(n)}(p_i; m_r). \quad (16.21)$$

Using then the relation (16.5) between microscopic and renormalized correlation functions, one recovers the scaling behaviour of equation (15.60) derived in Section 15.5.

In this approach, the large momentum behaviour of the right-hand side of equation (16.19) can also be analysed and, therefore, corrections to the leading behaviour can be calculated, using the SDE (see Section 16.8).

Remark. If g_r is fixed with $g_r < g_r^*$, then the large momentum behaviour of correlation functions is given by the UV fixed point $g_r = 0$, that is, by perturbation theory. However, as we have already discussed, this corresponds to a situation where the bare coupling constant g goes to zero as $(m_r/\Lambda)^\varepsilon$, which is a non-generic situation.

16.2 Critical domain: Homogeneous RG equations

To characterize the behaviour of correlation functions in the whole critical domain, we now use a different strategy. First, we consider the field integral representation of n -point correlation functions corresponding to the Hamiltonian (16.1). If we formally expand it in powers of τ , we obtain an expansion in terms of *critical* correlation functions with $\frac{1}{2} \int d^d x \phi^2(x)$ (the most IR singular part of the energy operator) insertions (see also Section 9.10). As a consequence, to be able to define these correlation functions by their perturbative expansion, we now have to return to the framework developed in Sections 15.4 and 15.6 and to the ε -expansion.

However, even so, because the insertion of $\int d^d x \phi^2(x)$ is the insertion of the Fourier transform of $\phi^2(x)$ at zero momentum, the corresponding correlation functions are still IR divergent. Therefore, in the Hamiltonian (16.1), we first replace the constant τ by a field $\tau(x)$. We can then expand the vertex functions $\Gamma^{(n)}$, as functions of space variables, in powers of the field $\tau(x)$:

$$\Gamma^{(n)}(x_i; \tau, g, \Lambda) = \sum_{l=0}^{\infty} \frac{1}{l!} \int d^d y_1 \cdots d^d y_l \tau(y_1) \cdots \tau(y_l) \Gamma^{(l,n)}(y_j; x_i; \tau = 0, g, \Lambda). \quad (16.22)$$

As already noted in Section 9.10, by acting with the functional differential operator $\int d^d y \tau(y) \delta/\delta\tau(y)$ on equation (16.22), one generates in the right-hand side a factor l in front of $\Gamma^{(l,n)}$. One then verifies that equation (15.67) implies (see equation (9.81)):

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) - \eta_2(g) \int d^d y \tau(y) \frac{\delta}{\delta \tau(y)} \right] \Gamma^{(n)}(x_i; \tau, g, \Lambda) = 0. \quad (16.23)$$

To calculate $\Gamma^{(n)}$, it is possible to perform a partial summation of perturbation theory in order to introduce the propagator $[-\nabla^2 + \tau(x)]^{-1}$, for example, by using the loop expansion. In this new perturbative expansion, the constant τ limit can be taken and leads to a massive theory: no IR divergence is generated anymore. Then, equation (16.23), in the Fourier representation, becomes [72]

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}^{(n)}(p_i; \tau, g, \Lambda) = 0. \quad (16.24)$$

Equation (16.24) is the formal analogue of equation (9.82), and differs from equation (16.7), which also applies to the non-critical (*i.e.* massive) theory, by the property that it is *homogeneous*.

16.3 Scaling properties above T_c

Within this new formalism, we first discuss again the behaviour of correlation functions in the critical domain above T_c , which we have already examined in Section 16.1. We integrate equation (16.24) using the method of characteristics. In addition to the functions $g(\lambda)$ and $Z(\lambda)$ of equations (15.51), we now define a function $\tau(\lambda)$ such that [51],

$$\lambda \frac{d}{d\lambda} \left[Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; \tau(\lambda), g(\lambda), \lambda \Lambda) \right] = 0, \quad (16.25)$$

and which is determined by imposing consistency with equation (16.24).

The consistency equations are:

$$\lambda \frac{d}{d\lambda} g(\lambda) = \beta(g(\lambda)), \quad g(1) = g, \quad (16.26)$$

$$\lambda \frac{d}{d\lambda} \ln \tau(\lambda) = -\eta_2(g(\lambda)), \quad \tau(1) = \tau, \quad (16.27)$$

$$\lambda \frac{d}{d\lambda} \ln Z(\lambda) = \eta(g(\lambda)), \quad Z(1) = 1. \quad (16.28)$$

Dimensional analysis yields

$$\tilde{\Gamma}^{(n)}(p_i; \tau(\lambda), g(\lambda), \lambda\Lambda) = (\lambda\Lambda)^{d-n(d-2)/2} \tilde{\Gamma}^{(n)}(p_i/\lambda\Lambda; \tau(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1). \quad (16.29)$$

The critical domain is defined, in particular, by $|\tau| \ll \Lambda^2$ and this is the source of the IR singular behaviour observed in perturbation theory. If the equation for λ ,

$$\tau(\lambda) = \lambda^2\Lambda^2, \quad (16.30)$$

has a solution, then the theory at scale λ is no longer critical. Combining equations (16.25–16.30), one finds

$$\tilde{\Gamma}^{(n)}(p_i; \tau, g, \Lambda) = Z^{-n/2}(\lambda) m^{(d-n(d-2)/2)} \tilde{\Gamma}^{(n)}(p_i/m; 1, g(\lambda), 1), \quad (16.31)$$

with the notation

$$m = \lambda\Lambda. \quad (16.32)$$

The solution of equation (16.27) can be written as

$$\tau(\lambda) = \tau \exp \left[- \int_1^\lambda \frac{d\sigma}{\sigma} \eta_2(g(\sigma)) \right]. \quad (16.33)$$

Substituting this relation into equation (16.30) and introducing the function (15.70),

$$\nu(g) = [\eta_2(g) + 2]^{-1},$$

one then obtains

$$\ln(\tau/\Lambda^2) = \int_1^\lambda \frac{d\sigma}{\sigma} \frac{1}{\nu(g(\sigma))}. \quad (16.34)$$

One looks for a solution λ in the limit $\tau/\Lambda^2 \ll 1$. Since $\nu(g)$ is a positive function, at least for g small enough, as one can verify in the explicit expression (15.74), from equation (16.34), one infers that the value of the parameter λ is small and, thus, that $g(\lambda)$ is close to the IR fixed point g^* . In this limit, $\nu(g(\sigma))$ can be replaced, at leading order, by the exponent $\nu = \nu(g^*)$ and equation (16.34) implies

$$\ln(\tau/\Lambda^2) \sim \frac{1}{\nu} \ln \lambda. \quad (16.35)$$

Equation (16.28) then yields

$$Z(\lambda) \propto \lambda^\eta. \quad (16.36)$$

Finally, taking the large Λ , or the small λ limit, and using equations (16.35) and (16.36) in equation (16.31), one finds that vertex functions have the scaling form

$$\tilde{\Gamma}^{(n)}(p_i; \tau, g, \Lambda = 1) \underset{\substack{\tau \ll 1 \\ |p_i| \ll 1}}{\sim} m^{(d-n(d-2+\eta)/2)} G_+^{(n)}(p_i/m), \quad (16.37)$$

where $G_+^{(n)}$ are universal functions, and

$$m(\Lambda = 1) = \xi^{-1} \propto \tau^\nu. \quad (16.38)$$

Equations (16.37) and (16.17) express the same scaling property. However, one additional result has been obtained. Equation (16.37) shows that the quantity m is proportional to the inverse correlation length (the physical mass in particle physics). Equation (16.38) then shows that the divergence of the correlation length $\xi = m^{-1}$ at T_c is characterized by the exponent ν (a result that could also have been derived with the formalism of Section 16.1).

From equation (16.37), one derives the scaling form of connected correlation functions:

$$\tilde{W}^{(n)}(p_i; \tau, g, \Lambda = 1) \underset{\substack{\tau \ll 1 \\ |p_i| \ll 1}}{\sim} m^{(d-n(d+2-\eta)/2)} H_+^{(n)}(p_i/m), \quad (16.39)$$

$$W^{(n)}(x_i; \tau, g, \Lambda = 1) \underset{\substack{\tau \ll 1 \\ |x_i| \gg 1}}{\sim} m^{(d-2+\eta)n/2} I_+^{(n)}(mx_i). \quad (16.40)$$

For $\tau \neq 0$, the correlation functions are finite at zero momentum and behave like

$$\tilde{W}^{(n)}(0; \tau, g, \Lambda) \propto \tau^{\nu(d-n(d+2-\eta)/2)}. \quad (16.41)$$

In particular, $n = 2$ corresponds to the magnetic susceptibility χ . Therefore,

$$\chi = \tilde{W}^{(2)}(p = 0; \tau, g, \Lambda) \propto \tau^{-\nu(2-\eta)}. \quad (16.42)$$

The exponent which characterizes the divergence of χ is usually denoted by γ . Equation (16.41) establishes the relation between exponents

$$\gamma = \nu(2 - \eta). \quad (16.43)$$

Finally, for the critical theory to exist when τ goes to zero, different powers of τ have to cancel in equation (16.37). From this observation, one recovers equation (15.60) in the form

$$\tilde{\Gamma}^{(n)}(\lambda p_i; \tau, g, \Lambda = 1) \underset{\substack{\tau^\nu \ll \lambda \\ |p_i| \ll 1}}{\propto} \lambda^{d-n(d-2+\eta)/2}. \quad (16.44)$$

16.4 Correlation functions with ϕ^2 insertions

A differentiation of the field integral with respect to $\tau(x)$, before taking the uniform τ limit, generates correlation functions with $[\frac{1}{2}\phi^2(x)]$ insertions (insertions of the Hamiltonian or configuration energy density, in the statistical formulation). By differentiating equation (16.23) with respect to $\tau(y_1), \dots, \tau(y_l)$, before taking the same limit, one derives RG equations for the corresponding vertex functions. One verifies that the resulting equation, except for $l = 2, n = 0$, is

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) - \eta_2(g) \left(l + \tau \frac{\partial}{\partial \tau} \right) \right] \tilde{\Gamma}^{(l,n)}(q_i; p_j; \tau, g, \Lambda) = 0. \quad (16.45)$$

The equation can be solved exactly in the same way as equation (16.24). We set $\Lambda = 1$. Then for $\tau \ll 1, |q_i| \ll 1, |p_j| \ll 1$, one finds the scaling form ($m \propto \tau^\nu$),

$$\tilde{\Gamma}^{(l,n)}(q_i; p_j; \tau, g, \Lambda = 1) \sim m^{[d-l/\nu-n(d-2+\eta)/2]} G_+^{(l,n)}(q_i/m; p_j/m). \quad (16.46)$$

The discussion then exactly follows the lines of Section 16.3, and we do not repeat it here. Rather, we focus on the case $n = 0, l = 2$, which corresponds to the energy density two-point correlation function. Starting from equation (15.67), and using the method described previously, one obtains

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \eta_2(g) \left(2 + \tau \frac{\partial}{\partial \tau} \right) \right] \tilde{\Gamma}^{(2,0)}(q; \tau, g, \Lambda) = \Lambda^{-\varepsilon} B(g). \quad (16.47)$$

The function $\Lambda^{-\varepsilon} C_2(g)$ of equation (15.75) is still a solution of the inhomogeneous equation. Equation (16.46) is then replaced by

$$\tilde{\Gamma}^{(2,0)}(q; \tau, g, \Lambda = 1) \underset{\substack{\tau \ll 1 \\ |q| \ll 1}}{\sim} m^{(d-2/\nu)} G_+^{(2,0)}(q/m) + C_2(g). \quad (16.48)$$

At zero momentum, $\tilde{\Gamma}^{(2,0)}(q = 0)$ is also the second derivative of the free energy (the connected vacuum amplitude) with respect to the temperature τ , that is, the *specific heat*. Its behaviour for $T \rightarrow T_{c+}$, is thus

$$\tilde{\Gamma}^{(2,0)}(0; \tau, g, \Lambda = 1) \sim A^+ \tau^{-(2-\nu d)} + C_2(g). \quad (16.49)$$

The specific heat exponent is called α . Therefore, we have established the scaling relation

$$\alpha = 2 - \nu d. \quad (16.50)$$

Integrating $\tilde{\Gamma}^{(2,0)}$ twice with respect to τ , one obtains the thermodynamic potential density,

$$\mathcal{G}(\tau, g, \Lambda) = \lim_{\Omega \rightarrow \infty} \Omega^{-1} \Gamma(\tau, g, \Lambda),$$

where Ω is the volume, for a vanishing magnetization M (and thus also the free energy in zero magnetic field). It has the expansion,

$$\begin{aligned} \mathcal{G}(M = 0, \tau, g, \Lambda) \\ = \Lambda^d \left[C_0(g) + C_1(g) \frac{\tau}{\Lambda^2} + C_2(g) \frac{\tau^2}{2\Lambda^4} + \frac{A^+}{(2-\alpha)(1-\alpha)} \left(\frac{\tau}{\Lambda^2} \right)^{2-\alpha} \right], \end{aligned} \quad (16.51)$$

valid for $|\tau| \ll \Lambda^2$. The three first terms correspond to the beginning of the small τ expansion of the regular part of the free energy and depend explicitly on g through three functions C_0 , C_1 , and C_2 , while the fourth term characterizes the leading behaviour of the singular part of the free energy and is universal (it still depends on the normalization of the temperature but this normalization cancels in appropriate ratios).

16.5 Scaling properties in a magnetic field and below T_c

In Sections 16.3 and 16.4, we have determined the behaviour of correlation functions in the critical domain, above T_c . We now examine the behaviour in the critical domain in the ordered phase ($M \neq 0$).

In order to pass continuously from the disordered ($T > T_c$) to the ordered phase ($T < T_c$), it is necessary to add an interaction that explicitly breaks the symmetry to the Hamiltonian. We thus consider correlation functions in the presence of an external magnetic field. Actually, it is more convenient to consider correlation functions at fixed magnetization. Therefore, we first discuss the relation between field and magnetization, called the *equation of state*.

As we have already noted in the mean-field analysis, in the ordered phase some qualitative differences appear between systems which have a discrete and a continuous symmetry. We consider first the case of the ϕ^4 (Ising-like) theory with a discrete \mathbb{Z}_2 symmetry. In Section 16.6, we discuss the N -vector model with $O(N)$ symmetry and $(\phi^2)^2$ interaction, an example of continuous symmetry.

16.5.1 The equation of state

The equation of state relates the magnetization M , expectation value of $\phi(x)$ in a constant magnetic field H , to the temperature. The thermodynamic potential density, as a function of M , is by definition

$$\mathcal{G}(M, \tau, g, \Lambda) = \sum_{n=0}^{\infty} \frac{M^n}{n!} \tilde{\Gamma}^{(n)}(p_i = 0; \tau, g, \Lambda). \quad (16.52)$$

The magnetic field H is given by

$$H = \frac{\partial \mathcal{G}}{\partial M} = \sum_{n=1}^{\infty} \frac{M^n}{n!} \tilde{\Gamma}^{(n+1)}(p_i = 0; \tau, g, \Lambda). \quad (16.53)$$

Noting that $n \equiv M(\partial/\partial M)$, one immediately derives, from the RG equation (16.24), the RG equation [51]

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \eta(g) \left(1 + M \frac{\partial}{\partial M} \right) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] H(M, \tau, g, \Lambda) = 0. \quad (16.54)$$

To integrate equation (16.54) by the method of characteristics, one introduces, in addition to $g(\lambda)$, $t(\lambda)$ and $Z(\lambda)$ of equations (16.26–16.28), a new function $M(\lambda)$ that satisfies

$$\lambda \frac{d}{d\lambda} \ln M(\lambda) = -\frac{1}{2} \eta[g(\lambda)], \quad \text{with } M(1) = M. \quad (16.55)$$

A comparison between equations (16.55) and (16.28) shows that $M(\lambda)$ is given by

$$M(\lambda) = M Z^{-1/2}(\lambda).$$

The solution of equation (16.54) can then be written as

$$H(M, \tau, g, \Lambda) = Z^{-1/2}(\lambda) H[M(\lambda), \tau(\lambda), g(\lambda), \lambda \Lambda]. \quad (16.56)$$

Dimensional analysis shows that

$$H(M, \tau, g, \Lambda) = \Lambda^{3-\varepsilon/2} H(M/\Lambda^{1-\varepsilon/2}, \tau/\Lambda^2, g, 1). \quad (16.57)$$

Applying the relation to the right-hand side of equation (16.56), one obtains

$$H(M, \tau, g, \Lambda) = (\lambda\Lambda)^{3-\varepsilon/2} Z^{-1/2}(\lambda) H[M(\lambda)/(\lambda\Lambda)^{1-\varepsilon/2}, \tau(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1]. \quad (16.58)$$

Again one can use the arbitrariness of λ to move outside the critical domain, in order to remove the critical singularities in the right-hand side of equation (16.58). Here, a natural choice is

$$M(\lambda) = (\lambda\Lambda)^{1-\varepsilon/2}, \quad (16.59)$$

which implies, using the solution of equation (16.55), that

$$\ln(M/\Lambda^{1-\varepsilon/2}) = \frac{1}{2} \int_1^\lambda \frac{d\sigma}{\sigma} [d - 2 + \eta(g(\sigma))]. \quad (16.60)$$

In the critical domain, the magnetization is small:

$$M \ll \Lambda^{1-\varepsilon/2}.$$

For $d \geq 2$ and g small, the expression $d - 2 + \eta(g)$ is positive because $\eta(g)$ is positive. This again implies that λ is small and thus $g(\lambda)$ is close to g^* . In this limit, equation (16.60) implies

$$M\Lambda^{\varepsilon/2-1} \propto \lambda^{(d-2+\eta)/2}. \quad (16.61)$$

From equation (16.27), one infers

$$\tau(\lambda)/\lambda^2 \propto \tau\lambda^{-1/\nu}, \quad (16.62)$$

and we have already seen that

$$Z(\lambda) \propto \lambda^\eta. \quad (16.63)$$

Finally, replacing $\tau(\lambda)$ and $Z(\lambda)$ by their asymptotic forms (16.62) and (16.63), and using equation (16.61) to eliminate λ , one concludes that, in the critical domain, the equation of state takes the general *scaling form*, proposed by Widom [122],

$$H(M, \tau, g, 1) \sim M^\delta f(\tau M^{-1/\beta}), \quad (16.64)$$

where $f(x)$ is, up to normalizations, a *universal function* and

$$\beta = \frac{\nu}{2}(d - 2 + \eta), \quad (16.65)$$

$$\delta = \frac{d + 2 - \eta}{d - 2 + \eta}. \quad (16.66)$$

Equations (16.65) and (16.66) relate the traditional critical exponents that characterize the vanishing of the spontaneous magnetization, and the singular relation between magnetic field and magnetization at T_c , respectively, to the exponents η and ν introduced previously.

Valid for $d < 4$, these latter two relations seem to be inconsistent with the values of the mean-field exponents for $d > 4$. To understand this point, it is necessary to remember that for $d > 4$, g^* vanishes, and all terms in H , except the term linear in M , come from corrections to expression (16.64) (see Section 17.1).

16.5.2 Properties of the universal function $f(x)$

(i) Griffith's analyticity [148]: equation (16.53) shows that H has a regular expansion in odd powers of M for $\tau > 0$. This implies that when the variable x becomes large and positive, $f(x)$ has the expansion

$$f(x) = \sum_{p=0}^{\infty} a_p x^{\gamma-2p\beta}. \quad (16.67)$$

(ii) For $M \neq 0$, when τ vanishes the theory remains massive. In the loop expansion, the corresponding propagator is massive. It follows that one can expand $\Gamma(M, \tau)$ and, therefore, $H(M, \tau)$ in powers of τ without meeting IR divergences. Therefore, $f(x)$ is infinitely differentiable at $x = 0$.

(iii) The appearance of a spontaneous magnetization below T_c , implies that the function $f(x)$ has a zero for $x = x_0$ with $x_0 < 0$,

$$f(x_0) = 0, \quad x_0 < 0. \quad (16.68)$$

Then, equation (16.64) leads, for $\tau < 0$, to the relation

$$M = |x_0|^{-\beta} (-\tau)^\beta, \quad \text{for } H = 0. \quad (16.69)$$

Equation (16.69) exhibits the singular behaviour of the spontaneous magnetization when the temperature approaches the critical temperature from below.

16.5.3 Correlation functions for non-vanishing magnetization

We now examine the behaviour of correlation functions in a field. All expressions are again be written for Ising-like systems. The generalization to the N -vector model with $O(N)$ symmetry is briefly discussed in Section 16.6.

The vertex functions, at fixed magnetization M , are obtained by expanding the generating functional $\Gamma(M)$ of vertex functions around $M(x) = M$,

$$\Gamma^{(n)}(x_1, \dots, x_n; \tau, M, g, \Lambda) = \frac{\delta^n}{\delta M(x_1) \cdots \delta M(x_n)} \Gamma(M, \tau, g, \Lambda) \Big|_{M(x)=M}. \quad (16.70)$$

The expansion in powers of M of the right-hand side of the equation, after Fourier transformation, then yields

$$\tilde{\Gamma}^{(n)}(p_1, \dots, p_n; \tau, M, g, \Lambda) = \sum_{s=0}^{\infty} \frac{M^s}{s!} \tilde{\Gamma}^{(n+s)}(p_1, \dots, p_n, 0, \dots, 0; \tau, 0, g, \Lambda). \quad (16.71)$$

From the RG equations satisfied by the correlation functions in zero magnetization (equations (16.24)), one derives [51]

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \eta(g) \left(n + M \frac{\partial}{\partial M} \right) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}^{(n)}(p_i; \tau, M, g, \Lambda) = 0. \quad (16.72)$$

The equation can be solved by exactly the same method as equation (16.54). One obtains the universal scaling form of vertex functions,

$$\tilde{\Gamma}^{(n)}(p_i; \tau, M, g, \Lambda = 1) \sim m^{[d-n(d-2+\eta)/2]} G^{(n)}(p_i/m, \tau m^{-1/\nu}), \quad (16.73)$$

for $|p_i| \ll 1$, $|t| \ll 1$, $M \ll 1$ and with here the definition

$$m = M^{\nu/\beta}. \quad (16.74)$$

The right-hand side of equation (16.73) now depends on two different length or mass scales: $m = M^{\nu/\beta}$ and τ^ν . The scaling forms of correlation functions differ by the powers of m as in equations (16.39) and (16.40).

16.5.4 Correlation functions in zero field below T_c : Spontaneous symmetry breaking

We have argued previously that, for $M \neq 0$, correlation functions are regular functions of τ near $\tau = 0$. Therefore, it is possible to cross the critical temperature and to then take the zero external magnetic field limit. In the limit, M becomes the *spontaneous magnetization* which is given, as a function of τ , by equation (16.69). After elimination of M in favour of τ in equation (16.73), one finds, below T_c ($\tau < 0$) in the critical domain, a behaviour of the form

$$\tilde{\Gamma}^{(n)}(p_i; \tau, M(\tau, H=0), g, 1) \sim m^{d-n(d-2+\eta)/2} G_-^{(n)}(p_i/m), \quad (16.75)$$

with

$$m = |x_0|^{-\nu} (-\tau)^\nu. \quad (16.76)$$

Therefore, vertex and connected correlation functions have exactly the same scaling behaviour above and below T_c [51]. In particular, since traditionally one defines below T_c ,

$$m^{-1} = \xi \propto (-\tau)^{-\nu'}, \quad [\tilde{\Gamma}^{(2)}(0)]^{-1} = \chi \sim (-\tau)^{-\gamma'}, \quad (16.77)$$

this establishes the relations $\nu' = \nu$ and $\gamma' = \gamma$. However, the universal functions $G_+^{(n)}$ and $G_-^{(n)}$ are different.

The extension of these considerations to the functions with ϕ^2 insertions, $\tilde{\Gamma}^{(l,n)}$ is straightforward. In particular, the same method yields the singular behaviour of the specific heat below T_c :

$$\tilde{\Gamma}^{(2,0)}(q=0, M(H=0, \tau)) - \Lambda^{-\varepsilon} C_2(g) \underset{\text{for } \tau \rightarrow 0_-}{\sim} A^-(-\tau)^{-\alpha}, \quad (16.78)$$

which, similarly, proves that the exponents above and below T_c are the same: $\alpha' = \alpha$.

Note that the constant term $\Lambda^{-\varepsilon} C_2(g)$ which depends explicitly on g is the same above and below T_c , in contrast with the coefficient of the singular part.

The derivation of the equality of exponents above and below T_c , relies on the existence of a path which avoids the critical point, along which the correlation functions are regular, and the RG equations everywhere satisfied.

Remark. The universal functions characterizing the behaviour of correlation functions in the critical domain still depend on the normalization of the physical parameters τ , H , M , distances or momenta. Physical quantities that are independent of these normalizations are *universal pure numbers*. In addition to critical exponents, examples are provided by the ratios of the amplitudes of the singularities above and below T_c , like A^+/A^- for the specific heat.

16.6 The N -vector model

We now generalize the results to models in which the order parameter is an N -component vector, and which have symmetries such that the Landau–Ginzburg–Wilson Hamiltonian has still the form of a ϕ^4 -like QFT. We first consider a simple but important example: the $O(N)$ symmetric model and then discuss briefly the general situation.

16.6.1 The $O(N)$ symmetric N -vector model: IR fixed point

The $O(N)$ symmetric effective Hamiltonian has the form

$$\mathcal{H}(\phi) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} (r_c + \tau) \phi^2(x) + \frac{1}{4!} g \Lambda^\varepsilon (\phi^2(x))^2 \right] d^d x, \quad (16.79)$$

in which ϕ is an N -component vector.

Above T_c and in zero field, the RG equations have exactly the same form as in the Ising-like case $N = 1$. The RG β -function has the expansion

$$\beta(g) = -\varepsilon g + \frac{N+8}{48\pi^2} g^2 + O(g^3, g^2\varepsilon). \quad (16.80)$$

At leading order in ε , $\beta(g)$ has a zero g^* which is an IR fixed point:

$$g^* = \frac{48\pi^2}{N+8} \varepsilon + O(\varepsilon^2), \quad (16.81)$$

$$\omega \equiv \beta'(g^*) = \varepsilon + O(\varepsilon^2) \quad (16.82)$$

and, therefore, all the scaling relations derived for $T \geq T_c$ in zero field in Sections 15.5, 15.6 can also be proved for the N -vector model with $O(N)$ symmetry. We give the expressions of the other RG functions at leading order in Section 16.7, by specializing expressions obtained for the general N -vector model.

16.6.2 Correlation functions in a field or below T_c

The addition of a magnetic field term in an $O(N)$ symmetric Hamiltonian has several consequences. First, the magnetization and the magnetic field are now vectors. The scaling forms derived previously apply to the modulus of these vectors. Second, the continuous $O(N)$ symmetry of the Hamiltonian is broken linearly in the dynamical variables by the addition of a magnetic field (see Chapter 13). Since the field and the magnetization distinguish one direction in vector space, there now exist 2^n n -point functions, each spin being either along the magnetization or transverse to it.

As we have shown in Chapter 13, these different correlation functions are related by a set of identities, called Ward–Takahashi (WT) identities, which have been discussed there in a general framework. In the example of the $O(N)$ symmetry, we recall here the simplest one, involving the two-point function at zero momentum, also verified in the mean-field approximation in Section 14.5.3. In terms of $\tilde{\Gamma}_T$, the inverse two-point function transverse to \mathbf{M} at zero momentum or the transverse susceptibility, it reads

$$\tilde{\Gamma}_T(p=0) = \chi_T^{-1} = \frac{H}{M}. \quad (16.83)$$

We recognize, in a different notation, equation (13.43). As we have already discussed in Section 13.4, it follows from this equation that if H goes to 0 below T_c , H/M , and, therefore, $\tilde{\Gamma}_T$ at zero momentum vanish. This last result implies the existence of $(N-1)$ Goldstone modes corresponding to the spontaneous breaking of the $O(N)$ symmetry with a residual $O(N-1)$ symmetry.

Finally, note that the inverse longitudinal two-point function $\tilde{\Gamma}_L(p)$ is singular at zero momentum in zero field below T_c , as one can infer from its Feynman graph expansion (Fig. 16.1). This IR singularity is not generated by critical fluctuations but by the Goldstone modes. It is characteristic of continuous symmetries. It implies that the longitudinal two-point correlation function does not decrease exponentially at large distance.

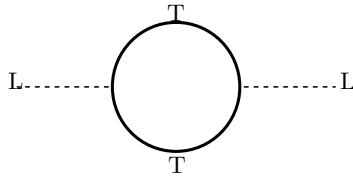


Fig. 16.1 One-loop Goldstone mode contribution to $\tilde{\Gamma}_L$

We discuss this problem more thoroughly in Chapter 19. In particular, the correlation length (16.76) becomes a crossover scale between critical behaviour and Goldstone mode dominated behaviour at larger distances.

Emergent symmetries. Previous results also apply to models in which the Hamiltonian has a symmetry smaller than the group $O(N)$, but is still such that the effective Hamiltonian has the form (16.79), because the quadratic and quartic group invariants are uniquely determined. Then, the $O(N)$ symmetry is dynamically generated in the critical domain. Only a close examination of the leading corrections to the critical behaviour reveals the difference. We have already encountered such a phenomenon: the hypercubic symmetry of the lattice has led to a $O(d)$ continuum space symmetry in the critical domain.

16.7 The general N -vector model

One can find interesting physical systems in which the effective Hamiltonian is not $O(N)$ invariant. A first category consists in systems in which there are several correlation lengths. In such situations generically, when the temperature varies, only one correlation length becomes infinite at a time. Then, the components of the dynamic variables which are non-critical do not contribute to the IR singularities. They can be integrated out in much the same way as the auxiliary fields in Pauli–Villars’s regularization scheme of Section 8.4.2. The effect is to renormalize the effective local Hamiltonian for the critical components. This remark is related to the decoupling theorem of particle physics [47]. Therefore, one can restrict the discussion to theories with only one correlation length.

Models with one correlation length. These models consist in systems in which the Hamiltonian is invariant under a symmetry group, subgroup of $O(N)$, which admits a unique quadratic invariant ϕ^2 but several quartic invariants.

A general Hamiltonian in this case has the form

$$\begin{aligned} \mathcal{H}(\phi) = & \int d^d x \left\{ \frac{1}{2} [(\nabla \phi(x))^2 + (r_c + \tau) \phi^2(x)] \right. \\ & \left. + \frac{\Lambda^\varepsilon}{4!} \sum_{i,j,k,l=1}^N g_{ijkl} \phi_i(x) \phi_j(x) \phi_k(x) \phi_l(x) \right\}. \end{aligned} \quad (16.84)$$

The symmetry implies that the two-point vertex function $\tilde{\Gamma}_{ij}^{(2)}$ is necessarily proportional to the unit matrix:

$$\tilde{\Gamma}_{ij}^{(2)}(p, \tau, g) = \delta_{ij} \tilde{\Gamma}^{(2)}(p, \tau, g), \quad (16.85)$$

and that the symmetric tensor g_{ijkl} has special properties, which in perturbation theory, take the form of successive algebraic conditions (see Section 10.6).

16.7.1 RG equations

In Section 10.6, we have discussed the renormalization of a general ϕ^4 QFT, and derived the corresponding RG equations. We can apply the formalism here [149, 51].

We first sketch the derivation of the RG equations for a multi-component critical theory. Since the field amplitude renormalization constant is independent of the components, the relation between bare and renormalized vertex functions takes the form

$$\tilde{\Gamma}_{r;i_1,i_2,\dots,i_n}^{(n)}(p_k, m_r, g_r, \mu) = Z^{n/2} \tilde{\Gamma}_{i_1,i_2,\dots,i_n}^{(n)}(p_k, \tau, g, \Lambda), \quad (16.86)$$

in which g stands for g_{ijkl} and g_r for $g_{r;ijkl}$.

Differentiating the equation with respect to Λ at g_r , m_r and μ fixed, one obtains the RG equation [149],

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial}{\partial g_{i'j'k'l'}} - \frac{n}{2} \eta(g) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right) \tilde{\Gamma}_{i_1,i_2,\dots,i_n}^{(n)} = 0, \quad (16.87)$$

with the definition

$$\sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial g_{r;ijkl}}{\partial g_{i'j'k'l'}} = -\varepsilon g_{r;ijkl}. \quad (16.88)$$

These equations can be integrated by the same method as before. One introduces scale-dependent coupling constants $g_{ijkl}(\lambda)$ obeying the flow equation

$$\lambda \frac{d}{d\lambda} g_{ijkl}(\lambda) = \beta_{ijkl}(g(\lambda)). \quad (16.89)$$

The large distance properties of such theories are then governed by fixed points, solution of the equations

$$\beta_{ijkl}(g^*) = 0, \quad \forall i, j, k, l. \quad (16.90)$$

The local stability properties of fixed points are governed by the eigenvalues of the matrix

$$M_{ijkl,i'j'k'l'} = \frac{\partial \beta_{ijkl}(g^*)}{\partial g_{i'j'k'l'}}. \quad (16.91)$$

If the real parts of all eigenvalues are positive, the fixed point is locally stable. The global properties depend on the complete solutions of equation (16.89), which determine the basin of attraction in coupling space of each IR stable fixed point. We do not discuss this problem further here and refer to the literature where a number of specific models have been considered (see Refs. [51, 149–152]).

The $\phi_i(x)\phi_j(x)$ insertions. It is also useful to consider correlation functions with $\frac{1}{2}\phi_i(x)\phi_j(x)$ insertions. Their renormalization involves a multiplication of each insertion by the matrix $\zeta_{ij,kl}^{(2)}$. This leads to the RG equation

$$\begin{aligned} & \left[\Lambda \frac{\partial}{\partial \Lambda} + \sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial}{\partial g_{i'j'k'l'}} - \frac{n}{2} \eta(g) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}_{j_1 k_1 \dots j_l k_l, i_1 \dots i_n}^{(l,n)} \\ & - \sum_{m=1}^l \eta_{j_m k_m, b_m c_m}^{(2)} \tilde{\Gamma}_{j_1 k_1 \dots b_m c_m \dots j_l k_l, i_1 \dots i_n}^{(l,n)} = 0, \end{aligned} \quad (16.92)$$

with the definition

$$\eta_{ij,kl}^{(2)} = - \sum_{i',j',k',l'} \beta_{i'j'k'l'} \left(\frac{\partial \zeta^{(2)}}{\partial g_{i'j'k'l'}} \left[\zeta^{(2)} \right]^{-1} \right)_{ij,kl}. \quad (16.93)$$

Since the insertions of ϕ^2 , which are generated by a differentiation with respect to τ , are multiplicatively renormalized, the matrix $\eta_{ij,kl}^{(2)}(g)$ has δ_{kl} as eigenvector, and

$$\sum_k \eta_{ij,kk}^{(2)}(g) = \eta_2(g) \delta_{ij}. \quad (16.94)$$

RG functions. At one-loop order, the RG functions are given by (equations (10.86) and (10.87))

$$\beta_{ijkl} = -\varepsilon g_{ijkl} + \frac{1}{16\pi^2} \sum_{m,n} (g_{ijmn} g_{mnkl} + 2 \text{ terms}) + O(g^3), \quad (16.95)$$

$$\eta(g) \delta_{ij} = \frac{1}{24} \frac{1}{(8\pi^2)^2} \sum_{k,l,m} g_{iklm} g_{jklm} + O(g^3), \quad (16.96)$$

$$\eta_{ij,kl}^{(2)} = -\frac{1}{16\pi^2} g_{ijkl} + O(g^2). \quad (16.97)$$

The condition (16.94) then implies (equations (10.68) and (10.69))

$$\sum_k g_{ijkk} = \gamma_1 \delta_{ij}, \quad \sum_{k,l,m} g_{iklm} g_{jklm} = \gamma_2 \delta_{ij}, \quad (16.98)$$

and summing over i, j , one can rewrite equation (16.96) as

$$\eta(g) = \frac{1}{(8\pi^2)^2} \frac{\gamma_2}{24} + O(g^3, g^2 \varepsilon). \quad (16.99)$$

16.7.2 Stability of the $O(N)$ symmetric fixed point

Among the possible fixed points, one always finds, in addition to the trivial Gaussian fixed point, the $O(N)$ symmetric fixed point. We can study its local stability at leading order in ε . We first specialize the expressions (16.95–16.97) to the case of the $(\phi^2)^2$ QFT with $O(N)$ symmetry. We then have to substitute

$$g_{ijkl} = \frac{g}{3} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}). \quad (16.100)$$

After a short calculation, the expression (16.80) of the β -function is recovered and, in addition,

$$\eta(\tilde{g}) = \frac{(N+2)}{72} \tilde{g}^2 + O(\tilde{g}^3), \quad (16.101)$$

where the notation (10.64) has been used:

$$\tilde{g} = N_d g, \quad N_d = 2(4\pi)^{-d/2} / \tilde{\Gamma}(d/2).$$

Introducing the identity matrix \mathbf{I} and the projector \mathbf{P} ,

$$I_{ij,kl} = \frac{1}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (16.102)$$

$$P_{ij,kl} = \frac{1}{N} \delta_{ij} \delta_{kl}, \quad (16.103)$$

one can express the matrix $\eta_{ij,kl}^{(2)}$ as

$$\eta^{(2)} = -(N\mathbf{P} + 2\mathbf{I}) \frac{\tilde{g}}{6} + O(\tilde{g}^2). \quad (16.104)$$

The trace of the matrix $\eta^{(2)}$ yields $\eta_2(g)$. The second eigenvalue of the matrix $\eta'_2(g)$, given by its traceless part, corresponds to a symmetry breaking mass term, and as we have discussed at the beginning of this section, describes the crossover from a situation with one correlation length to a situation in which some components of the order parameter decouple. It is traditional to introduce a new function $\varphi(g)$ and to parametrize it as

$$\eta'_2(g) = \frac{\varphi(g)}{\nu(g)} - 2. \quad (16.105)$$

The fixed point value $\varphi = \varphi(g^*)$ is called the *crossover exponent* [153]. Finally, the stability conditions are given by the eigenvalues of the matrix (16.91). Setting

$$g_{ijkl} = g_{ijkl}^* + s_{ijkl}, \quad (16.106)$$

at leading order one finds

$$(Ms)_{ijkl} = -\varepsilon s_{ijkl} + \frac{\varepsilon}{N+8}(\delta_{ij}s_{mmkl} + 5 \text{ terms} + 12s_{ijkl}). \quad (16.107)$$

Taking s_{ijkl} proportional to g_{ijkl}^* , one recovers the exponent ω . More generally, the eigenvectors can be classified according to their trace properties. We expand

$$s_{ijkl} = ug_{ijkl}^* + (v_{ij}\delta_{kl} + 5 \text{ terms}) + w_{ijkl}, \quad (16.108)$$

in which the tensors v_{ij} and w_{ijkl} are traceless:

$$\sum_i v_{ii} = 0, \quad \sum_k w_{ijkk} = 0. \quad (16.109)$$

The three eigenvalues corresponding to u , w , and v are, respectively,

$$\omega = \varepsilon + O(\varepsilon^2), \quad \omega_{\text{anis.}} = \varepsilon \frac{4-N}{N+8} + O(\varepsilon^2), \quad \omega' = \frac{8\varepsilon}{N+8} + O(\varepsilon^2). \quad (16.110)$$

The perturbation proportional to v_{ij} does not satisfy the trace condition (16.98). Therefore, it lifts the degeneracy between the correlation lengths of the different components of the order parameter. It induces a crossover to a situation in which some components decouple. However, one easily verifies that the corresponding eigenvalue ω' leads to effects subleading for ε small with respect to the eigenvalue $\eta'_2(g^*)$. Within the class of interactions satisfying equation (16.98), the relevant eigenvalue is $\omega_{\text{anis.}}$. We find the very interesting result that the $O(N)$ symmetric fixed point is stable against any perturbation for N smaller than some value N_c [149, 51]. This is an example of *emergent symmetry*: correlation functions in the critical domain have a larger symmetry than microscopic correlation functions. The calculation of $\omega_{\text{anis.}}$ at order ε , in models with cubic anisotropy [150], yields

$$N_c = 4 - 2\varepsilon + O(\varepsilon^2). \quad (16.111)$$

16.7.3 Gradient flow

Denoting g_α a set of variables (here the set of coupling constants) parametrizing a manifold, the flow equation

$$\beta_\alpha(g) = \sum_\beta T_{\alpha\beta}(g) \frac{\partial U}{\partial g_\beta}, \quad (16.112)$$

in which the function U is the potential, defines a gradient flow for g_α if $T_{\alpha\beta}$ is a symmetric and positive matrix. The form of the equation is invariant under a change of parametrization (more precisely a diffeomorphism) of the manifold.

It can be verified that the β -function, at two-loop order as given by (16.95), can indeed be written as [154]

$$\beta_{ijkl} = \frac{\partial U(g)}{\partial g_{ijkl}},$$

with

$$\begin{aligned} U(g) = & -\frac{1}{2}\varepsilon \sum_{i,j,k,l} g_{ijkl}g_{ijkl} + \frac{1}{(4\pi)^2} \sum_{i,j,k,l,m,n} g_{ijkl}g_{klmn}g_{mnij} \\ & + \frac{1}{(4\pi)^4} \sum_{i,j,k,l,m,n,p,q} \left(\frac{3}{2}g_{ijkl}g_{ijmn}g_{pqkm}g_{pqln} + \frac{1}{12}g_{ijkl}g_{ijkm}g_{npql}g_{npqm} \right). \end{aligned} \quad (16.113)$$

At three-loop order, the RG flow equation for g_{ijkl} takes the more general form (16.112), where the matrix T is positive, at least for all g_{ijkl} small.

The structure of gradient flow implies that the RG flow follows curves of monotonous decrease of the potential U , the fixed points are extrema of the function U , and the stable fixed point corresponds to the lowest value of the potential. Moreover, the matrix of derivatives of the β -function at a fixed point is symmetric, and thus has real eigenvalues.

At leading order in ε , a fixed point g^* is solution of the equation

$$\begin{aligned} \varepsilon g_{ijkl}^* &= \frac{1}{16\pi^2} \sum_{m,n} (g_{ijmn}^*g_{mnkl}^* + \text{2 terms}) \\ \Rightarrow \varepsilon \sum_{i,j,k,l} g_{ijkl}^*g_{ijkl}^* &= \frac{3}{(4\pi)^2} \sum_{i,j,k,l,m,n} g_{ijkl}^*g_{klmn}^*g_{mnij}^* \end{aligned}$$

and, therefore, the value of the potential can then be written as

$$U(g^*) = -\frac{1}{6}\varepsilon \sum_{i,j,k,l} g_{ijkl}^*g_{ijkl}^* + O(\varepsilon^4).$$

Comparing with the expression (16.99) for the RG function $\eta(g)$, we find the relation between the potential at the fixed point and the corresponding exponent,

$$\eta = \frac{1}{6N} \frac{1}{(4\pi)^4} \sum_{i,j,k,l} g_{ijkl}^*g_{ijkl}^* + O(\varepsilon^3) = -\frac{1}{N\varepsilon} \frac{1}{(4\pi)^4} U(g^*) + O(\varepsilon^3).$$

Therefore, at leading order in ε , the stable fixed point, which corresponds to the lowest value of the potential, also corresponds to the largest value of the exponent η . Numerical estimates suggest that this property remains true beyond the ε -expansion [152].

16.8 Asymptotic expansion of the two-point function

In the critical domain, when points are separated by distances much smaller than the correlation length ξ , the correlation functions tend towards the correlation functions of the critical theory ($T = T_c$), for example,

$$\tilde{\Gamma}^{(2)}(p) \underset{\xi^{-1} \ll p \ll \Lambda}{\propto} p^{2-\eta}. \quad (16.114)$$

The right-hand side is actually the first term of an asymptotic expansion in the variable $p\xi$ for $p\xi$ large. The leading term has been obtained by using the property that, at large non exceptional momenta, the derivative $\partial\tilde{\Gamma}^{(n)}(p_1, \dots, p_n)/\partial\tau$ is asymptotically negligible with respect to $\tilde{\Gamma}^{(n)}(p_1, \dots, p_n)$. However, since

$$\frac{\partial}{\partial\tau}\tilde{\Gamma}^{(n)}(p_1, \dots, p_n) = \tilde{\Gamma}^{(1,n)}(0; p_1, \dots, p_n),$$

the derivative $\partial\tilde{\Gamma}^{(n)}/\partial\tau$ cannot be evaluated with the same method, because the momenta are exceptional. As we have explained in Section 11.3, it is necessary to use the SDE of operator products.

16.8.1 Asymptotic expansion from SDE

We focus on the two-point function. We have to evaluate $\tilde{\Gamma}^{(1,2)}(0; p, -p)$. However, we cannot apply directly the SDE to $\tilde{\Gamma}^{(1,2)}$ because this would involve $\tilde{\Gamma}^{(2,0)}$, which requires additional renormalizations.

Therefore, we differentiate once more with respect to τ ,

$$\frac{\partial^2}{(\partial\tau)^2}\tilde{\Gamma}^{(2)}(p) = \tilde{\Gamma}^{(2,2)}(0, 0; p, -p). \quad (16.115)$$

To $\tilde{\Gamma}^{(2,2)}$ we can now apply the SDE,

$$\tilde{\Gamma}^{(2,2)}(0, 0; p, -p) \sim B(p)\tilde{\Gamma}^{(3,0)}(0, 0, 0), \quad \text{for } \xi^{-1} \ll p \ll \Lambda. \quad (16.116)$$

As shown in Section 11.4, $B(p)$ satisfies an RG equation that can be obtained by applying the differential operator

$$D \equiv \Lambda \frac{\partial}{\partial\Lambda} + \beta(g) - \left(\frac{1}{\nu(g)} - 2 \right) \tau \frac{\partial}{\partial\tau}, \quad (16.117)$$

on both sides of equation (16.116) and using the RG equations (16.45).

One finds

$$[D + \nu^{-1}(g) - 2 - \eta(g)] B(p) \sim 0. \quad (16.118)$$

For $\tau = 0$ and $g = g^*$, the equation becomes

$$\left(\Lambda \frac{\partial}{\partial\Lambda} + \frac{1}{\nu} - 2 - \eta \right) B(p) \sim 0. \quad (16.119)$$

Equation (16.116) shows also that $B(p)$ has canonical dimension ε . It follows that

$$B(p) \propto \Lambda^\varepsilon (p/\Lambda)^{2-\eta-(1-\alpha)/\nu}. \quad (16.120)$$

A differentiation of equation (16.49) with respect to τ yields

$$\tilde{\Gamma}^{(3,0)}(0,0,0) \sim \Lambda^{-\varepsilon}(\tau/\Lambda^2)^{-1-\alpha}. \quad (16.121)$$

Finally, integrating equation (16.115) twice with respect to τ , and using the set of equations (16.116), (16.120) and (16.121), one obtains an expansion of the form [155]

$$\tilde{\Gamma}^{(2)}(p)_{\xi^{-1} \ll p \ll \Lambda} = p^{2-\eta}(a + b\tau p^{-1/\nu} + c\tau^{1-\alpha} p^{-(1-\alpha)/\nu} + \dots), \quad (16.122)$$

a result partially anticipated on physical grounds [156]. Successive corrections to expression (16.122) can then be obtained by using systematically the SDE beyond leading order.

Note that the effect of the differentiation with respect to τ has been simply to generate the regular terms in the temperature that have an order in τ comparable to the singular terms for $\varepsilon \rightarrow 0$.

16.8.2 Next to leading terms in a field or below T_c

It is also possible to obtain expressions in a field or below T_c by expanding correlation functions in powers of the magnetization and applying the SDE to each term. The results now differ between Ising-like systems and the N -vector model. For Ising-like systems, one finds [155],

$$\tilde{\Gamma}^{(2)}(p, \tau, M) = p^{2-\eta} \left[a + b\tau/p^{1/\nu} + G_1(\tau/M^{1/\beta})(p/M^{\nu/\beta})^{-(1-\alpha)/\nu} + \dots \right], \quad (16.123)$$

in which the function G_1 can be related to the free energy and thus the equation of state by

$$G_1(x) = \int_1^\infty ds s^{\delta-1/\beta} \left[f'(0) - f'(x/s^{1/\beta}) \right] + \frac{f'(0)}{\delta - (1/\beta) + 1}. \quad (16.124)$$

In the $O(N)$ symmetric case, the SDE involves the second operator of dimension 2,

$$\mathcal{O}_{ij}[\phi(x)] = \phi_i(x)\phi_j(x) - \frac{\delta_{ij}}{N}\phi^2(x). \quad (16.125)$$

This operator is multiplicatively renormalizable. Correlation functions with the insertion of \mathcal{O}_{ij} satisfy the RG equations,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) - \left(\frac{\varphi(g)}{\nu(g)} - 2 \right) \right] \tilde{\Gamma}_{\mathcal{O}_{ij}}^{(n)}(p_i; g, \Lambda) = 0, \quad (16.126)$$

in which the RG function $\varphi(g)$ has been defined in equation (16.105). As a consequence, a new term is present at the same order in the asymptotic expansion of the two-point function, which becomes [155],

$$\begin{aligned} \tilde{\Gamma}^{(2)}(p, \tau, M) = & p^{2-\eta} \left\{ \left[a + b\tau/p^{1/\nu} + G_1(\tau/M^{1/\beta})(p/M^{\nu/\beta})^{-(1-\alpha)/\nu} \right] \delta_{ij} \right. \\ & \left. + G_2(\tau/M^{1/\beta})(p/M^{\nu/\beta})^{-d+\varphi/\nu} \left(\frac{\delta_{ij}}{N} - \frac{M_i M_j}{M^2} \right) \right\} + \dots, \end{aligned} \quad (16.127)$$

in which φ is the crossover exponent, and G_2 a new universal function which may be calculated in an ε -expansion. At order ε , it is given by

$$G_2(x) = 1 + \frac{\varepsilon}{2(N+8)} [(x+3)\ln(x+3) - (x+1)\ln(x+1)] + O(\varepsilon^2). \quad (16.128)$$

16.9 Some universal quantities as ε expansions

Many models with second order phase transitions have been investigated, the effective QFTs identified and then various universal quantities calculated by field theoretical methods. We can report here only a limited number of significant results. Therefore, we consider only the important example of the $O(N)$ symmetric N -vector model described by the Hamiltonian (16.79),

$$\mathcal{H}(\phi) = \int \left\{ \frac{1}{2} [\nabla \phi(x)]^2 + \frac{1}{2} (r_c + \tau) \phi^2(x) + \frac{1}{4!} g \Lambda^\varepsilon (\phi^2(x))^2 \right\} d^d x,$$

in which ϕ is an N -component field.

We describe results for critical exponents, the equation of state, and a few amplitude ratios. We discuss more thoroughly critical exponents, because they make a detailed and precise comparison between QFT, other theoretical methods and experiments possible.

In QFT, universal quantities have been calculated by two methods: the ε expansion, initiated by Wilson and Fisher [75], which we have systematically discussed in previous chapters, and perturbation theory at fixed dimension (see Sections 16.1 and 41.3.1). In both cases, the expansion is divergent for all values of the expansion parameter. The rate of divergence can be obtained from instanton calculus, as we explain in Chapter 40. There exist methods to deal with divergent series (see Sections 41.1–41.2). In the case of the ϕ^4 QFT, methods based on the Borel transformation has been extensively used (Chapter 40).

Initially, many quantities have been calculated up to order ε^2 by elementary techniques [75, 157]. The order ε^3 has required more involved QFT techniques [158]. Later, a series of ingenious tricks, and the systematic use of computer algebra, within the framework of dimensional regularization and the minimal subtraction scheme (see Chapter 10), have led to a determination of critical exponents in the N -vector model, up to five [159], then six-loop order for all N [160], and even seven loops for $N = 1$ [161].

16.9.1 RG functions. Critical exponents

Although the RG functions of the $(\phi^2)^2$ theory and, therefore, the critical exponents are known up to six-loop order for the $O(N)$ model and up to seven loops for $N = 1$, we give here, for illustration, only three successive terms in the expansion, referring to the literature for higher order results. In terms of the variable

$$\tilde{g} = N_d g, \quad N_d = \frac{2}{(4\pi)^{d/2} \Gamma(d/2)}, \quad (16.129)$$

the RG functions $\beta(\tilde{g})$ and $\eta_2(\tilde{g})$ at three-loop order, $\eta(\tilde{g})$ at four-loop order are

$$\begin{aligned} \beta(\tilde{g}) = & -\varepsilon \tilde{g} + \frac{(N+8)}{6} \tilde{g}^2 - \frac{(3N+14)}{12} \tilde{g}^3 \\ & + \frac{[33N^2 + 922N + 2960 + 96(5N+22)\zeta(3)]}{12^3} \tilde{g}^4 + O(\tilde{g}^5), \end{aligned} \quad (16.130)$$

$$\eta(\tilde{g}) = \frac{(N+2)}{72} \tilde{g}^2 \left[1 - \frac{(N+8)}{24} \tilde{g} + \frac{5(-N^2 + 18N + 100)}{576} \tilde{g}^2 \right] + O(\tilde{g}^5), \quad (16.131)$$

$$\eta_2(\tilde{g}) = -\frac{(N+2)}{6} \tilde{g} \left[1 - \frac{5}{12} \tilde{g} + \frac{(5N+37)}{48} \tilde{g}^2 \right] + O(\tilde{g}^4), \quad (16.132)$$

in which $\zeta(s)$ is Riemann's ζ -function, $\zeta(3) = 1.20205690315\dots$.

The zero $\tilde{g}^*(\varepsilon)$ of the β -function then is

$$\begin{aligned}\tilde{g}^*(\varepsilon) = & \frac{6\varepsilon}{(N+8)} \left[1 + \frac{3(3N+14)}{(N+8)^2} \varepsilon + \left(\frac{18(3N+14)^2}{(N+8)^4} \right. \right. \\ & \left. \left. - \frac{33N^2 + 922N + 2960 + 96(5N+22)\zeta(3)}{8(N+8)^3} \right) \varepsilon^2 \right] + O(\varepsilon^4).\end{aligned}\quad (16.133)$$

The values of the critical exponents η , γ and the correction exponent ω ,

$$\eta = \eta(\tilde{g}^*), \quad \gamma = \frac{2-\eta}{2+\eta_2(\tilde{g}^*)}, \quad \omega = \beta'(\tilde{g}^*),$$

follow:

$$\begin{aligned}\eta = & \frac{\varepsilon^2(N+2)}{2(N+8)^2} \left\{ 1 + \frac{(-N^2 + 56N + 272)}{4(N+8)^2} \varepsilon + \frac{1}{16(N+8)^4} [-5N^4 - 230N^3 \right. \\ & \left. + 1124N^2 + 17920N + 46144 - 384(5N+22)(N+8)\zeta(3)] \varepsilon^2 \right\} + O(\varepsilon^5),\end{aligned}\quad (16.134)$$

$$\begin{aligned}\gamma = & 1 + \frac{(N+2)}{2(N+8)} \varepsilon + \frac{(N+2)}{4(N+8)^3} (N^2 + 22N + 52) \varepsilon^2 + \frac{(N+2)}{8(N+8)^5} \\ & \times [N^4 + 44N^3 + 664N^2 + 2496N + 3104 - 48(5N+22)(N+8)\zeta(3)] \varepsilon^3 \\ & + O(\varepsilon^4),\end{aligned}\quad (16.135)$$

$$\begin{aligned}\omega = & \varepsilon - \frac{3(3N+14)}{(N+8)^2} \varepsilon^2 + \frac{[33N^2 + 922N + 2960 + 96(5N+22)\zeta(3)]}{4(N+8)^3} \varepsilon^3 \\ & - 18 \frac{(3N+14)^2}{(N+8)^4} \varepsilon^3 + O(\varepsilon^4).\end{aligned}\quad (16.136)$$

All other exponents can be obtained from the scaling relations derived in Sections 16.3–16.5. Note that the β function at order g^4 involves $\zeta(3)$. At higher orders $\zeta(5)$ and $\zeta(7)$ successively appear. In Table 40.3, we give the values of the critical exponents γ and η obtained by simply adding the successive terms of the ε expansion for $\varepsilon = 1$ and $N = 1$. We immediately observe a striking phenomenon: the sums first seem to settle near some reasonable value, and then begin to diverge with increasing oscillations.

We argue in Chapter 40 that the ε expansion is divergent for all values of ε .

Remark. The definition of the β function by minimal subtraction has an intrinsic meaning, unlike other definitions, since then $\beta(g)$ can be expressed in terms of $\omega(\varepsilon)$. Setting

$$\beta(g) = g(-\varepsilon + s(g)),$$

then

$$g(s) = \frac{s}{\beta_2} \exp \left[\int_0^s ds' \left(\frac{1}{\omega(s')} - \frac{1}{s'} \right) \right].$$

16.9.2 The scaling equation of state

The scaling equation of state provides an interesting example of a universal function. Its ε -expansion has been obtained up to order ε^2 for arbitrary N [147], and order ε^3 for $N = 1$ [162]. We set

$$H = M^\delta f(x = \tau/M^{1/\beta}), \quad (16.137)$$

in which the normalizations of x and the function $f(x)$ are such that

$$f(0) = 1, \quad f(-1) = 0. \quad (16.138)$$

It is also convenient to set

$$y = x + 1, \quad z = x + 3, \quad \rho = z/4y. \quad (16.139)$$

The expansion up to order ε^2 is then

$$f(x) = 1 + x + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + O(\varepsilon^3), \quad (16.140)$$

with

$$f_1(x) = \frac{1}{2(N+8)} [(N-1)y \ln y + 3z \ln z - 9y \ln 3 + 6x \ln 2], \quad (16.141a)$$

$$\begin{aligned} f_2(x) = & \left[\frac{1}{2(N+8)} \right]^2 \{ [N-1+6 \ln 2 - 9 \ln 3 + (N-1) \ln y] [3z \ln z + (N-1)y \ln y \right. \\ & + 6x \ln 2 - 9y \ln 3] + \frac{1}{2}(10-N)y (\ln^2 z - \ln^2 3) + 36 (\ln^2 z - y \ln^2 3 + x \ln^2 2) \\ & - 54 \ln 2 (\ln z + x \ln 2 - y \ln 3) + 3 \ln \frac{27}{4} (N-1)y \ln y + \frac{-4N^2 + 17N + 212}{N+8} \\ & \times [z \ln z + 2x \ln 2 - 3y \ln 3] + (N-1)y \ln y \ln z - \frac{1}{2}N(N-1)y \ln^2 y \\ & + \frac{N-1}{N+8} (19N+92)y \ln y - 2(N-1) [(x+6)J_1(x) - 6yJ_1(0)] \\ & \left. - 6(N-1) [J_2(x) - yJ_2(0)] + 4(N-1) [J_3(x) - yJ_3(0)] \right\}, \end{aligned} \quad (16.141b)$$

where

$$J_i(x) = I_i(\rho), \quad (16.142)$$

and

$$I_1(\rho) = \int_0^\infty \frac{du \ln u}{u(1-u)} \left[(1-u/\rho)^{1/2} \theta(\rho-u) - 1 \right], \quad (16.143)$$

$$I_2(\rho) = \rho \frac{d}{d\rho} I_1(\rho), \quad (16.144)$$

$$I_3(\rho) = I_1(\rho) + 2I_2(\rho). \quad (16.145)$$

The expressions (16.141) are not uniform, and valid only for x of order 1. For x large, that is, for small magnetization M , the magnetic field has a regular expansion in odd powers of M , that is, in the variable $x^{-\beta}$ (Section 16.5). Therefore, it is convenient to introduce Josephson's parametrization [163], which leads to a representation uniform in both limits.

16.9.3 Parametric representation of the equation of state

We set

$$x = x_0 (1 - \theta^2) \theta^{-1/\beta}, \quad \theta > 0, \quad (16.146)$$

where x_0 is an arbitrary positive constant. More directly, we can parametrize M and t in terms of two variables R and θ , setting

$$\begin{aligned} M &= R^\beta \theta, \\ \tau &= x_0 R (1 - \theta^2), \\ H &= R^\beta h(\theta). \end{aligned} \quad (16.147)$$

Then, the function

$$h(\theta) = \theta^\delta f(x(\theta)) \quad (16.148)$$

is an odd function of θ regular near $\theta = 1$, which is x small, and near $\theta = 0$ which is x large. For the special choice

$$x_0 = 3(3/2)^{1/2\beta-1}, \quad (16.149)$$

the equation of state at order ε takes the rather simple form

$$h(\theta) = \theta (3 - 2\theta^2) \left[1 + \frac{\varepsilon(N-1)}{2(N+8)} \ln(3 - 2\theta^2) \right] + O(\varepsilon^2). \quad (16.150)$$

For $N = 1$, the equation is especially simple and corresponds to the so-called linear parametric model in which $h(\theta)$ is a cubic odd function of θ . One verifies that it is still possible to adjust x_0 at order ε^2 to preserve this form. However, at order ε^3 , which is also known for $N = 1$, the introduction of a term proportional to θ^5 becomes necessary. One finds

$$h(\theta) = h_0 \theta (b^2 - \theta^2) (1 + c\theta^2) + O(\varepsilon^4), \quad (16.151)$$

in which h_0 is the field normalization constant, and b, c are given by

$$b^2 = \frac{3}{2} \left(1 - \frac{\varepsilon^2}{12} \right), \quad c = -\frac{\varepsilon^3}{18} \left(\zeta(3) + \frac{I-1}{4} \right), \quad (16.152)$$

with

$$I = \int_0^1 dx \frac{\ln[x(1-x)]}{1-x(1-x)} = \frac{4}{9}\pi^2 - \frac{2}{3}\psi'(1/3) = -2.3439072386\dots. \quad (16.153)$$

The constant x_0 is given by

$$x_0 = b^{1/\beta} / (b^2 - 1). \quad (16.154)$$

Remark. In the case $N > 1$, the function $h(\theta)$ has still a singularity on the coexistence curve, due to the presence of Goldstone modes in the ordered phase. The nature of this singularity can be obtained from the study of the non-linear σ -model, in Chapter 19. We show that the behaviour of correlation functions below T_c in a theory with a spontaneously broken continuous symmetry is governed by the zero temperature IR fixed point. Therefore, the coexistence curve singularities can be obtained from a low temperature expansion (for details, see Sections 19.11 and 19.12).

In all cases, as already stated, the essential property of the parametric representation is that it automatically satisfies the different requirements about the regularity properties of the equation of state and leads to uniform approximations.

The comparison with the numerical results for the Ising model ($N = 1$) and the Heisenberg model $N = 3$ in three dimensions shows that the successive ε and ε^2 corrections improve the mean-field approximation.

From the parametric representation of the equation of state, it is also possible to derive a representation for the singular part of the free energy density. Setting

$$F(M, \tau) \equiv \Omega^{-1} \Gamma_{\text{sing.}}(M, \tau) = R^{2-\alpha} g(\theta), \quad (16.155)$$

(Ω is the volume) one finds for $g(\theta)$ the differential equation

$$h(\theta) (1 - \theta^2 + 2\beta\theta^2) = 2(2 - \alpha)\theta g(\theta) + (1 - \theta^2) g'(\theta). \quad (16.156)$$

The integration constant is obtained by requiring the regularity of $g(\theta)$ at $\theta = 1$. Note that if one expands

$$h(\theta) (1 - \theta^2 + 2\beta\theta^2) = X_0 + X_1(1 - \theta^2) + X_2(1 - \theta^2)^2 + O((1 - \theta^2)^3),$$

then for $\alpha \rightarrow 0$

$$g(\theta) \sim -\frac{X_2}{2\alpha}(1 - \theta^2)^2. \quad (16.157)$$

In the same way, the inverse magnetic susceptibility is given by

$$\chi^{-1} = R^\gamma g_2(\theta), \quad (16.158)$$

with

$$g_2(\theta) (1 - \theta^2 + 2\beta\theta^2) = 2\beta\delta\theta h(\theta) + (1 - \theta^2) h'(\theta). \quad (16.159)$$

In particular, these expressions can be used to calculate various universal ratios of amplitudes.

16.9.4 Amplitude ratios

Some simple universal numbers, apart from critical exponents, have been calculated: ratios of amplitudes of singularities near T_c [164–167]. We first consider two examples which can be derived directly from the equation of state.

The specific heat. The singular part of the specific heat C_H , that is, the $\phi^2(x)$ two-point correlation function at zero momentum, behaves near T_c as

$$C_H = A^\pm |\tau|^{-\alpha}, \quad \tau = T - T_c \rightarrow \pm 0. \quad (16.160)$$

The ratio A^+/A^- is universal. It is directly related to the function $g(\theta)$ defined by equation (16.156):

$$\frac{A^+}{A^-} = (b^2 - 1)^{2-\alpha} \frac{g(0)}{g(b)}. \quad (16.161)$$

At order ε^2 , one finds

$$\begin{aligned} \frac{A^+}{A^-} &= 2^{\alpha-2} N \left\{ 1 + \varepsilon + \left[3N^2 + 26N + 100 + (4 - N)(N - 1)\zeta(2) \right. \right. \\ &\quad \left. \left. - 6(5N + 22)\zeta(3) - 9(4 - N)\lambda \right] \frac{\varepsilon^2}{2(N + 8)^2} \right\} + O(\varepsilon^3), \end{aligned} \quad (16.162)$$

with

$$\zeta(2) = \pi^2/6 = 1.64493406684\dots,$$

while λ is defined in terms of the integral I given in equation (16.153):

$$\lambda = -I/2 = \frac{1}{3}\psi'(1/3) - \frac{2}{9}\pi^2 = 1.17195361934\dots.$$

The evaluation (16.157) shows that, for α small, this is a poor representation since $A^+/A^- = 1 + O(\alpha)$. A better representation then is (we give only the two first terms),

$$\frac{A^+}{A^-} = 2^\alpha (1 - K\alpha/\varepsilon), \quad \text{with} \quad K = \frac{1}{2}(N+8) + \frac{N^2 + 4N + 28}{2(N+8)}\varepsilon + O(\varepsilon^2).$$

The magnetic susceptibility. The magnetic susceptibility in zero field can also be calculated from the function $g_2(\theta)$ defined by equation (16.159). As we know, below T_c , the susceptibility diverges for systems with Goldstone modes. We restrict ourselves, therefore, to $N = 1$. Defining

$$\chi = C^\pm |\tau|^{-\gamma}, \quad \text{for } \tau \rightarrow \pm 0, \quad (16.163)$$

one obtains

$$\frac{C^+}{C^-} = \frac{2(1+cb^2)(b^2-1)^{1-\gamma}}{[1-b^2(1-2\beta)]} \quad (16.164a)$$

$$= \frac{2^{\gamma+1}}{6\beta-1} \left[1 + \left(\frac{2\lambda+1}{4} - \zeta(3) \right) \frac{\varepsilon^3}{12} \right] + O(\varepsilon^4). \quad (16.164b)$$

The ratio C^+/C^- can be expressed, at order ε^2 , entirely in terms of critical exponents. This form follows naturally from the parametric representation of the equation of state. The ε^3 relative correction is of the order of only 3%.

The correlation length. We define here the correlation length in terms of the ratio of the two first moments of the two-point correlation function:

$$\tilde{\Gamma}^{(2)}(p) = \tilde{\Gamma}^{(2)}(0) (1 + \xi_1^2 p^2) + O(p^4). \quad (16.165)$$

The function ξ_1^2 has the scaling form

$$\xi_1^2(M, \tau) = M^{-2\nu/\beta} f_\xi(\tau/M^{1/\beta}). \quad (16.166)$$

Otherwise, it shares all the properties of the equation of state. It can be written in parametric form as

$$\xi_1^2(M, \tau) = R^{-2\nu} g_\xi(\theta). \quad (16.167)$$

At order ε , for $N = 1$, for example, one finds

$$g_\xi(\theta) = g_\xi(0) \left(1 - \frac{5}{18} \varepsilon \theta^2 \right) + O(\varepsilon^2). \quad (16.168)$$

Setting in zero field,

$$\xi_1 = f_1^\pm |\tau|^{-\nu}, \quad \text{for } \tau \rightarrow \pm 0, \quad (16.169)$$

quantity that exists only for $N = 1$, one can use the determination of g_ξ to calculate the ratio

$$f_1^+/f_1^- = 2^\nu \left[1 + \frac{5}{24} \varepsilon + \frac{1}{432} \left(\frac{295}{24} + 2I \right) \varepsilon^2 \right] + O(\varepsilon^3), \quad (16.170)$$

in which the constant I is given by equation (16.153).

An additional universal constant. To the relation between exponents,

$$2 - \alpha = d\nu$$

is associated a universal combination, which involves only amplitudes of singularities when T_c is approached from above,

$$R_\xi^+ = f_1^+ (\alpha A^+)^{1/d}. \quad (16.171)$$

Indeed, from the definitions (16.155) and (16.169), one infers

$$(R_\xi^+)^d = (1 - \alpha)(2 - \alpha)\tau^{\alpha-2}F(0, \tau)\tau^{\nu d}(\xi_1)^d = (1 - \alpha)(2 - \alpha)F(0, \tau)(\xi_1)^d,$$

where the last product is normalization independent. The ε -expansion of R_ξ^+ is

$$(R_\xi^+)^d = \sigma_d \frac{N}{2} \nu (1 - \alpha) \left[1 + \eta \left(\frac{-11}{2} + \frac{14}{3} \lambda \right) \right] + O(\varepsilon^3), \quad (16.172)$$

with

$$\sigma_d = \Gamma(1 + \varepsilon/2)\Gamma(1 - \varepsilon/2)N_d$$

(the loop factor N_d has been defined in equation (16.129)).

Other universal ratios. Clearly, it is possible to define an infinite number of other universal ratios. We give here a few other examples, which have been considered in the literature. Let us first define some additional amplitudes. On the critical isotherm, the correlation length behaves as

$$\xi_1 = f_1^c / H^{2/(d+2-\eta)}, \quad (16.173)$$

the magnetic susceptibility as

$$\chi = C^c / H^{1-1/\delta}; \quad (16.174)$$

the spontaneous magnetization vanishes as

$$M = B(-\tau)^\beta, \quad (16.175)$$

and the spin–spin correlation function in momentum space at T_c behaves as

$$\chi(p) = \left[\tilde{\Gamma}^{(2)}(p) \right]^{-1} = D p^{\eta-2}. \quad (16.176)$$

One can then define the following universal ratio,

$$R_c = \alpha A^+ C^+ / B^2, \quad (16.177)$$

which corresponds to the relation between exponents

$$\alpha + 2\beta + \gamma = 2.$$

Indeed, using this relation, one verifies that R_c is proportional to $F(0, \tau)M^{-2}\chi$, which is normalization independent. The ε -expansion of R_c is,

$$R_c = \frac{N}{N+8} 2^{-2\beta-1} \varepsilon \left[1 + \left(1 - \frac{30}{(N+8)^2} \right) \varepsilon \right] + O(\varepsilon^3). \quad (16.178)$$

One can also construct the three following combinations [165]:

$$Q_1 = C^c \delta / (B^{\delta-1} C^+)^{1/\delta}, \quad (16.179)$$

$$Q_2 = (f_1^c/f_1^+)^{2-\eta} C^+/C^c, \quad (16.180)$$

$$Q_3 = D (f_1^+)^{2-\eta} / C^+, \quad (16.181)$$

which correspond to the relations $\gamma = \beta(\delta - 1)$, the explicit expression of δ and $\gamma = \nu(2 - \eta)$. Moreover, Q_1 and Q_3 are normalization independent, because $H\chi/M$ and $p\xi$, respectively, are. For Q_2 , this property follows immediately from the definition. Thus, all three quantities are *universal*.

The quantity Q_1 is related to R_χ defined in Ref. [166],

$$R_\chi = Q_1^{-\delta}. \quad (16.182)$$

Their ε -expansions can be written as

$$R_\chi = 3^{(\delta-3)/2} 2^{\gamma+(1-\delta)/2} \left[1 + \left(\frac{2\lambda+1}{4} - \zeta(3) \right) \frac{\varepsilon^3}{18} \right] + O(\varepsilon^4), \quad (16.183)$$

$$Q_2 = 1 + \frac{\varepsilon}{18} + \left(\frac{23}{9} + \frac{4}{3}\lambda \right) \frac{\varepsilon^2}{54} + O(\varepsilon^3), \quad (16.184)$$

$$Q_3 = 1 - \left(\frac{8}{3}\lambda + 5 \right) \frac{\varepsilon^2}{216} + O(\varepsilon^3). \quad (16.185)$$

Numerical results are given in Table 41.6, and compared with various high temperature (HT) series and experimental determinations.

It is worth mentioning that universal ratios of amplitudes of corrections to the leading critical behaviour have also been calculated. We expand any physical quantity, for $\tau = T - T_c$ small, as

$$f(\tau) = A_f |\tau|^{-\lambda_f} \left(1 + a_f |\tau|^\theta + \dots \right), \quad (16.186)$$

where the correction exponent θ (sometimes also called Δ_1) is given by

$$\theta = \omega\nu. \quad (16.187)$$

The ratio of correction amplitudes a_{f_1}/a_{f_2} corresponding to two different quantities f_1 and f_2 is also universal. A few such ratios have been calculated. One example is the ratio of corrections involving the correlation length and the susceptibility above T_c :

$$\frac{a_\chi^+}{a_\xi^+} = 2 \left\{ 1 - \frac{\varepsilon}{N+8} - \left[\frac{2\lambda}{3(N+8)} - \frac{N^2 - 15N - 124}{2(N+8)^3} \right] \varepsilon^2 \right\} + O(\varepsilon^3). \quad (16.188)$$

16.10 Conformal bootstrap

At an IR fixed point, a QFT is scale invariant, but is also expected to be conformal invariant (Section A13.3). In dimension 2, the conformal group is infinite dimensional, and this has made a systematic construction of many conformal invariant theories possible. However, in higher dimensions, the conformal group reduces to $SO(d + 1, 1)$, and only determines the form of the two- and three-point functions.

In recent years, a new method has been developed to determine critical exponents in dimensions $d > 2$. It combines conformal invariance and short distance expansion (Section 11.3) [91]. Describing the whole method precisely goes beyond this work, but a few simple elements can be given, in the example of the \mathbb{Z}_2 symmetric universality class, which contains the Ising model.

At the IR fixed point, all scaling operators can be classified according to their scaling dimension, spin and \mathbb{Z}_2 parity. Moreover, the scaling operators can be further divided into primaries, and descendants obtained by differentiation of the primaries.

SDE. The SDE of the product of two primary operators involves only primary operators $O_i(x)$ and descendants and, therefore, can be expressed only in terms of primaries in the form

$$O_i(x)O_j(y) = \sum_k f_{ijk} C_{ijk}(y, \partial_y) O_k(y),$$

where the functions C_{ijk} are determined by conformal invariance, spins and dimensions of the primary operators, and the f_{ijk} are additional constants, which play the role of structure constants of the operator algebra.

Moreover, in the case of the Ising-model universality class, the \mathbb{Z}_2 parity further restricts the set of operators in the SDE.

Associativity of the SDE. One considers the expectation value of the product of four primaries,

$$\langle O_1(x_1)O_2(x_2)O_3(x_3)O_4(x_4) \rangle,$$

and applies the SDE first on O_1O_2 and O_3O_4 , and then on O_1O_4 and O_3O_2 . In both cases, the four-point function is reduced to a combination of products of two-point functions. Assuming the associativity of the SDE, one can identify the two expansions. Presumably, the resulting equations determine a discrete set of parameters corresponding to \mathbb{Z}_2 symmetric critical theories.

Truncating the expansion, ordered according to increasing scaling dimensions, one then verifies that the resulting equations yield inequalities involving all the parameters of the conformal theory, in particular, scaling dimensions. The most recent results [435] constrain exponents with a remarkable precision, and yield values consistent with other theoretical estimates (Sections 41.5 and 41.6).

17 Critical phenomena: Corrections to scaling behaviour

In Chapters 15 and 16, while deriving the scaling behaviour of correlation functions, we have always kept only the leading term in the critical region. We examine now the different corrections to the leading behaviour [168–170, 80].

For instance, when we have solved the renormalization group (RG) equations, we have so far replaced the effective coupling constant $g(\lambda)$ at scale λ by g^* , neglecting the small difference $g(\lambda) - g^*$, which vanishes only if $g = g^*$. Moreover, to establish RG equations, we have neglected corrections subleading by powers of Λ , and effects of other couplings of higher canonical dimensions. Subleading terms related to the motion of $g(\lambda)$, which give the leading corrections for ε small, can easily be derived from the solutions of the RG equations studied previously, and are discussed first. The situations below and at four dimensions (the upper-critical dimension) have to be examined separately. The second type of corrections involves additional considerations, and is examined in the second part of the chapter.

The last section is devoted to one physical application, provided by systems with strong dipolar forces, which have 3 as upper-critical dimension.

17.1 Corrections to scaling: Generic dimensions

Dimensions $d < 4$. In dimensions $d < 4$, to characterize the corrections to scaling due to an initial value of the ϕ^4 coupling g different from the fixed point value g^* , it is convenient to solve the RG equations (16.24) by a slightly different method, which is based on introducing a set of coupling constant-dependent renormalizations:

$$\begin{cases} \ln \tilde{Z}(g) = - \int_{g^*}^g \frac{dg'}{\beta(g')} [\eta(g') - \eta], \\ \tilde{M}(g) = M Z^{-1/2}(g), \\ \tilde{\tau}(g) = \tau \exp \left[\int_{g^*}^g \frac{dg'}{\beta(g')} \left(\frac{1}{\nu(g')} - \frac{1}{\nu} \right) \right], \end{cases} \quad (17.1)$$

and a new coupling constant \tilde{g} , which characterizes the deviation of g from g^* :

$$\tilde{g} = (g - g^*) \exp \left[\int_{g^*}^g dg' \left(\frac{\omega}{\beta(g')} - \frac{1}{(g' - g^*)} \right) \right]. \quad (17.2)$$

($\omega = \beta'(g^*)$, see equation (16.11).) Then, in the vertex functions in the Fourier representation, we substitute

$$\begin{aligned} \tilde{\Gamma}^{(n)}(p_i; \tau, M, g, \Lambda) &= \tilde{Z}^{-n/2}(g) \tilde{\Gamma}^{(n)}(p_i; \tilde{\tau}(g), \tilde{M}(g), g^*, \Lambda) \\ &\times C^{(n)}(p_i; \tilde{\tau}(g), M(g), \tilde{g}, \Lambda). \end{aligned} \quad (17.3)$$

The function $C^{(n)}$ satisfies the boundary condition,

$$C^{(n)}(p_i; \tau, M; 0, \Lambda) = 1. \quad (17.4)$$

The finite renormalizations (17.1) eliminate the trivial deviations from the fixed point theory that correspond simply to a finite renormalizations of the different scaling variables. Equation (16.72) then implies

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \omega \tilde{g} \frac{\partial}{\partial \tilde{g}} - \frac{\eta}{2} M \frac{\partial}{\partial M} - \left(\frac{1}{\nu} - 2 \right) \tau \frac{\partial}{\partial \tau} \right] C^{(n)}(p_i, \tau, M, \tilde{g}, \Lambda) = 0. \quad (17.5)$$

Solving this equation by expanding $C^{(n)}$ in powers of \tilde{g} , one obtains

$$C^{(n)}(p_i, \tau, M, \tilde{g}, 1) = 1 + \sum_{s=1}^{\infty} \tilde{g}^s \tau^{s\omega\nu} D_s^{(n)}(p_i \tau^{-\nu}, \tau M^{-1/\beta}), \quad (17.6)$$

in which Λ has been set equal to 1.

Therefore, the exponent ω , which characterizes the approach to the fixed point, also characterizes the leading corrections to the critical behaviour.

For models with one coupling constant, $\beta(g)$ has the form

$$\beta(g) = -\varepsilon g + ag^2 + O(g^3, g^2\varepsilon). \quad (17.7)$$

In such a case, one always finds

$$\omega = \varepsilon + O(\varepsilon^2), \quad \omega\nu = \varepsilon/2 + O(\varepsilon^2). \quad (17.8)$$

Note that, to render all terms in the expansion (17.6) dimensionless, in the sense of scaling dimensions, we can assign to \tilde{g} the dimension $-\omega$.

Dimension $d = 4$. In exactly four dimensions, the situation is more subtle, because the $\phi^4(x)$ operator is marginal, and the approach to the fixed point is only logarithmic. This question is examined in next section.

Scaling for $d > 4$. So far, we have considered corrections to scaling for $d < 4$. In four dimensions or above, the fixed point corresponds to $g^* = 0$, that is, to the Gaussian fixed point and, therefore, the leading contributions to all correlation functions, except the two-point function, come from corrections to scaling, since these functions vanish at the fixed point. It is simple to verify that this special feature of the Gaussian fixed point explains the apparent contradiction between some RG predictions like relation between exponents involving explicitly the dimension d (called hyperscaling relations) and mean-field exponents: it is necessary to take into account the dimension of the ϕ^4 coupling constant g which, according to the preceding discussion, is $-\omega$:

$$\omega = d - 4, \quad \text{for } d > 4. \quad (17.9)$$

Let us consider, for example, the mean-field equation of state, valid for all dimensions $d > 4$:

$$H = \tau M + \frac{1}{6} g M^3.$$

The magnetization has dimension $(d-2)/2$ and the magnetic field H dimension $(d+2)/2$ in agreement with the general expressions for the exponents given by equations (16.65) and (16.66). These values are consistent with the product τM , since τ has dimension 2. Then, the dimension of gM^3 is $4-d+3(d-2)/2=(d+2)/2$, which is indeed the dimension of H .

17.2 Logarithmic corrections at the upper-critical dimension

The upper-critical dimension is the dimension at which deviations from mean-field theory appear (and the effective quantum field theory is just renormalizable). For the ϕ^4 field theory, this dimension is 4. In this dimension, there generally exists a marginal operator, here $\int \phi^4(x) d^d x$ and, therefore, as we have indicated in the general discussion of Section 16.1, logarithmic corrections to the mean-field behaviour are expected. We have already examined some consequences in the framework of particle physics in Sections 9.11 and 9.12, in particular, the issue of *triviality*. By contrast, the dimension 4 is not of physical relevance for statistical problems. However, its study is of special pedagogical value, because exact predictions can be derived from RG arguments. Because the infrared (IR) fixed point corresponds to $g^* = 0$, no assumption about the fixed point theory is required. Finally, we note some physical systems have $d = 3$ as upper-critical dimension, for example, tricritical systems (Section 14.8), or ferroelectrics with dipolar uniaxial long range forces (see Section 17.5).

We will study here only the equation of state and the specific heat, the generalization to correlation functions being straightforward. For $\varepsilon = 0$, the relation (16.58) becomes,

$$H(M, \tau, g, \Lambda) = Z^{-1/2}(\lambda)(\lambda\Lambda)^3 H(M(\lambda)/\lambda\Lambda, \tau(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1). \quad (17.10)$$

The various functions are given by,

$$\ln \lambda = \int_g^{g(\lambda)} \frac{dg'}{\beta(g')}, \quad (17.11)$$

$$\ln Z(\lambda) = \int_g^{g(\lambda)} dg' \frac{\eta(g')}{\beta(g')}, \quad (17.12)$$

$$\ln(\tau(\lambda)/\tau) = - \int_g^{g(\lambda)} dg' \frac{\eta_2(g')}{\beta(g')}, \quad (17.13)$$

and $M(\lambda)/M = Z^{-1/2}(\lambda)$.

In the $(\phi^2)^2$ field theory, ϕ being a N -component vector field, that is, in the $O(N)$ symmetric N -vector model, for g small, the expansions of the RG functions are (equations (10.89), (10.90), and (10.91), but with a different normalization of the coupling g),

$$\begin{aligned} \beta(g) &= \frac{(N+8)}{6} \frac{g^2}{8\pi^2} - \frac{(3N+14)}{12} \frac{g^3}{(8\pi^2)^2} O(g^4), \\ \eta(g) &= \frac{(N+2)}{72} \left(\frac{g}{8\pi^2} \right)^2 + O(g^3), \\ \eta_2(g) &= - \frac{(N+2)}{6} \frac{g}{8\pi^2} + O(g^2). \end{aligned} \quad (17.14)$$

We now introduce a physical momentum scale $\mu \ll \Lambda$, like the correlation length when it is finite, and choose then $\lambda = \mu/\Lambda$. The solution of equation (17.11) in this regime,

$$\ln \lambda = - \frac{48\pi^2}{(N+8)g(\lambda)} + \frac{3(3N+14)}{(N+8)^2} \ln(g(\lambda)) + O(1),$$

yields the effective coupling

$$g(\lambda) \equiv g_r \underset{\lambda \rightarrow 0}{=} \frac{48\pi^2}{N+8} \frac{1}{|\ln \lambda|} - \frac{3(3N+14)}{(N+8)^2} \frac{\ln |\ln \lambda|}{(\ln \lambda)^2} + O\left(\frac{1}{(\ln \lambda)^2}\right). \quad (17.15)$$

The factor $Z(\lambda)$ yields inessential finite renormalizations of H and M given by

$$\ln Z(\lambda) = \ln \zeta(g) + \frac{2\pi^2}{3} \frac{N+2}{N+8} g_r + O(g_r^2),$$

with

$$\ln \zeta(g) = \int_g^0 dg' \frac{\eta(g')}{\beta(g')}.$$

Therefore,

$$M(\lambda) = M \zeta^{-1/2}(g) [1 + O(g_r)].$$

Also,

$$\tau(\lambda) = T(g)(g(\lambda))^{(N+2)/(N+8)} \tau [1 + O(g_r)], \quad (17.16)$$

with

$$\ln T(g) = \int_g^0 dg' \left[\frac{N+2}{(N+8)g'} - \frac{\eta_2(g')}{\beta(g')} \right].$$

Since g_r is small, $H(\lambda)$ in the right-hand side of equation (17.10), as well all other physical quantities, can be calculated from perturbation theory. Note here the power of the RG method: we started from a theory with a coupling constant of order 1 and perturbative coefficients increasing like powers of $\ln(\tau/\Lambda^2)$ or equivalent. Direct perturbation theory is obviously useless. In contrast, in the effective theory at scale λ , the coupling constant $g(\lambda)$ is small, and the perturbative coefficients are of order 1.

The field H has the perturbative expansion,

$$H = \tau M + \frac{g}{6} M^3 + \dots .$$

We renormalize H , M , and τ by the factors $\zeta^{1/2}(g), T(g)$. Then, from the relations (17.10) and (17.15), one derives

$$H(M, \tau, g, \Lambda = 1) = \tau M g_r^{(N+2)/(N+8)} + \frac{1}{6} g_r M^3 + \dots . \quad (17.17)$$

At T_c (*i.e.*, $\tau = 0$), the physical length scale can be chosen such that $M(\lambda)/\lambda = 1$ (with $\Lambda = 1$), and thus $\lambda \propto M/\Lambda$. It follows that

$$H \propto M^3 / |\ln M| . \quad (17.18)$$

For $|\tau| \neq 0$, it is more convenient to use the length scale defined by

$$\tau(\lambda) = \lambda^2 \Rightarrow g(\lambda) \propto 1 / \ln |\tau|. \quad (17.19)$$

Then, the spontaneous magnetization is given by ($\tau < 0$)

$$M \propto (|\tau|)^{1/2} |\ln |\tau||^{3/(N+8)}, \quad (17.20)$$

and the susceptibility in zero field by

$$\chi^{-1} \propto |\tau| |\ln |\tau||^{-(N+2)/(N+8)}. \quad (17.21)$$

Finally, the specific heat satisfies the RG equation,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \eta_2(g) \left(2 + \tau \frac{\partial}{\partial \tau} \right) \right] \tilde{\Gamma}^{(2,0)} = B(g), \quad (17.22)$$

and $B(g)$ has the expansion

$$B(g) = \frac{N}{16\pi^2} + O(g). \quad (17.23)$$

It is simple to verify that a function $C_2(g)$, solution of equation (17.22), and thus of

$$\beta(g)C'_2(g) - 2\eta_2(g)C_2(g) = B(g),$$

is necessarily singular at the origin. For example, one can take a solution of the form

$$C_2(g) = \frac{3N}{(N-4)} \frac{1}{g} + O(1), \quad \text{for } N \neq 4. \quad (17.24)$$

For $N = 4$, an additional logarithmic singularity is present. The combination $\tilde{\Gamma}^{(2,0)} - C_2(g)$, solution of the homogeneous RG equation, is, for g small, dominated by the pole of $C_2(g)$:

$$\tilde{\Gamma}^{(2,0)}(0; \tau, g, \Lambda) - C_2(g) \sim -\frac{3N}{(N-4)} \frac{1}{g(\lambda)} \zeta^{-2}(\lambda), \quad (17.25)$$

with

$$\zeta(\lambda) = \exp \left[\int_g^{g(\lambda)} \frac{dg'}{\beta(g')} \eta_2(g') \right] \propto [g(\lambda)]^{-(N+2)/(N+8)}. \quad (17.26)$$

Collecting all factors, one obtains the behaviour of the specific heat:

$$\tilde{\Gamma}^{(2,0)}(0; \tau, g, \Lambda) - C_2(g) \propto |\ln \tau|^{(4-N)/(N+8)} \left[1 + O\left(\frac{\ln |\ln \tau|}{\ln \tau}\right) \right]. \quad (17.27)$$

Particle physics. For the viewpoint of particle physics, the renormalization by a logarithmic factor in equation (17.16) is irrelevant compared to the Gaussian renormalization, and does not alleviate the fine-tuning problem. However, the equation analogous to equation (17.13) in the *renormalized theory*, describing a finite renormalization associated with a change of physical scale, is useful.

Remark. It is apparent from these expressions that a parametrization in terms of the variables τ or M leads to rather complicated expressions. A more efficient way of writing all these results is to introduce a parametric representation in terms of the effective coupling constant $g(\lambda)$ and to calculate λ in terms of $g(\lambda)$ from equation (17.11). We parametrize $\beta(g)$ as

$$\beta(g) = \beta_2 g^2 + \beta_3 g^3 + O(g^4), \quad (17.28)$$

and set

$$s = g(\lambda). \quad (17.29)$$

Then, λ is given by

$$\lambda = s^{-\beta_3/\beta_2^2} e^{-1/\beta_2 s} \tilde{\lambda}(s). \quad (17.30)$$

In equation (17.30), the function $\tilde{\lambda}(s)$ is a regular function of s for s small. The renormalization factor $Z(\lambda)$ is a regular function of s . Finally, equation (17.13) yields

$$\frac{\tau(\lambda)}{\lambda^2 \Lambda^2} = \frac{\tau}{\Lambda^2} s^{[2\nu_1 + \beta_3/\beta_2]/\beta_2} e^{2/\beta_2 s} [\tilde{\tau}(s)]^{-1}, \quad (17.31)$$

in which $\tilde{\tau}(s)$ is a regular function of s , and $\nu(g)$ has been parametrized as

$$\nu(g) = (2 + \eta_2(g))^{-1} = \tfrac{1}{2} + \nu_1 g + O(g^2). \quad (17.32)$$

Then, we determine λ from the condition (17.19), and equation (17.31) parametrizes τ as a function of s . At leading order, all critical behaviours are described by a singular factor of the form occurring in equations (17.30) or (17.31), multiplied by regular series in $s = g(\lambda)$.

17.3 Irrelevant operators and the question of universality

In generic dimensions, we now examine the contributions coming from irrelevant operators (see also Section 11.1) [168–170]. We again stress that these operators have been found to be irrelevant at the Gaussian fixed point, near four dimensions. We still rely on the assumption that dimensions vary continuously when the IR fixed point moves away from the Gaussian fixed point. Finally, the analysis is local, we consider only the neighbourhood of the fixed point.

We first recall power counting arguments for a general theory with an action $\mathcal{S}(\phi)$ (for details, see Chapters 8 and 11):

$$\mathcal{S}(\phi) = \int d^d x \left[\tfrac{1}{2} (\nabla \phi(x))^2_\Lambda + \tfrac{1}{2} r \phi^2(x) + \sum_\alpha u_\alpha \mathcal{O}_\alpha(\phi, x) \right], \quad (17.33)$$

in which $\mathcal{O}_\alpha(\phi, x)$ is a local monomial of degree n_α in ϕ with k_α derivatives. The dimension $[u_\alpha]$ of the coupling constant u_α is then (consequence of equation (8.15))

$$[u_\alpha] = d - k_\alpha - \tfrac{1}{2} n_\alpha (d - 2). \quad (17.34)$$

We treat all interactions in action (17.33) in perturbation theory. Order by order in the loop expansion, we evaluate the divergent part of the corresponding Feynman diagrams, and add counter-terms to the action to render the theory finite. Since the action (17.33) contains all possible monomials in the field, any counter-term is a linear combination of the operators $\mathcal{O}_\alpha(\phi)$.

In order for a product of constants u_β to appear in a counter-term proportional to an operator \mathcal{O}_α , it is necessary and generically sufficient that the condition

$$\Delta = d - [\mathcal{O}_\alpha] - \sum_l [u_{\beta_l}] \geq 0, \quad (17.35)$$

is satisfied. Then, the coefficient $\delta u_\alpha(\Lambda)$ of the counter-term proportional to \mathcal{O}_α diverges like a positive power of the cut-off Λ (or a power of logarithm if $\Delta = 0$),

$$\delta u_\alpha(\lambda) \sim \Lambda^\Delta. \quad (17.36)$$

We now return to the question of irrelevant operators. We restrict ourselves to dimensions $d = 4 - \varepsilon$, $\varepsilon > 0$ and small, since as we have seen, this is the only situation in which a reliable analysis is possible. The first operators we have, for example, in mind are ϕ^6 , $\phi^2(\nabla\phi)^2\dots$, which are operators of dimension 6 in four dimensions.

To introduce the cut-off Λ in the effective field theory, we rescale the lengths and the field ϕ (equations (15.40) and (15.41)). Therefore, each coupling constant u_α is the product of a dimensionless quantity g_α by a power of the cut-off:

$$u_\alpha = g_\alpha \Lambda^{-\delta_\alpha}, \quad (17.37)$$

where δ_α , as given by equation (17.34), has the form

$$\delta_\alpha = -d + k_\alpha + \frac{1}{2}n_\alpha(d-2). \quad (17.38)$$

If δ_α is positive, the corresponding operator \mathcal{O}_α leads to a non-renormalizable theory, and we have already stated that it is irrelevant. In the tree approximation, the statement follows from equation (17.37): the operator gives contributions vanishing with a power of the cut-off. However, in higher orders, the statement is less trivial, since divergences at large cut-off coming from the momentum integration may cancel the powers coming from the coupling constants. To understand what happens, it is necessary to analyse the counter-terms generated by these operators using equations (17.35) and (17.36).

The total power Δ' of the cut-off, which multiplies the operator \mathcal{O}_α in a counter-term, is the sum of the power Δ generated by the divergence of perturbation theory (equation (17.35)), and the powers already present in the coefficients u_β (equation (17.37)):

$$\Delta' = \Delta - \sum_l \delta_{\beta_l} = \Delta + \sum_l [u_{\beta_l}], \quad (17.39)$$

and, therefore, using the definition (17.35),

$$\Delta' = d - [\mathcal{O}_\alpha] = [u_\alpha]. \quad (17.40)$$

The conclusion is simple: due to the divergences of perturbation theory, irrelevant operators indeed give non-vanishing contributions, but these contributions can be cancelled by changing the amplitudes of the relevant or marginal terms in the Hamiltonian, because $\Delta' \geq 0$ is equivalent to $[u_\alpha] \geq 0$.

Example. The leading new corrections come from operators \mathcal{O}_i^6 which have dimension 6 in four dimensions. The corresponding interactions have the form

$$\Lambda^{2\varepsilon-2} \int d^d x \phi^6(x), \quad \Lambda^{\varepsilon-2} \int d^d x (\phi(x) \nabla \phi(x))^2, \quad \Lambda^{-2} \int d^d x (\nabla^2 \phi(x))^2.$$

In terms of the renormalized operators, they have the expansion

$$\int d^d x \mathcal{O}_i^6(x) = \int d^d x \left\{ \sum_j Z_{ij} [\mathcal{O}_j^6(x)]_r + \sum_j A_{ij} [\mathcal{O}_j^4(x)]_r + B [\phi^2(x)]_r \right\}.$$

We have denoted by $\mathcal{O}_j^4(x)$ the two operators of dimension 4 in four dimensions: ϕ^4 and $(\nabla\phi)^2$. In the framework of the ε -expansion, the coefficients Z_{ij} diverge like powers of $\ln \Lambda$, A_{ij} like Λ^2 and B like Λ^4 , up to powers of $\ln \Lambda$.

Taking into account the powers of Λ in front of the interaction terms, we see that only the contributions proportional to operators of dimensions 4 and 2 are divergent. If we cancel these contributions by subtracting to the operators of dimension 6 a suitable combination of bare operators of dimensions 4 and 2 (bare and renormalized operators are linearly related), we obtain the true new corrections which decrease like $\Lambda^{-2+O(\varepsilon)}$.

Discussion. The analysis also clarifies the interpretation of the constants r and g , which parametrize the ϕ^4 Hamiltonian. These are not the parameters that are generated directly by the microscopic theory but, instead, effective parameters taking into account the effect of neglected irrelevant operators. However, the analysis of previous chapters is, at leading order, not modified. Indeed, the change in ϕ^2 corresponds only to a modification of the critical temperature which is a non-universal quantity. Moreover, below four dimensions, we have shown that many physical quantities (universal quantities) do not depend on g either, since g can be replaced by its fixed point value g^* . Finally, a change in the cut-off procedure corresponds generally to a change in the coefficients of the irrelevant part of the propagator ($\phi\Delta^2\phi\dots$). The effect of such a change is obtained from the previous analysis also. We can now clarify the concept of universality: below four dimensions all dimensionless quantities in which g can be replaced by g^* , the IR fixed point value, and which do not depend on the normalizations of the field ϕ , the deviation from the critical temperature t , and of the magnetic field are universal. Obvious examples are ratios of amplitude of singularities below and above T_c , ratios of amplitudes of leading corrections to scaling, the rescaled equation of state (relation between H and M), the renormalized correlation functions as defined in Chapters 15 and 16, and so on.

Emergent symmetries. A simple consequence of the analysis is the following: if one adds an irrelevant operator to a renormalizable Hamiltonian, which breaks its symmetry, then the symmetry of the critical theory is broken if and only if the irrelevant operator generates by renormalization relevant or marginal operators breaking the symmetry. An application is the following: on the lattice, operators of the form $\sum_\mu \int \phi(x)\partial_\mu^4\phi(x)d^d x$ which break the $O(d)$ symmetry of the effective $\phi^4(x)$ action are present. However, these operators have a hypercubic symmetry and, since the only relevant operators they can generate, like $\int (\nabla\phi)^2 d^d x$, due to the hypercubic symmetry, are $O(d)$ symmetric, the critical theory has an *emergent $O(d)$ symmetry*.

By contrast, the addition of a naively irrelevant term like $\int \phi^5(x)d^d x$ to a Hamiltonian, which is symmetric in the reflection $\phi \mapsto -\phi$, generates relevant terms linear in ϕ , which are equivalent to the addition of a magnetic field, and breaks the reflection symmetry.

17.4 Corrections coming from irrelevant operators. Improved action

For simplicity, we consider the effect, in the critical theory, of an operator \mathcal{O}_α at first order only in the corresponding coupling constant u_α . The following discussion applies in the framework of the ε -expansion, and relies on the results of Chapter 11 concerning the renormalization of composite operators.

17.4.1 Corrections to scaling

In Section 17.3, we have shown that an operator \mathcal{O}_α gives contributions equivalent to all operators of lower canonical dimensions. For example, $\phi^8(x)$ first generates effects equivalent to $\phi^2(x)$, $\phi^4(x)$ and $(\nabla\phi(x))^2$ and all operators of dimension 6, and then genuine new corrections. To isolate these corrections, it is necessary to subtract from the operator a linear combination of all operators which have a lower dimensions at $d = 4$, that is, to perform an additive renormalization.

However, we can omit all operators that are total derivatives, since only the space integrals appear in the action. We define a subtracted operator $\bar{\mathcal{O}}_\alpha(\phi)$ by

$$\bar{\mathcal{O}}_\alpha(\phi) = \mathcal{O}_\alpha(\phi) - \sum_{\beta \text{ such that } [\mathcal{O}_\beta] < [\mathcal{O}_\alpha]} C_{\alpha\beta}(\Lambda, g) \mathcal{O}_\beta(\phi). \quad (17.41)$$

Let us again illustrate this point with the example of the operators of dimension 6, like $\phi^6(x)$. One then subtracts a linear combination of operators of dimensions 2 and 4:

$$\bar{\phi}^6(x) = \phi^6(x) - C_1 \phi^2(x) - C_2 [\nabla \phi(x)]^2 - C_3 \phi^4(x).$$

The coefficients C_1 , C_2 , and C_3 can be determined by a set of renormalization conditions at zero momentum:

$$\begin{aligned} \tilde{\Gamma}_{\bar{\phi}^6}^{(2)}(p) &= O(p^4 \times \text{powers of } \ln p), \text{ for } p \rightarrow 0, \\ \tilde{\Gamma}_{\bar{\phi}^6}^{(4)}(p_i = 0) &= 0. \end{aligned} \quad (17.42)$$

The first condition implies, in particular, that the critical temperature is not changed. These conditions are not affected by IR divergences, because the correlation functions with an operator insertion have positive dimensions.

After such additive renormalizations, the bare operators are related to the completely renormalized operators \mathcal{O}_α^r by

$$\mu^{-\delta_\alpha} \int d^d x \mathcal{O}_\alpha^r(x) = \sum_\beta Z_{\alpha\beta}(g, \Lambda/\mu) \Lambda^{-\delta_\beta} \int d^d x \bar{\mathcal{O}}_\beta(x), \quad (17.43)$$

in which μ is the renormalization scale. Additional renormalization conditions at scale μ for the insertion of renormalized operators determine the matrix $Z_{\alpha\beta}$.

The relation between correlation functions $\tilde{\Gamma}_{\bar{\mathcal{O}}_\alpha}^{(n)}$ with one $\Lambda^{-\delta_\alpha} \int d^d x \bar{\mathcal{O}}_\alpha$ insertion, and the renormalized functions with $\mu^{-\delta_\alpha} \int d^d x \mathcal{O}_\alpha(x)$ insertion, then is

$$\sum_\beta Z_{\alpha\beta} Z^{n/2} \tilde{\Gamma}_{\bar{\mathcal{O}}_\beta}^{(n)}(p_i; g, \Lambda) = \left[\tilde{\Gamma}_{\mathcal{O}_\alpha}^{(n)}(p_i, g_r, \mu) \right]_r. \quad (17.44)$$

This leads to the RG equations,

$$\sum_\beta \left\{ \left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right] \delta_{\alpha\beta} - \eta_{\alpha\beta}(g) \right\} \tilde{\Gamma}_{\bar{\mathcal{O}}_\beta}^{(n)} = 0, \quad (17.45)$$

with

$$\eta_{\alpha\beta}(g) = - \sum_\gamma \left(\Lambda \frac{\partial}{\partial \Lambda} Z_{\alpha\gamma} \right) (Z^{-1})_{\gamma\beta}. \quad (17.46)$$

The effects of the insertions of the operators $\bar{\mathcal{O}}_\alpha$ are then governed by the eigenvalues η_α of the matrix $\eta_{\alpha\beta}(g^*)$ [171]. From the relation (17.43), we infer that the renormalization matrix $Z_{\alpha\beta}$ has the form

$$Z_{\alpha\beta} = \delta_{\alpha\beta} (\Lambda/\mu)^{\delta_\alpha} (1 + O(g)). \quad (17.47)$$

Therefore, the eigenvalues η_α are at leading order given by

$$\eta_\alpha = \delta_\alpha + O(\varepsilon).$$

The genuine new contributions of the irrelevant operators of canonical dimension $d + \delta_\alpha$ in four dimensions, are suppressed by powers $\Lambda^{-\delta_\alpha+O(\varepsilon)}$ of the cut-off Λ . For operators of dimension 6, $\delta_6 = 2$. In an infinitesimal neighbourhood of dimension 4, these operators remain irrelevant. Our general analysis, which is based upon the idea that the critical behaviour of ferromagnetic systems can be described by an effective ϕ^4 field theory remains valid beyond the ε -expansion, as long as this property remains true.

Important irrelevant effects. In some cases, the irrelevant effects may be especially important. An example is provided by systems where the initial theory has only a discrete symmetry, while the symmetry of the critical theory is continuous. In the low temperature phase, the critical theory has Goldstone-mode singularities. These singularities are suppressed by corrections due to irrelevant operators.

17.4.2 Fixed point in Hamiltonian space and improved action

We have noted that by adding to the $\phi^4(x)$ field theory irrelevant interactions, we could modify correlation functions by terms in the ε -expansion of order $1/\Lambda^2$ (up to logarithms). In the right-hand side of the RG equations (15.46), we have neglected terms of the same order. Symanzik [90] has given a perturbative proof that, by adding to the Hamiltonian the proper linear combination of irrelevant operators, it is possible to cancel exactly these corrections. The coefficients of the linear combination are functions of the ϕ^4 coupling constant g . For example, the complete set of operators of dimension 6 can be used to cancel exactly the corrections of order $1/\Lambda^2$ in the right-hand side of the RG equations (15.46), the operators of dimension 8 to cancel the order $1/\Lambda^4$ (up to logarithms), and so on. An infinite iteration of this procedure leads to a theory that depends on only one ϕ^4 coupling constant, and which satisfies the RG equations exactly. It is actually a ‘renormalized’ theory constructed without using the renormalization procedure, but by considering an infinite sequence of Hamiltonians.

From the general RG point of view, the Hamiltonians which lead to correlation functions satisfying RG equations exactly belong to a one parameter line in Hamiltonian space which goes from the Gaussian fixed point to the non-trivial IR fixed point.

Conversely, by directly constructing an RG for a cut-off field theory, it is possible to prove the existence of the renormalized field theory [59].

As an application, Symanzik [90] has advocated the use of such improved actions (adding, for example, all terms of dimension 6) for numerical lattice simulations. However, it should be mentioned, that the applications of this ingenious idea (in particular, in numerical simulations for particle physics) require very precise numerical data and large lattices. Indeed, because the improved action involves more extended interactions on the lattice (like second nearest neighbours), the effective size of the lattice is reduced, increasing finite size effects. But statistics then becomes a serious problem, and more complicated interactions slow down numerical calculations.

17.5 Application: Uniaxial systems with strong dipolar forces

In Section 17.2, we have stressed that, for critical systems at the upper-critical dimension, the RG leads to exact predictions. Unfortunately, in the case of the N -vector model, the upper-critical dimension is 4 and, therefore, the predictions are not useful for macroscopic phase transitions. The main application is numerical physics, for example, the Higgs sector ($N = 4$) of the Standard Model of particle physics has been investigated numerically.

However, the critical dimensions change with long-range forces [172, 173]. Therefore, here we describe another system, for which precise measurements have been made, and which has dimension 3 as the upper-critical dimension: a uniaxial ferromagnet or ferroelectric system with strong dipolar forces [174, 169, 175].

Dipolar forces. We consider a spin system in d dimensions in which the d -component spins S^μ interact both through short range and dipolar forces:

$$-\beta\mathcal{H}(\mathbf{S}) = \sum_{\mathbf{x}, \mathbf{x}'} \left[\sum_{\mu, \nu} V_{\mu\nu}(\mathbf{x} - \mathbf{x}') S_\mathbf{x}^\mu S_{\mathbf{x}'}^\nu + \gamma (\mathbf{S}_\mathbf{x} \cdot \nabla_\mathbf{x})(\mathbf{S}_{\mathbf{x}'} \cdot \nabla_{\mathbf{x}'}) \frac{1}{|\mathbf{x} - \mathbf{x}'|^{d-2}} \right], \quad (17.48)$$

$V_{\mu\nu}(\mathbf{x})$ being the short range potential, and γ a parameter. We assume that the long range dipolar forces are strong enough to play a role in the part of the critical domain accessible to experiments. In addition, we assume that the lattice is strongly anisotropic in such a way that only one component of the spin \mathbf{S} is critical, and the effective Hamiltonian can be simplified into

$$-\beta\mathcal{H}(S) = \sum_{\mathbf{x}, \mathbf{x}'} S_\mathbf{x} S_{\mathbf{x}'} \left(V(\mathbf{x} - \mathbf{x}') + \gamma \partial_z \partial_{z'} \frac{1}{|\mathbf{x} - \mathbf{x}'|^{d-2}} \right), \quad (17.49)$$

in which S now denotes the component of the spin vector \mathbf{S} along the $z \equiv x_d$ axis.

After Fourier transformation, the Hamiltonian can be written as

$$-\beta\mathcal{H}(S) = \int d^d q \tilde{S}(\mathbf{q}) \tilde{S}(-\mathbf{q}) \left(\tilde{V}(\mathbf{q}) + \tilde{\gamma} \frac{q_z^2}{q^2} \right), \quad (17.50)$$

where $\tilde{\gamma}$ is proportional to γ and $\tilde{V}(\mathbf{q})$ is a regular function of \mathbf{q} which, due to hypercubic symmetry, has the expansion

$$\tilde{V}(\mathbf{q}) = \tilde{V}(0) + \frac{1}{6} \nabla^2 \tilde{V}(0) \mathbf{q}^2 + O(q^4). \quad (17.51)$$

In the critical domain, in which $|\mathbf{q}|$ is small, the two terms coming from the short range potential and the dipolar forces are of the same order of magnitude:

$$q^2 \sim q_z^2 / q^2. \quad (17.52)$$

This implies that q_z , the z component of the vector \mathbf{q} , is much smaller than the other components \mathbf{q}_\perp :

$$|q_z| \sim (\mathbf{q}_\perp)^2. \quad (17.53)$$

Therefore, we can simplify further the interaction potential. Finally, in the case of an even spin distribution, we can reproduce the configuration energy by an effective Hamiltonian $\mathcal{H}(\phi)$ of the form (in terms of ‘bare’ parameters τ_0, A_0, u_0), in the Fourier representation,

$$\begin{aligned}\mathcal{H}(\phi) = & \frac{1}{2} \int d^d q \phi(-\mathbf{q}) (\mathbf{q}_\perp^2 + A_0^2 q_z^2 / \mathbf{q}_\perp^2 + r_c + \tau_0) \phi(\mathbf{q}) \\ & + \frac{u_0}{4!} \int d^d q_1 \cdots d^d q_4 \delta^{(d)}(\sum \mathbf{q}_i) \phi(\mathbf{q}_1) \cdots \phi(\mathbf{q}_4),\end{aligned}\quad (17.54)$$

where r_c is the critical value and τ_0 parametrizes the deviation from the critical temperature.

The upper-critical dimension. Usual power counting is modified, because space is no longer isotropic. In units of the transverse components of \mathbf{q} , the dimension of q_z is 2 (equation (17.53)):

$$[q_z] = 2 \Rightarrow [z] = -2.$$

The volume element in configuration space $d\mathbf{x}_\perp dz$ has thus canonical dimension $-(d+1)$. This implies that power counting analysis is the same as in the conventional ϕ^4 theory in $(d+1)$ dimensions. In particular, the upper-critical dimension is given by

$$d+1=4 \Rightarrow d=3.$$

The RG methods thus lead to exact analytic predictions in the physical dimension $d=3$.

RG equations. From expression (17.54), we infer the propagator (A and τ are the renormalized parameters),

$$\tilde{\Delta}(\mathbf{q}) = \frac{\mathbf{q}_\perp^2}{(\mathbf{q}_\perp^2)^2 + A^2 q_z^2 + \tau \mathbf{q}_\perp^2}. \quad (17.55)$$

Diagrams calculated with this propagator are regular for \mathbf{q} small, therefore,

$$\int d^d q \phi(-\mathbf{q}) \frac{q_z^2}{\mathbf{q}_\perp^2} \phi(\mathbf{q}),$$

never appears as a counter-term.

The renormalized Hamiltonian thus has the form ($Z, Z_2, Z_g, \delta m^2$ are renormalization constants)

$$\begin{aligned}\mathcal{H}_r = & \frac{1}{2} \int d^d q \phi(-\mathbf{q}) (Z \mathbf{q}_\perp^2 + A^2 q_z^2 / \mathbf{q}_\perp^2 + \delta m^2 + Z_2 \tau) \phi(\mathbf{q}) \\ & + \frac{\mu^\varepsilon}{4!} A g Z_g(g) \int d^d q_1 \cdots d^d q_4 \delta^{(d)}(\sum \mathbf{q}_i) \phi(\mathbf{q}_1) \cdots \phi(\mathbf{q}_4),\end{aligned}\quad (17.56)$$

in which μ is the renormalization scale and ε now is defined as

$$d = 3 - \varepsilon. \quad (17.57)$$

We introduce also a bare dimensionless coupling constant

$$u_0 = \Lambda^\varepsilon A_0 g_0. \quad (17.58)$$

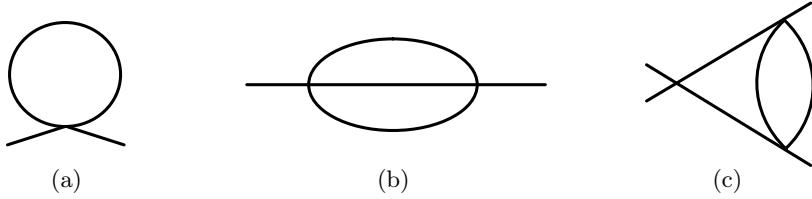


Fig. 17.1 Diagrams needed at two-loop order

The relation between bare and renormalized correlation functions reads

$$\tilde{\Gamma}_r^{(n)}(p_i; \tau, g, A, \mu) = Z^{n/2} \tilde{\Gamma}^{(n)}(p_i; \tau_0, g_0, A_0, \Lambda). \quad (17.59)$$

In addition, comparing expressions (17.54) and (17.56), one finds

$$A = Z^{1/2} A_0. \quad (17.60)$$

RG equations follow:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0} + \frac{1}{2} \eta(g_0) \left(A_0 \frac{\partial}{\partial A_0} - n \right) - \eta_2(g_0) \tau_0 \frac{\partial}{\partial \tau_0} \right] \tilde{\Gamma}^{(n)}(p_i; g_0, A_0, \tau_0, \Lambda) = 0. \quad (17.61)$$

Two-loop calculation of RG functions. For reasons explained in Section 15.4, for practical calculations, we use the renormalized theory and minimal subtraction (see Chapter 10). In what follows, to simplify the notation we omit the index r on renormalized functions. The renormalized RG equations are formally identical to equations (17.61):

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} + \frac{1}{2} \eta(g) \left(A \frac{\partial}{\partial A} - n \right) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}^{(n)}(p_i; g, A, \tau, \Lambda) = 0, \quad (17.62)$$

but the RG functions are different at two-loop order.

The calculations are here somewhat similar to the dynamic calculations as described in Section A36, because if we identify the z direction with time, the propagators have the same denominators. The combinatorial factors of Feynman diagrams are those of the ϕ^4 field theory. Only the expressions of the diagrams differ. We need only their values at vanishing external z components of momenta, and this simplifies the integration over momentum variables in the z direction (called ω hereafter). In Fig. 17.1, we display the three diagrams needed at two-loop order. In the critical (or massless) theory, in terms of the Fourier transform of the propagator,

$$\tilde{\Delta}(\omega, k) = \frac{1}{k^2 + \omega^2/k^2},$$

the expressions are

$$\begin{aligned} (a) &\equiv \frac{1}{(2\pi)^d} \int d\omega d^{d-1}k \tilde{\Delta}(\omega, k) \tilde{\Delta}(\omega, p+k) \\ &= \frac{1}{(2\pi)^{d-1}} \frac{1}{2} \int \frac{d^{d-1}k}{k^2 + (p+k)^2} = \frac{1}{(16\pi)^{(d-1)/2}} \Gamma(\varepsilon/2) p^{-\varepsilon}. \\ (b) &\equiv \frac{1}{(2\pi)^{2d}} \int d\omega_1 d\omega_2 d^{d-1}k_1 d^{d-1}k_2 \tilde{\Delta}(\omega_1, k_1) \tilde{\Delta}(\omega_2, k_2) \tilde{\Delta}(\omega_1 + \omega_2, k_1 + k_2 + p) \\ &= \frac{1}{(2\pi)^{d-2}} \frac{1}{4} \int \frac{d^{d-1}k_1 d^{d-1}k_2}{k_1^2 + k_2^2 + (k_1 + k_2 + p)^2} = \frac{1}{(16\pi)^{d-1}} \frac{3}{8} \left(\frac{16}{27} \right)^{(d-1)/2} \Gamma(2-d) p^{2d-4}. \end{aligned}$$

After integration over the corresponding ω variables, the diagram (c) takes the form

$$(c) = \frac{1}{(2\pi)^{2(d-1)}} \frac{1}{8} [C_1(p_1, p_2) + C_1(p_2, p_1) + C_2(p_1, p_2)], \quad \text{with}$$

$$C_1(p, q) = \int \frac{d^{d-1}k_1 d^{d-1}k_2}{[k_1^2 + (k_1 + p + q)^2][k_1^2 + k_2^2 + (k_1 + k_2 + p)^2]},$$

$$C_2(p, q) = \int \frac{d^{d-1}k_1 d^{d-1}k_2}{[(k_1 + p + q)^2 + k_2^2 + (k_1 + k_2 + p)^2]} \frac{1}{[k_1^2 + k_2^2 + (k_1 + k_2 + p)^2]}.$$

The final result is

$$C_1(\varepsilon) \simeq \frac{1}{(16\pi)^{d-1}} \left(\frac{2}{\sqrt{3}} \right)^\varepsilon \frac{\Gamma(1+\varepsilon)}{\varepsilon^2} (|p_1 + p_2|)^{-2\varepsilon} + O(1),$$

$$C_2(\varepsilon) = \frac{2}{3} \frac{1}{(16\pi)^2 \varepsilon} + O(1).$$

The renormalization constants Z_g , Z and Z_2 are then determined up to order g^2 :

$$Z_g = 1 + \frac{3N_d}{\varepsilon} g + \left(\frac{9}{\varepsilon^2} - \frac{3}{\varepsilon} \ln \frac{4}{3} - \frac{2}{\varepsilon} \right) N_d^2 g^2 + O(g^3), \quad (17.63)$$

$$Z = 1 - \frac{2}{27} N_d^2 \frac{g^2}{\varepsilon} + O(g^3), \quad (17.64)$$

$$Z_2^{-1} = 1 - N_d \frac{g}{\varepsilon} + \left(-\frac{1}{\varepsilon^2} + \frac{1}{3\varepsilon} + \frac{1}{2\varepsilon} \ln \frac{4}{3} \right) N_d^2 g^2 + O(g^3), \quad (17.65)$$

with

$$N_d = (16\pi)^{\varepsilon/2-1} \Gamma(1 + \varepsilon/2). \quad (17.66)$$

The RG functions follow:

$$\beta(g) = -\varepsilon \left[\frac{d}{dg} \ln(gZ_g Z^{-3/2}) \right]^{-1}, \quad (17.67)$$

$$\eta(g) = \beta(g) \frac{d}{dg} \ln Z(g), \quad (17.68)$$

$$\eta_2(g) = \frac{1}{\nu(g)} - 2 = \frac{d}{dg} \ln(Z_2 Z^{-1}). \quad (17.69)$$

Therefore, setting $\tilde{g} = N_d g$, one finds

$$N_d \beta(\tilde{g}) = -\varepsilon \tilde{g} + 3\tilde{g}^2 - \left(-6 \ln \frac{4}{3} + \frac{34}{9} \right) \tilde{g}^3 + O(\tilde{g}^4), \quad (17.70)$$

$$\eta(\tilde{g}) = \frac{4}{27} \tilde{g}^2 + O(\tilde{g}^3), \quad (17.71)$$

$$\eta_2(\tilde{g}) = -\tilde{g} + \left(\frac{14}{27} + \ln \frac{4}{3} \right) \tilde{g}^2 + O(\tilde{g}^3). \quad (17.72)$$

Scaling behaviour below three dimensions. Dimensional analysis in the critical theory yields

$$\begin{aligned} \tilde{\Gamma}^{(n)}(\lambda \mathbf{p}_\perp, \rho p_z; \tau, g, A, \mu) &= \lambda^{n+(n-2)(1-d)/2} \rho^{(2-n)/2} \\ &\times \tilde{\Gamma}^{(n)}(\mathbf{p}_\perp, p_z; \tau/\lambda^2, g, A\rho/\lambda^2, \mu/\lambda). \end{aligned} \quad (17.73)$$

In $d = 3 - \varepsilon$ dimensions, the model has an IR fixed point $g^*(\varepsilon)$. At the fixed point, one finds

$$\tilde{\Gamma}^{(n)}(\mathbf{p}_\perp, p_z, \tau, A = \mu = 1) = \tau^{\gamma-(n-2)d_\phi} \tilde{\Gamma}^{(n)}(\mathbf{p}_\perp \tau^\nu, p_z / \tau^{\nu(2-\eta/2)}), \quad (17.74)$$

with

$$\gamma = \nu(2 - \eta), \quad d_\phi = \frac{1}{2}(d - 1 + \eta).$$

At two-loop order, the fixed point coupling constant and the critical exponents are

$$\tilde{g}^*(\varepsilon) = \frac{\varepsilon}{3} + \left(\frac{2}{9} \ln \frac{4}{3} + \frac{34}{243} \right) \varepsilon^2 + O(\varepsilon^3), \quad (17.75)$$

$$\eta = \frac{4}{243} \varepsilon^2 + O(\varepsilon^3), \quad (17.76)$$

$$\nu^{-1} = 2 - \frac{\varepsilon}{3} - \left(\frac{1}{9} \ln \frac{4}{3} + \frac{20}{243} \right) \varepsilon^2 + O(\varepsilon^3), \quad (17.77)$$

$$\omega = \varepsilon - \left(\frac{2}{3} \ln \frac{4}{3} + \frac{34}{81} \right) \varepsilon^2 + O(\varepsilon^3). \quad (17.78)$$

Logarithmic corrections to mean-field behaviour in three dimensions. In three dimensions, the RG equations can be solved by the method indicated in Section 17.2.

For the effective coupling constant at scale λ , one finds

$$\frac{1}{g(\lambda)} \underset{\lambda \rightarrow 0}{\sim} \frac{3}{16\pi} \ln \frac{1}{\lambda} \left[1 - 2 \frac{(17 + 27 \ln 4/3)}{81} \frac{\ln |\ln \lambda|}{|\ln \lambda|} + O\left(\frac{1}{\ln \lambda}\right) \right]. \quad (17.79)$$

A short calculation then yields, for example, the susceptibility in zero field,

$$\chi^{-1} \sim C_{\pm} |\ln \tau|^{-1/3} \left[1 + \frac{1}{243} (108 \ln(4/3) + 41) \frac{\ln |\ln \tau|}{|\ln \tau|} + O\left(\frac{1}{|\ln \tau|}\right) \right], \quad (17.80)$$

or the specific heat,

$$C = A_{\pm} |\ln \tau|^{1/3} \left[1 - \frac{1}{243} (108 \ln(4/3) + 41) \frac{\ln |\ln \tau|}{|\ln \tau|} + O\left(\frac{1}{|\ln \tau|}\right) \right], \quad (17.81)$$

with the universal ratio

$$\frac{A_+}{A_-} = \frac{1}{4}. \quad (17.82)$$

The specific heat has been measured in a high precision experiment on the dipolar Ising ferromagnet LiTbF₄ by Ahlers *et al* [176]. Fitting the specific heat by

$$\begin{aligned} C_+ &= \frac{A_+}{b^z} \{ [1 + b \ln(a/|\tau|)]^z - 1 \} + B, \\ C_- &= \frac{A_-}{b^{z'}} \left\{ [1 + b \ln(a/|\tau|)]^{z'} - \frac{1}{4} \right\} + B, \end{aligned}$$

they find

$$\begin{aligned} \frac{A_+}{A_-} &= 0.244 \pm 0.009, \\ z = z' &= 0.336 \pm 0.024, \end{aligned}$$

results which agree nicely with the theoretical predictions.

18 $O(N)$ -symmetric vector models for N large

In Chapters 15–17, we have derived universal properties of $O(N)$ symmetric critical systems within the framework of the formal $\varepsilon = 4 - d$ expansion. Therefore, it is reassuring to verify that the results obtained in this way remain valid, at least in some limiting case, even when ε is not infinitesimal. We thus show in this chapter that, in the example of the $O(N)$ symmetric $(\phi^2)^2$ field theory, the same universal properties can also be derived at fixed dimension in the large N limit and, more generally, order by order in an $1/N$ -expansion. For example, early calculations of critical exponents are reported in Refs. [177–179]. For early reviews, see, for example, the contributions Refs. [180]. Moreover, large N techniques are also useful, because they provide an insight into other non-perturbative questions, including issues relevant to four-dimensional physics like renormalons and triviality.

Using a large N expansion, we exhibit a remarkable relation between the $(\phi^2)^2$ field theory and the non-linear σ -model (Chapter 19), valid to all orders.

Large N techniques can be generalized to many systems like fermion theories (see Chapter 20), to critical dynamics and so on (for a recent general review, see Ref. [181]), provided the fields, or order parameters, transform by the fundamental representations of the $O(N)$ or $U(N)$ groups. By contrast, theories involving $N \times N$ matrices have found solutions, for N large, only in some examples in zero and one dimensions [182].

18.1 The large N action

We consider the partition function of an $O(N)$ symmetric model,

$$\mathcal{Z} = \int [d\phi(x)] \exp [-\mathcal{S}(\phi)], \quad (18.1)$$

where ϕ is a N -component vector field, and where the Euclidean action $\mathcal{S}(\phi)$ has the form

$$\mathcal{S}(\phi) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + NU(\phi^2(x)/N) \right] d^d x, \quad (18.2)$$

$U(\rho)$ being a general positive polynomial. The reason for the peculiar N dependence will become clear shortly. A cut-off Λ , consistent with the symmetry, is implied.

For

$$U(\rho) = \frac{1}{2}r\rho + \frac{1}{4!}u\rho^2, \quad \text{with } u > 0, \quad (18.3)$$

one recovers the action (15.41) of the $(\phi^2)^2$ field theory,

$$\mathcal{S}(\phi) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} u (\phi^2(x))^2 / N \right] d^d x, \quad (18.4)$$

with

$$u = N\Lambda^{4-d}g. \quad (18.5)$$

The solution of the general model, in the large N limit, is based on an idea of mean-field approximation: it is expected that, for N large, $O(N)$ invariant quantities self-average and, therefore, have small fluctuations. For example,

$$\langle \phi^2(x) \phi^2(y) \rangle \underset{N \rightarrow \infty}{\sim} \langle \phi^2(x) \rangle \langle \phi^2(y) \rangle.$$

This argument suggests taking $\phi^2(x)$ as a dynamical variable, in place of ϕ . We thus introduce two additional fields, λ and ρ , and impose the condition $N\rho(x) = \phi^2(x)$ by an integral over λ . For each point x in space, we use the identity,

$$1 = N \int d\rho \delta(\phi^2 - N\rho) = \frac{N}{4i\pi} \int d\rho d\lambda e^{\lambda(\phi^2 - N\rho)/2}, \quad (18.6)$$

where the integration contour in the complex λ plane is parallel to the imaginary axis. The insertion of this identity into the field integral (18.1) yields another representation of the partition function [51],

$$\mathcal{Z} = \int [d\phi][d\rho][d\lambda] \exp [-\mathcal{S}(\phi, \rho, \lambda)], \quad (18.7)$$

where

$$\mathcal{S}(\phi, \rho, \lambda) = \int \left[\frac{1}{2} (\nabla \phi(x))^2 + NU(\rho(x)) + \frac{1}{2} \lambda(x) (\phi^2(x) - N\rho(x)) \right] d^d x. \quad (18.8)$$

The action is then quadratic in the field and the Gaussian integration over ϕ can be performed; the dependence on N of the partition function becomes explicit. Actually, it is convenient to separate the components of ϕ into one component σ , and $(N - 1)$ components $\boldsymbol{\pi}$, and integrate over $\boldsymbol{\pi}$ only (for $T < T_c$, it may even be convenient to integrate over only $(N - 2)$ components). For N large, the difference is negligible. To generate σ correlation functions, we also add a source $H(x)$ (a space-dependent magnetic field) to the action. After integration over $\boldsymbol{\pi}(x)$, the partition function in a field becomes

$$\mathcal{Z}(H) = \int [d\sigma][d\rho][d\lambda] \exp \left[-\mathcal{S}_N(\sigma, \rho, \lambda) + \int d^d x H(x) \sigma(x) \right], \quad (18.9)$$

with

$$\begin{aligned} \mathcal{S}_N(\sigma, \rho, \lambda) = & \int \left[\frac{1}{2} (\nabla \sigma(x))^2 + NU(\rho(x)) + \frac{1}{2} \lambda(x) (\sigma^2(x) - N\rho(x)) \right] d^d x \\ & + \frac{1}{2} (N - 1) \text{tr} \ln (-\nabla^2 + \lambda(\cdot)). \end{aligned} \quad (18.10)$$

We take the large N limit with a fixed function $U(\rho)$.

Ansatz. We assume that the expectation values of fields scale with N like

$$\sigma = O(N^{1/2}), \quad \rho, \lambda = O(1). \quad (18.11)$$

With this ansatz, \mathcal{S}_N is of order N for N large, and the field integral can be calculated by the *steepest descent method*.

ϕ^2 field and energy operator. In this formalism, the ρ -field, which is proportional to the ϕ^2 field, is thus the energy operator.

Other methods. The large N limit itself can be obtained by several other algebraic methods but, in contrast with the method we explain here, several of these methods cannot be extended beyond leading order. For example, Schwinger–Dyson equations, for N large, lead to a self-consistent equation for the two-point function. From the point of view of stochastic quantization or critical dynamics, the Langevin equation also becomes linear and self-consistent for N large. One replaces $\phi^2(x, t)$ by $\langle \phi^2(x, t) \rangle$ ($\langle \cdot \rangle$ means noise expectation value) at leading order. A variational principle of mean-field type, designed to provide the best Gaussian approximation, also yields the large N result.

More general $O(N)$ -invariant scalar field theories. The algebraic method described here can easily be generalized to actions that depend on several N -component fields. In a general $O(N)$ symmetric field theory, the composite fields with small fluctuations are the $O(N)$ invariant polynomials. One thus introduces pairs of fields and Lagrange multipliers for all independent scalar products constructed from the N -component fields.

For example, in the case of two fields ϕ_1, ϕ_2 , an $O(N)$ -invariant potential is only a function of the three $O(N)$ invariant quantities, the scalar products $\phi_1 \cdot \phi_2, \phi_1^2$, and ϕ_2^2 . One then introduces three fields $\rho_{ij}(x)$, and implements in the field integral the relations

$$N\rho_{ij}(x) = \phi_i(x) \cdot \phi_j(x),$$

by integrating over three $\lambda_{ij}(x)$ fields, as in the identity (18.6). One then substitutes $N\rho_{ij}(x)$ for $\phi_i(x) \cdot \phi_j(x)$ in the interactions. The field integral becomes Gaussian in the fields ϕ_i , the integration over ϕ_i can be performed and the dependence in N becomes explicit. The large N limit again can be derived by the steepest descent method.

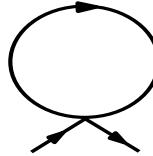


Fig. 18.1 The one-loop diagram Ω_d

18.2 Large N limit: Saddle point equations, critical domain

We look for uniform saddle points $(\sigma(x), \rho(x), \lambda(x)$ space-independent), $\sigma(x) = \sigma$, $\rho(x) = \rho$, and $\lambda(x) = m^2$, because the λ saddle point value must be positive.

18.2.1 Saddle point equations in zero field

The action density \mathcal{E} in zero field ($H = 0$) then becomes

$$\mathcal{E}(\sigma, \rho, m^2, \Lambda) = NU(\rho) + \frac{1}{2}m^2(\sigma^2 - N\rho) + \frac{1}{2}\frac{N}{(2\pi)^d} \int^\Lambda d^d k \ln(k^2 + m^2), \quad (18.12)$$

where the symbol \int^Λ means that the momentum integral has been regularized with a cut-off Λ . We choose, as a regularization, to replace in the action ∇ by $\nabla_\Lambda = \nabla \kappa^{1/2}(-\nabla^2/\Lambda^2)$, where $\kappa(s)$ is a positive polynomial for $s \geq 0$ and $\kappa(0) = 1$.

A differentiation of \mathcal{E} with respect to σ, ρ and m^2 then yields the saddle point equations,

$$m^2\sigma = 0, \quad (18.13)$$

$$U'(\rho) - \frac{1}{2}m^2 = 0, \quad (18.14)$$

$$\sigma^2/N - \rho + \Omega_d(m, \Lambda) = 0, \quad (18.15)$$

where $\Omega_d(m, \Lambda)$ is the diagram of Fig. 18.1, which appears in all large N discussions:

$$\Omega_d(m, \Lambda) = \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 \kappa(p^2/\Lambda^2) + m^2}. \quad (18.16)$$

One verifies that the equations (18.13–18.15) have solutions scaling in N consistently with the ansatz (18.11).

Discussion. From the origin of the tr ln term in expression (18.10), we infer that, at leading order, m is the mass of the $(N - 1)$ π components. Equation (18.13) implies either $\sigma = 0$ or $m = 0$. If $\sigma = 0$ and $m \neq 0$, the $O(N)$ symmetry is unbroken and the N fields ϕ_i have the same mass m . By contrast, if $\sigma \neq 0$, then $m = 0$, the $O(N)$ symmetry is spontaneously broken, and the π -field is massless. The appearance of these $(N - 1)$ massless fields is a manifestation of the Goldstone phenomenon. However, equation (18.15) shows that the solution $m = 0$ exists only for $d > 2$, because at $d = 2$, the integral is infrared (IR) divergent. This property is consistent with Mermin–Wagner–Coleman’s theorem [118]: in a system with only short range forces a continuous symmetry cannot be broken for $d \leq 2$, in the sense that the expectation value σ of the order parameter necessarily vanishes. In the large N limit, the origin of the property is clear: the Goldstone modes would be massless, as we expect from general arguments and verify here, and this induces an IR instability for $d \leq 2$.

18.2.2 The $(\phi^2)^2$ field theory: Phases and exponents

The discussion from now on is specific to the example of the $(\phi^2)^2$ field theory, where $U(\rho)$ is given by equation (18.3). We note that the large N limit is taken at uN fixed. Moreover, we assume $d > 2$, except if stated explicitly otherwise.

In this example, the ρ -integral is Gaussian and can be performed. The result of the integration corresponds to replace $\rho(x)$ by the solution of the classical equation

$$\lambda(x) = r + \frac{1}{6}u\rho(x), \quad (18.17)$$

and equation (18.14) becomes

$$m^2 = r + \frac{1}{6}u\rho. \quad (18.18)$$

After elimination of $\rho(x)$, the large N action reduces to

$$\begin{aligned} \mathcal{S}_N(\sigma, \lambda) &= \frac{1}{2} \int d^d x \left[(\nabla_\Lambda \sigma(x))^2 + \lambda(x)\sigma^2(x) - \frac{3N}{u}\lambda^2(x) + \frac{6N}{u}r\lambda(x) \right] \\ &\quad + \frac{1}{2}(N - 1) \text{tr ln}(-\nabla_\Lambda^2 + \lambda(\cdot)). \end{aligned} \quad (18.19)$$

The low temperature phase: $d > 2$. We first assume that σ , the expectation value of the field, does not vanish and thus $m = 0$. Equations (18.15) and (18.18) then yield

$$\sigma^2/N = \rho - \rho_c = (6/u)(r_c - r), \quad (18.20)$$

where we have introduced the constants ρ_c and r_c ,

$$\rho_c = \Omega_d(0, \Lambda), \quad r_c = -u\rho_c/6 < 0. \quad (18.21)$$

For $\rho - \rho_c > 0$, that is, $r < r_c$, the field $\sigma(x)$ has a non-vanishing expectation value, and the symmetry is spontaneously broken. The value $r = r_c$, where the expectation value vanishes, corresponds to the critical temperature T_c . In terms of the parameter

$$\tau = 6(r - r_c)/u, \quad (18.22)$$

which characterizes the deviation from the critical temperature, one finds

$$\sigma = \sqrt{-\tau} \equiv (-\tau)^\beta \Rightarrow \beta = \frac{1}{2}. \quad (18.23)$$

In the N large limit, the exponent β remains mean-field like, in all dimensions $d > 2$.

The high-temperature phase for $d > 2$. Above T_c , σ vanishes. From expression (18.19), one then infers the σ -propagator. In the Fourier representation, for $\Lambda \rightarrow \infty$, one finds

$$\tilde{\Delta}_\sigma(p) = \frac{1}{p^2 + m^2}. \quad (18.24)$$

It has a free-field form, and, at this order, m is the physical mass, or the inverse of the correlation length ξ , of the field σ and thus of all components of the ϕ field.

In the critical limit, $\tau = 0$, m vanishes and thus from the form (18.24) of the σ -propagator, one finds

$$\eta = 0 \Rightarrow d_\phi = \frac{1}{2}(d - 2). \quad (18.25)$$

The critical exponent η remains Gaussian in all dimensions d .

From equation (18.15), one infers that $\partial\rho/\partial m$ is negative, and thus r is an increasing function of m . The minimum value of r , obtained for $m = 0$, is r_c .

Using the definitions (18.21) and (18.22) inside equations (18.18) and (18.15), one then finds

$$m^2 = (u/6)(\rho - \rho_c + \tau), \quad (18.26a)$$

$$\rho - \rho_c = -m^2 D_d(m, \Lambda), \quad (18.26b)$$

where $D_d(m, \Lambda)$ is given by

$$\begin{aligned} D_d(m, \Lambda) &= \frac{1}{m^2}(\Omega_d(0, \Lambda) - \Omega_d(m, \Lambda)) \\ &= \frac{1}{(2\pi)^d} \int \frac{d^d p}{[p^2 \kappa(p^2/\Lambda^2) + m^2][p^2 \kappa(p^2/\Lambda^2)]}. \end{aligned} \quad (18.27)$$

For $m \ll \Lambda$, the leading contributions are

$$D_d(m, \Lambda) = C(d)m^{d-4} - a(d)\Lambda^{d-4} + O(m^2\Lambda^{d-6}), \quad (18.28)$$

where $a(d)$ depends explicitly on the regularization scheme, and $C(d)$ is universal. For $d < 4$, it can be calculated for $\Lambda \rightarrow \infty$, and one finds

$$C(d) = -\frac{\pi}{2 \sin(\pi d/2)} N_d = -\frac{1}{(4\pi)^{d/2}} \Gamma(1 - d/2), \quad (18.29)$$

where $N_d = 2/(4\pi)^{d/2}\Gamma(d/2)$ is the usual loop factor. For $d < 4$, $a(d)$ characterizes the leading correction in equation (18.28) and, for $d \rightarrow 4$, its divergence ($a(d) \sim 1/8\pi^2(4-d)$) cancels the divergence of $C(d)$.

(i) For $d > 4$, the function $D_d(m, \Lambda)$ has a limit for $m = 0$ (equation (18.28)). This implies $\rho - \rho_c \sim a(d)\Lambda^{d-4}m^2$ and, therefore, at leading order:

$$m = 1/\xi \propto \sqrt{\tau} \Rightarrow \nu = \frac{1}{2}. \quad (18.30)$$

For $d > 4$, the correlation exponent ν has a mean-field or Gaussian value.

(ii) By contrast, for $2 < d < 4$, the function $D_d(m, \Lambda)$ for m small is given by the first term in the expansion (18.28),

$$D_d(m, \Lambda) \sim C(d)m^{d-4}. \quad (18.31)$$

The leading m -dependent contribution, for $m \rightarrow 0$, in equation (18.26a) now comes from $\rho - \rho_c$. One concludes

$$m = \xi^{-1} \sim \tau^{1/(d-2)}, \quad (18.32)$$

which shows that the correlation exponent ν is no longer Gaussian:

$$\nu = \frac{1}{d-2}. \quad (18.33)$$

One verifies that the exponents β, ν, η satisfy the scaling relation $\beta = \nu d_\phi$, proved within the framework of the ε -expansion.

(iii) For $d = 4$, the leading m -dependent contribution in equation (18.26a) still comes from $\rho - \rho_c$. Moreover,

$$D_4(m, \Lambda) \underset{m \rightarrow 0}{\sim} \frac{1}{8\pi^2} \ln(\Lambda/m).$$

The correlation length no longer has a power-law behaviour but, by contrast, a mean-field behaviour modified by a logarithm. This is typical of a situation where the Gaussian fixed point is stable in the presence of a marginal operator.

Two dimensions. Examining equation (18.15) for $\sigma = 0$ and $d = 2$, one finds that the correlation length becomes large only for $r \rightarrow -\infty$. This peculiar situation will be discussed in the framework of the non-linear σ -model (Chapter 19).

18.2.3 Singular free energy and scaling equation of state

In a constant magnetic field H in the σ direction, the free energy density $W(H)$ is given by (V_d is the space volume)

$$W(H) = \lim_{V_d \rightarrow \infty} \frac{\ln \mathcal{Z}(H)}{V_d} = -NU(\rho) - \frac{1}{2}m^2(\sigma^2 - N\rho) + H\sigma - \frac{N}{2(2\pi)^d} \int^\Lambda d^d p \ln(1+m^2/p^2),$$

where the saddle point values σ, ρ, m^2 are the solutions of equations (18.14) and (18.15) and the modified saddle point equation (18.13),

$$m^2\sigma = H. \quad (18.34)$$

The magnetization M is given by

$$M = \frac{\partial W}{\partial H} = \sigma,$$

because partial derivatives of \mathcal{W} with respect to m^2, σ vanish as a consequence of the saddle point equations. The thermodynamic potential density $\mathcal{G}(M)$ is the Legendre transform of $W(H)$.

In the $(\phi^2)^2$ field theory, using the form (18.19) of the action, one finds

$$\begin{aligned} \mathcal{G}(M)/N &= \lim_{V_d \rightarrow \infty} \Gamma(M)/NV_d = (HM - W(H))/N \\ &= -\frac{3}{2u}m^4 + \frac{3r}{u}m^2 + \frac{1}{2}m^2M^2/N + \frac{1}{2}\frac{1}{(2\pi)^d} \int^\Lambda d^d p \ln[(p^2 + m^2)/p^2]. \end{aligned}$$

We note that (equations (18.21, 18.27))

$$\frac{d}{dm^2} \frac{1}{(2\pi)^d} \int^\Lambda d^d p \ln[(p^2 + m^2)/p^2] = \Omega(m, \Lambda) = \rho_c - m^2 D_d(m, \Lambda).$$

The integral can thus be evaluated for Λ large in terms of r_c and the expansion (18.31):

$$\frac{1}{(2\pi)^d} \int^\Lambda d^d p \ln[(p^2 + m^2)/p^2] = -2 \frac{C(d)}{d} m^d - \frac{6r_c}{u} m^2 + \frac{a(d)}{2} m^4 \Lambda^{d-4} + O(m^6 \Lambda^{d-6}).$$

The thermodynamic potential density becomes

$$\mathcal{G}(M)/N = \frac{3}{2} \left(\frac{1}{u^*} - \frac{1}{u} \right) m^4 + \frac{3(r - r_c)}{u} m^2 + \frac{1}{2} m^2 M^2/N - \frac{C(d)}{d} m^d, \quad (18.35)$$

where we have defined

$$u^* = \frac{6}{a(d)} \Lambda^{4-d}. \quad (18.36)$$

Dimensions $d < 4$. We now discuss only the physical situation $d < 4$. Then, for m small, the term proportional to m^4 is negligible with respect to the singular term m^d . At leading order in the critical domain

$$\mathcal{G}(M)/N = \frac{1}{2} \tau m^2 + \frac{1}{2} m^2 M^2/N - C(d)m^d/d, \quad (18.37)$$

where τ , the deviation from the critical temperature, is defined by equation (18.22).

The general property of the Legendre transformation and the saddle point equation (18.15) imply that the derivative of Γ with respect to m vanishes:

$$\frac{2}{N} \frac{\partial \mathcal{G}}{\partial m^2} = \tau + M^2/N - C(d)m^{d-2} = 0,$$

an equation which makes it possible to calculate m directly from \mathcal{G} :

$$m = \left[\frac{1}{C(d)} (\tau + M^2/N) \right]^{1/(d-2)}.$$

It follows that the thermodynamic potential density, at leading order in the critical domain, is given by

$$\mathcal{G}(M) \sim N \frac{(d-2)}{2d} \frac{1}{(C(d))^{2/(d-2)}} (\tau + M^2/N)^{d/(d-2)}. \quad (18.38)$$

Various quantities can be derived from $\mathcal{G}(M)$, for example, the equation of state by differentiating with respect to M . The resulting equation of state has the expected scaling form

$$H = \frac{\partial \mathcal{G}}{\partial M} = h_0 M^\delta f(\tau/M^2), \quad (18.39)$$

in which h_0 is a normalization constant, the exponent δ is given by

$$\delta = \frac{d+2}{d-2}, \quad (18.40)$$

in agreement with the general scaling relation $\delta = d/d_\phi - 1$, and the function $f(x)$ is

$$f(x) = (1+x)^{2/(d-2)}. \quad (18.41)$$

The asymptotic form of $f(x)$ for x large implies $\gamma = 2/(d-2)$, again in agreement with the scaling relation $\gamma = \nu(2-\eta)$. Taking into account the values of the critical exponents γ and β , it is then easy to verify that the function f satisfies all required properties like, for example, Griffith's analyticity (see Section 16.5). In particular, the equation of state can be cast into the parametric form $\{M, \tau\} \mapsto \{R, \theta\}$ with

$$M = R^\beta \theta, \quad \tau = 3R(1 - \theta^2),$$

and, therefore,

$$H = h_0 R^{\beta\delta} \theta (3 - 2\theta^2)^{2/(d-2)}.$$

Leading corrections to scaling. The m^4 term in equation (18.35) yields the leading corrections to the scaling behaviour. It is subleading by the power of τ ,

$$m^4/m^d = O(\tau^{(4-d)/(d-2)}).$$

We conclude

$$\omega\nu = (4-d)/(d-2) \Rightarrow \omega = 4-d. \quad (18.42)$$

We have identified the exponent ω which governs the leading corrections to scaling. Note that, for the special value $u = u^*$, this correction vanishes.

Specific heat exponent. Amplitude ratios. Differentiating twice $\mathcal{G}(M)$ with respect to τ , one obtains the specific heat \mathcal{C} at fixed magnetization:

$$\mathcal{C} \propto (\tau + M^2)^{(4-d)/(d-2)}. \quad (18.43)$$

For $M = 0$, we identify the specific exponent

$$\alpha = \frac{4-d}{d-2}, \quad (18.44)$$

which indeed is equal to $2 - d\nu$, as predicted by scaling relations. Among the ratio of amplitudes, one can calculate, for example, R_ξ^+ and R_c (for definitions, see Section 16.9.4):

$$(R_\xi^+)^d = \frac{4N}{(d-2)^3} \frac{\Gamma(3-d/2)}{(4\pi)^{d/2}}, \quad R_c = \frac{4-d}{(d-2)^2}. \quad (18.45)$$

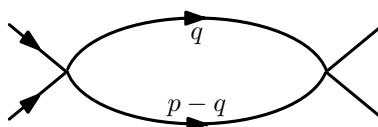


Fig. 18.2 The ‘bubble’ diagram $B_\Lambda(p, m)$

18.2.4 The ϕ^2 two-point function

Differentiating twice the action (18.19) with respect to $\lambda(x)$, then replacing the field $\lambda(x)$ by its expectation value m^2 , one obtains the λ -propagator $\tilde{\Delta}_\lambda(p)$ above T_c ,

$$\tilde{\Delta}_\lambda(p) = -\frac{2}{N} \left(\frac{6}{u} + B_\Lambda(p, m) \right)^{-1}, \quad (18.46)$$

where $B_\Lambda(p, m)$ is the ‘bubble’ diagram of Fig. 18.2:

$$B_\Lambda(p, m) = \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d q}{(q^2 + m^2) [(p - q)^2 + m^2]}. \quad (18.47)$$

The λ -propagator is negative, because the λ -field is imaginary. The λ -field propagator is related by equation (18.17) to the ρ , and thus, to the ϕ^2 two-point function, which is given by

$$\langle \tilde{\phi}^2(p) \tilde{\phi}^2(-p) \rangle = -\frac{12N}{u [1 + (u/6)B_\Lambda(p, m)]}. \quad (18.48)$$

At zero momentum, one recovers the specific heat up to an additive constant.

In the critical theory ($m = 0$ at this order), for $2 \leq d \leq 4$, at low momentum the denominator of $\tilde{\Delta}_\lambda$ is dominated by the integral

$$B_\Lambda(p, 0) = \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d q}{q^2(p - q)^2} \underset{2 < d < 4}{=} b(d)p^{-\varepsilon} - a(d)\Lambda^{-\varepsilon} + O(\Lambda^{d-6}p^2), \quad (18.49)$$

where

$$b(d) = -\frac{\pi}{\sin(\pi d/2)} \frac{\Gamma^2(d/2)}{\Gamma(d-1)} N_d, \quad (18.50)$$

and thus

$$\tilde{\Delta}_\lambda(p) \underset{p \rightarrow 0}{\sim} -\frac{2}{Nb(d)} p^\varepsilon. \quad (18.51)$$

We again verify consistency with scaling relations. In particular, we note that, in the large N limit, the dimension $[\lambda]$ of the field λ is

$$[\lambda] = \frac{1}{2}(d + \varepsilon) = 2, \quad (18.52)$$

a result important for the $1/N$ perturbation theory.

Remarks.

(i) For $d = 4$, the behaviour of the propagator is still dominated by the integral, which has a logarithmic behaviour $\tilde{\Delta}_\lambda \propto 1/\ln(\Lambda/p)$.

(ii) Therefore, note that, for $d \leq 4$, the contributions generated by the term proportional to $\lambda^2(x)$ in action (18.19) are always negligible in the critical domain.

18.3 Renormalization group (RG) functions and leading corrections to scaling

The RG functions. For a more detailed verification of the consistency of the large N results with the RG framework, we now calculate RG functions at leading order. One first verifies that, at leading order for Λ large, the solution m of equation (18.26) satisfies (setting $\varepsilon = 4 - d$)

$$\Lambda \frac{\partial m}{\partial \Lambda} + \varepsilon a(d) \Lambda^{-\varepsilon} \frac{u^2}{6} \frac{\partial m}{\partial u} = 0,$$

where the constant $a(d)$, defined in equation (18.31) depends on the cut-off procedure, but for ε small satisfies

$$a(d) \sim 1/(8\pi^2\varepsilon). \quad (18.53)$$

Then, we set (equations (18.5, 18.36))

$$u = Ng\Lambda^\varepsilon, \quad g^* = u^* \Lambda^{-\varepsilon}/N = 6/(Na(d)). \quad (18.54)$$

In the new variables Λ, g, τ , we obtain an equation which expresses that m is RG invariant,

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right) m(\tau, g, \Lambda) = 0, \quad (18.55)$$

with

$$\beta(g) = -\varepsilon g(1 - g/g^*), \quad (18.56)$$

$$\nu^{-1}(g) = 2 + \eta_2(g) = 2 - \varepsilon g/g^*. \quad (18.57)$$

When $a(d)$ is positive (but this not true for all regularizations, see the following discussion), one finds an IR fixed point g^* , as well as exponents $\omega = \varepsilon$, and $\nu^{-1} = d - 2$, in agreement with equations (18.42) and (18.33). In the framework of the ε -expansion, and thus for N finite but $(4 - d)$ small, ω is associated with the leading corrections to scaling. In the large N limit, ω remains smaller than 2 for $\varepsilon < 2$, and this extends the property for N large to all dimensions $2 \leq d \leq 4$.

Finally, applying the RG equations to the propagator (18.24), one obtains

$$\eta(g) = 0, \quad (18.58)$$

a result consistent with the value (18.25) found for η .

Leading corrections to scaling. From the general RG analysis, we expect the leading corrections to scaling to vanish for $u = u^*$. This property has already been verified for the free energy. We now consider the correlation length or mass m given by equation (18.26). If we keep the leading correction to the integral for m small (equation (18.31)), we find

$$\frac{6}{u} - \frac{6}{u^*} + C(d)m^{-\varepsilon} + O(m^2\Lambda^{d-6}) = \frac{\tau}{m^2}, \quad (18.59)$$

where equation (18.54) has been used. We see that the leading correction again vanishes for $u = u^*$. Actually, for $d \rightarrow 4$, all correction terms suppressed by powers of order ε vanish simultaneously, as expected from the RG analysis of the ϕ^4 field theory. Moreover, one verifies that the leading correction is proportional to $(u - u^*)\tau^{\varepsilon/(2-\varepsilon)}$, which leads to $\omega\nu = \varepsilon/(2 - \varepsilon)$, in agreement with equations (18.42) and (18.33).



Fig. 18.3 Diagrams contributing to the λ propagator in the large N limit

Similarly, if one keeps the leading correction to the λ -propagator in the critical theory (equation (18.49)), one finds

$$\tilde{\Delta}_\lambda(p) = -\frac{2}{N} \left(\frac{6}{u} - \frac{6}{u^*} + b(d)p^{-\varepsilon} \right)^{-1}, \quad (18.60)$$

where terms of order Λ^{-2} have been neglected. The leading corrections to the scaling behaviour again exactly cancel for $u = u^*$, as expected.

Diagrammatic interpretation. In the large N limit of the $(\phi^2)^2$ field theory, the λ propagator is given by the perturbative contributions coming from chains of ‘bubble’ diagrams of the form displayed in Fig. 18.3. These diagrams asymptotically form a geometric series which is summed by the algebraic techniques described in Section 18.1.

Discussion.

(i) One can show that a perturbation due to irrelevant operators is equivalent, at leading order in the critical region, to a modification of the $(\phi^2)^2$ coupling. This can be explicitly verified here. The amplitude of the leading correction to scaling has been found to be proportional to $6/u - a(d)\Lambda^{-\varepsilon}$, where the value of $a(d)$ depends on the cut-off procedure, and thus of contributions of irrelevant operators. We denote by u' the $(\phi^2)^2$ coupling constant in another scheme where a is replaced by a' . Identifying the leading correction to scaling, we find the homographic relation:

$$\frac{6\Lambda^\varepsilon}{u} - a(d) = \frac{6\Lambda^\varepsilon}{u'} - a'(d),$$

consistent with the special form (18.56) of the β -function.

(ii) *The sign of $a(d)$.* It is generally assumed that $a(d)$ is positive. This is indeed what one finds in some regularization schemes, like the simplest Pauli–Villars’s regularization where $a(d)$ is positive in all dimensions $2 < d < 4$. Moreover, $a(d)$ is always positive near four dimensions where it diverges like

$$a(d) \underset{d \rightarrow 4}{\sim} \frac{1}{8\pi^2\varepsilon}.$$

Then, there exists an IR fixed point, non-trivial zero of the β -function. For this value u^* , the leading corrections to scaling vanish.

However, for d fixed, $d < 4$, this is not a universal feature. For example, in the case of simple lattice regularizations, it has been shown that in $d = 3$ the sign is arbitrary. Then, if $a(d)$ is negative, the RG method for large N (at least in the perturbative framework) is confronted with a serious difficulty. Indeed, the coupling flows in the IR limit to large values where the large N expansion is no longer reliable. It is not known whether this signals a real physical problem, or is just an artefact of the large N limit.

Another way of stating the problem is to directly examine the relation between bare and renormalized coupling constant. Denoting by $g_{\text{r}} m^{4-d}$ the renormalized four-point function at zero momentum, one finds

$$m^{4-d} g_{\text{r}} = \frac{\Lambda^{4-d} g}{1 + \Lambda^{4-d} g N B_\Lambda(0, m)/6}. \quad (18.61)$$

In the limit $m \ll \Lambda$, the relation can be written as

$$\frac{1}{g_r} = \frac{(d-2)NC(d)}{12} + \left(\frac{m}{\Lambda}\right)^{4-d} \left(\frac{1}{g} - \frac{Na(d)}{6}\right). \quad (18.62)$$

We see that for $a(d) < 0$, the renormalized IR fixed point value cannot be reached by varying $g > 0$ for any finite value of m/Λ . In the same way, leading corrections to scaling can no longer be cancelled.

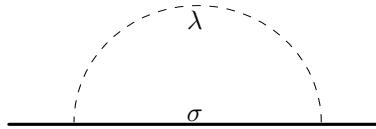


Fig. 18.4 The diagram contributing to $\Gamma_{\sigma\sigma}^{(2)}$ at order $1/N$: full line and dotted line correspond to the σ and the λ propagators, respectively

18.4 Small-coupling constant, large-momentum expansions for $d < 4$

In Section 18.6.6, we discuss more systematically the $1/N$ expansion, but the $1/N$ correction to the two-point function can already help us understanding why the massless field theory exists, but has no perturbative expansion for $d < 4$.

In the framework at the $1/N$ expansion, we have been able to define a critical theory ($T = T_c, m^2 = 0$) at fixed dimension $d < 4$, while the usual perturbative expansion is IR divergent. Note that, since the Gaussian fixed point is an ultraviolet (UV) fixed point, the small coupling expansion is also a large momentum expansion. To understand the phenomenon, we consider the $\langle\sigma\sigma\rangle$ correlation function at order $1/N$. At this order, only one diagram contributes (Fig. 18.4), containing two $\lambda^2\sigma$ vertices. After mass renormalization and in the large cut-off limit, one finds

$$\tilde{\Gamma}_{\sigma\sigma}^{(2)}(p) = p^2 + \frac{2}{N(2\pi)^d} \int \frac{d^d q}{(6/u) + b(d)q^{-\varepsilon}} \left(\frac{1}{(p+q)^2} - \frac{1}{q^2} \right) + O\left(\frac{1}{N^2}\right). \quad (18.63)$$

The integral cannot be expanded in a Taylor series in the coupling constant u . By contrast, an analytic study reveals that it has a small u expansion of the form

$$\sum_{k \geq 1} \alpha_k u^k p^{2-k\varepsilon} + \beta_k u^{(2+2k)/\varepsilon} p^{-2k}. \quad (18.64)$$

The coefficients α_k, β_k can be obtained by calculating the Mellin transform with respect to u of the integral. Indeed, if a function $f(u)$ behaves like u^t for u small, then the Mellin transform

$$M(s) = \int_0^\infty du u^{-1-s} f(u)$$

has a pole at $s = t$. Applying the transformation to the integral, and inverting q and u integrations, one obtains the integral

$$\int_0^\infty du \frac{u^{-1-s}}{(6/u) + b(d)q^{-\varepsilon}} = \frac{1}{6} \left(\frac{b(d)q^{-\varepsilon}}{6} \right)^{1-s} \frac{\pi}{\sin \pi s}.$$

Then, the result of the remaining q integration follows from the evaluation of the generic integral (18.97).

The terms with integer powers of u correspond to the formal perturbative expansion, where each integral is calculated for ε small enough. The coefficient α_k has poles at $\varepsilon = (2l + 2)/k$, values for which the corresponding power of p^2 is $-l$, that is, an integer. One verifies that β_l has a pole at the same value of ε , and that the singular contributions cancel in the sum. In the dimensions corresponding to the poles, logarithms of u appear in the expansion.

18.5 Dimension 4: Triviality issue for N large

A number of issues concerning the physics of the $(\phi^2)^2$ theory in *four dimensions* can be examined within the framework of the large N expansion. For simplicity, we consider here only the critical (*i.e.*, massless) theory.

Triviality and Landau ghost. In the large N limit, the renormalized coupling constant g_r , defined as the value of the vertex $\langle \sigma \sigma \sigma \rangle$ at momenta of order $\mu \ll \Lambda$, is given by

$$g_r = \frac{g}{1 + \frac{1}{6} N g B_\Lambda(\mu, 0)}, \quad (18.65)$$

where $B_\Lambda(p, 0)$ corresponds to the bubble diagram (Fig. 18.2), and behaves like

$$B_\Lambda(p, 0) \underset{p \ll \Lambda}{\sim} \frac{1}{8\pi^2} \ln(\Lambda/p) + \text{const.} \quad (18.66)$$

Therefore, when the ratio μ/Λ goes to zero, the renormalized coupling constant g_r vanishes, for that all initial values of g . This is the so-called *triviality* property (see Section 9.12). If one works formally and, ignoring the problem, expresses the leading contribution to the four-point function in terms of the renormalized constant:

$$\frac{g}{1 + \frac{N}{48\pi^2} g \ln(\Lambda/p)} = \frac{g_r}{1 + \frac{N}{48\pi^2} g_r \ln(\mu/p)},$$

one discovers that the function has a pole for

$$p = \mu e^{48\pi^2/(Ng_r)}.$$

This pole is non-physical and, therefore, the corresponding ‘particle’ is often called the Landau ghost. It is characteristic of theories which have the free theory as an IR limit. At higher orders, this pole makes the loop integrals diverge, as the example of the renormalized two-point function at order $1/N$ (expression (18.63)),

$$\tilde{\Gamma}_{\sigma\sigma}^{(2)}(p) = p^2 + \frac{g_r}{3N(2\pi)^4} \int \frac{d^4 q}{1 + N g_r \ln(\mu/q)/48\pi^2} \left(\frac{1}{(p+q)^2} - \frac{1}{q^2} - \text{subtraction} \right),$$

explicitly shows (the expression has still to be subtracted to take into account field renormalization). In an expansion in powers of g_r , each term instead is finite but one finds, after renormalization, UV contributions of the type

$$\int^\infty \frac{d^4 q}{q^6} \left(-\frac{Ng_r}{48\pi^2} \ln(\mu/q) \right)^k \underset{k \rightarrow \infty}{\propto} \left(\frac{Ng_r}{96\pi^2} \right)^k k!.$$

The perturbative manifestation of the Landau ghost is the appearance of (*renormalon*) contributions to the perturbation series, which are not Borel summable (see Section 40.3.2, for details), in contrast with the semi-classical contributions corresponding to the finite momentum integration, which are Borel summable, but invisible for N large. Finally, we have considered the massless limit to simplify expressions, because the UV problem is independent of the mass of the field ϕ .

18.6 The $(\phi^2)^2$ field theory and the non-linear σ -model for N large

Within the framework of the $1/N$ expansion, one can establish a remarkable relation, valid to all orders, between the $(\phi^2)^2$ field theory and the non-linear σ -model discussed in Chapter 19. Note that a cut-off Λ is implicitly assumed, where needed.

18.6.1 The non-linear σ -model

In the action (18.19), we shift the field $\lambda(x)$ by its expectation value m^2 , $\lambda(x) \mapsto m^2 + \lambda(x)$. The action becomes

$$\begin{aligned} S_N(\sigma, \lambda) = & \frac{1}{2} \int d^d x \left[(\nabla \sigma(x))^2 + m^2 \sigma^2(x) + \lambda(x) \sigma^2(x) - \frac{3}{u} \lambda^2(x) - \frac{6}{u} (m^2 - r) \lambda(x) \right] \\ & + \frac{(N-1)}{2} \text{tr} \ln [-\nabla^2 + m^2 + \lambda(\cdot)]. \end{aligned} \quad (18.67)$$

We have pointed out that the term proportional to $\int d^d x \lambda^2(x)$, which has dimension $(4-d)$ for large N in all dimensions, is irrelevant in the critical domain for $d < 4$ (this also applies to $d = 4$, where it is marginal and yields only logarithmic corrections). Actually, the constant part in the inverse propagator as written in equation (18.60) plays the role of a large momentum cut-off. We thus neglect the λ^2 term. We work then backwards, reintroduce the initial field ϕ and integrate over $\lambda(x)$. We find, formally,

$$\mathcal{Z} = \int [d\phi(x)] \prod_x \delta [\phi^2(x) - 6(m^2 - r)/u] \exp \left[- \int \frac{1}{2} (\nabla \phi(x))^2 d^d x \right]. \quad (18.68)$$

In this form, we recognize the partition function of the $O(N)$ symmetric non-linear σ -model, which is discussed in Chapter 19, in an unusual parametrization. Therefore, we have discovered a remarkable correspondence: to all orders in a $1/N$ expansion, the renormalized non-linear σ -model is identical to the renormalized $(\phi^2)^2$ field theory at the IR fixed point, that is, for generic $(\phi^2)^2$ coupling.

18.6.2 The large N limit

We define now the partition function of the non-linear σ -model, with a more standard parametrization (see Section 19.1), as

$$\mathcal{Z} = \int [d\phi(x)] \prod_x \delta(\phi^2(x) - 1) \exp [-S(\phi)], \quad (18.69)$$

with

$$S(\phi) = \frac{1}{2T} \int d^d x (\nabla \phi(x))^2. \quad (18.70)$$

In this parametrization, the coupling constant T plays the role of the temperature, as the lattice regularization of the continuum σ model shows (Section 19.4.2). Therefore, this parametrization is well adapted to a study of low temperature phenomena.

We then solve the σ -model directly in the large N limit. The solution makes it possible to compare explicitly all expressions in the $(\phi^2)^2$ field theory and the non-linear σ -model parametrizations.

We rewrite the partition function as

$$\mathcal{Z} = \int [d\phi(x)d\lambda(x)] \exp [-\mathcal{S}(\phi, \lambda)], \quad (18.71)$$

with

$$\mathcal{S}(\phi, \lambda) = \frac{1}{2T} \int d^d x \left[(\nabla \phi(x))^2 + \lambda(x)(\phi^2(x) - 1) \right], \quad (18.72)$$

where we integrate over $\lambda(x)$ on the imaginary axis. Integrating, as we did in Section 18.1, over $(N - 1)$ components of ϕ and calling σ the remaining component, we obtain

$$\mathcal{Z} = \int [d\sigma(x)d\lambda(x)] \exp [-\mathcal{S}_N(\sigma, \lambda)], \quad (18.73)$$

with

$$\begin{aligned} \mathcal{S}_N(\sigma, \lambda) &= \frac{1}{2T} \int \left[(\nabla \sigma(x))^2 + (\sigma^2(x) - 1) \lambda(x) \right] d^d x \\ &\quad + \frac{1}{2}(N - 1) \text{tr} \ln [-\nabla^2 + \lambda(\cdot)]. \end{aligned} \quad (18.74)$$

The large N limit is taken here at TN fixed. The saddle point equations, analogues of equations (18.13–18.15), are

$$m^2\sigma = 0, \quad (18.75a)$$

$$\sigma^2 = 1 - (N - 1)\Omega_d(m, \Lambda)T, \quad (18.75b)$$

where we have set $\langle \lambda(x) \rangle = m^2$, and the function $\Omega_d(m, \Lambda)$ is defined, in a regularized form, by equation (18.16). At low temperature, for $d > 2$, σ is different from 0 and thus m , which is the mass of the π -field, vanishes. Equation (18.75b) gives the spontaneous magnetization:

$$\sigma^2 = 1 - (N - 1)\Omega_d(0, \Lambda)T. \quad (18.76)$$

Setting

$$1/T_c = (N - 1)\Omega_d(0, \Lambda), \quad (18.77)$$

we can write equation (18.76) as

$$\sigma^2 = 1 - T/T_c. \quad (18.78)$$

Thus, T_c is the critical temperature, where σ vanishes with an exponent $\beta = \frac{1}{2}$.

By contrast, above T_c , σ vanishes and m , which is now the common mass of the π - and σ -field, is, for $2 < d < 4$ and $\Lambda \gg m$, given in terms of the function (18.27), by

$$\frac{1}{T_c} - \frac{1}{T} = (N - 1)m^2 D_d(m, \Lambda) \underset{d < 4}{\sim} (N - 1)C(d)m^{d-2}. \quad (18.79)$$

For $T \rightarrow T_c$, one recovers the scaling form (18.32) of the correlation length $\xi = 1/m$.

Adding a magnetic field, we can calculate the free energy, $\mathcal{W}(H) = T \ln \mathcal{Z}(H)$, and the thermodynamic potential $\Gamma(M)$ after Legendre transformation,

$$\mathcal{G}(M) = \Gamma(M)/V_d = \frac{d-2}{2d} \frac{1}{(NC(d))^{2/(d-2)}} (M^2 - 1 + T/T_c)^{d/(d-2)}, \quad (18.80)$$

a result which extends equation (18.38) to all temperatures below T_c . The calculation of other physical quantities and the expansion in $1/N$ follow from the considerations of previous sections and Section 18.6.6.

18.6.3 Renormalization group

Anticipating slightly some results of Chapter 19, we calculate RG functions at leading order for N large. First, we introduce a dimensionless temperature t , setting

$$T = \Lambda^{2-d} t, \quad T_c = \Lambda^{2-d} t_c.$$

We then express that the mass or correlation length solutions of equation (18.79) are RG invariant:

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} \right) [(N-1)D_d(m)] = \left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} \right) \Lambda^{d-2} \left(\frac{1}{t_c} - \frac{1}{t} \right) = 0.$$

Therefore,

$$\beta(t) = (d-2)t(1-t/t_c). \quad (18.81)$$

From the RG point of view, $T = 0$ is an IR stable fixed point, and T_c an unstable IR fixed point, and also a stable UV fixed point. Since these questions are analysed thoroughly in Chapter 19, we postpone further discussion.

Conversely, we can solve the equation in general, and find

$$m(t) \propto \Lambda \exp \left[\int^t \frac{dt'}{\beta(t')} \right]. \quad (18.82)$$

Using the β -function (18.81), we find an explicit form that coincides for $m \ll \Lambda$ with the solution of equation (18.79) only for $d < 4$. This indicates that the RG equations (19.92) are valid only for $d < 4$.

For later use, we note that the RG-invariant crossover length (19.106) is given, in the large- N limit, for $d < 4$, by

$$\begin{aligned} \xi(t) &= \frac{t^{1/(d-2)}}{\Lambda} \exp \left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{(d-2)t'} \right) dt' \right] = \frac{1}{\Lambda} \left(\frac{1}{t} - \frac{1}{t_c} \right)^{-1/(d-2)} \\ &= \left(\frac{1}{T} - \frac{1}{T_c} \right)^{-1/(d-2)}. \end{aligned} \quad (18.83)$$

In the same way, σ^2 , which is the square of the one-point function, is expected to satisfy

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} + \zeta(t) \right) \sigma^2 = 0.$$

Inserting the result (18.78), we obtain

$$\zeta(t) = (d-2)t/t_c. \quad (18.84)$$

18.6.4 Two dimensions: Critical domain. Borel summability.

For a dimension d close to 2, the critical temperature behaves like $t_c \sim (2\pi/N)(d-2)$ and, for $d=2$, it vanishes. Then, equation (18.75b) determines the mass m . Since $\sigma=0$, one finds

$$m \propto \Lambda e^{-2\pi/(Nt)}, \quad (18.85)$$

in agreement with the RG predictions

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} \right) m(t, \Lambda) = 0, \quad \beta(t) = -Nt^2/2\pi.$$

The expression calls for two remarks. First, although the critical temperature vanishes, one can define a critical domain, that is, a range of temperature for which the correlation length is large with respect to the microscopic scale:

$$m \ll \Lambda \Rightarrow t = T = O(1/\ln(\Lambda/m)).$$

This is a low temperature range that shrinks when the correlation length increases.

Second, the field two-point function takes in the large N - limit, the form

$$\tilde{\Gamma}_{\sigma\sigma}^{(2)}(p) = p^2 + m^2. \quad (18.86)$$

If we substitute for m the large N expression (18.85), and expand the two-point function in powers of the coupling constant t , we note that the mass contribution vanishes to all orders, preventing any perturbative calculation of the mass of the field. The perturbation series is trivially not Borel summable. Most likely, this property remains true for the model at finite N . On the other hand, if we break the $O(N)$ symmetry by a magnetic field, adding a term $-h \int d^2x \sigma(x)/T$ to the action, the physical mass becomes calculable in perturbation theory,

$$\frac{h^2}{m^4} = 1 - \frac{Nt}{2\pi} (\ln(\Lambda/m) + \text{const.}).$$

18.6.5 Corrections to scaling and the dimension 4

In equation (18.79), we have neglected corrections to scaling. By contrast, if we take into account the leading correction, we find

$$m^2 (C(d)m^{d-4} - a(d)\Lambda^{d-4}) \propto t - t_c,$$

where $a(d)$, as we have already explained, is a constant that explicitly depends on the cut-off procedure and can thus be varied by changing contributions of irrelevant operators. By comparing with the results of Section 18.3, we discover that, although the non-linear σ -model superficially depends on one parameter less than the corresponding $(\phi^2)^2$ field theory, since the ϕ^4 coupling has no equivalent, actually this parameter is hidden in the shape of the cut-off function. Because the non-linear σ -model is not perturbatively renormalizable in $d > 2$, it is more sensitive to the cut-off function. Therefore, the identity between the ϕ^4 theory and the non-linear σ model extends beyond the IR fixed point. This remark becomes important in the four-dimensional limit where most leading contributions come from the leading corrections to scaling. For example, for $d=4$ equation (18.79) takes a different form, the dominant term in the right-hand side is proportional to $m^2 \ln m$. We recognize, in the factor $\ln m$, the effective ϕ^4 coupling at mass scale m .

However, at finite N , to describe with usual perturbation theory and RG the physics of the non-linear σ model, it is necessary to introduce the operator $\int d^d x \lambda^2(x)$, which although irrelevant for $d < 4$, becomes marginal, and to go over to the $(\phi^2)^2$ field theory.

18.6.6 Higher orders in the large N expansion: Power counting

Higher order terms in the steepest descent calculation of the field integral (18.9) generate a systematic and consistent $1/N$ expansion [183].

We thus analyse the terms in the action (18.67) from the point of view of large N power counting. The dimension of the field $\sigma(x)$ is $(d-2)/2$. From the critical behaviour (18.51) of the λ -propagator, we have inferred the canonical dimension $[\lambda]$ of the field $\lambda(x)$:

$$2[\lambda] - \varepsilon = d \quad \Rightarrow \quad [\lambda] = 2.$$

Therefore, λ^2 has dimension $4 > d$, and thus is irrelevant for $d < 4$ (and marginally irrelevant for $d = 4$). The interaction term $\int \lambda(x) \sigma^2(x) d^d x$ has dimension 0. One verifies that the non-local interactions involving the λ -field, coming from the expansion of the $\text{tr} \ln$, have all also canonical dimension 0:

$$\left[\text{tr} \left[\lambda(x) (-\nabla^2 + m^2)^{-1} \right]^k \right] = k[\lambda] - 2k = 0.$$

This power counting property has the following implication: in contrast with usual perturbation theory, the $1/N$ expansion generates only logarithmic corrections to the leading long distance behaviour for any fixed dimension d , $2 < d \leq 4$. The situation is thus similar to the situation one encounters for the ε -expansion (at the IR fixed point), and one expects to be able to calculate universal quantities like critical exponents as a power series in $1/N$. However, because the interactions are not local, the results of renormalization theory do not immediately apply. We will now construct an alternative quasi-local field theory, for which the standard RG analysis is valid, and that reduces to the large N field theory in some limits.

18.7 The $1/N$ -expansion: An alternative field theory

To be able to use the standard results of renormalization theory, we reformulate the critical theory to deal with non-local interactions. If we neglect corrections to scaling, we can omit the term λ^2 in action (18.19) and use the representation (18.71) of the partition function:

$$\begin{aligned} \mathcal{Z} &= \int [d\lambda(x)] [d\phi(x)] \exp [-\mathcal{S}(\phi, \lambda)], \quad \text{with}, \\ \mathcal{S}(\phi, \lambda) &= \frac{1}{2T} \int d^d x \left[(\nabla \phi(x))^2 + \lambda(x) (\phi^2(x) - 1) \right]. \end{aligned} \quad (18.87)$$

The difficulty arises from the λ -propagator, absent in the perturbative formulation, and generated by the large N summation. We thus add to the action (18.87) a quadratic term in λ , which, in the tree approximation of standard perturbation theory, generates a λ -propagator of the form (18.51). The modified action \mathcal{S}_v reads

$$\begin{aligned} \mathcal{S}_v(\phi, \lambda) &= \frac{1}{2T} \int d^d x \left[(\nabla \phi(x))^2 + \lambda(x) (\phi^2(x) - 1) \right] \\ &\quad - \frac{1}{2v^2} \int d^d x \lambda(x) (-\nabla^2)^{-\varepsilon/2} \lambda(x). \end{aligned} \quad (18.88)$$

In the limit where the parameter v goes to infinity, the coefficient of the additional term vanishes, and the initial action is recovered.

We will now consider only the critical theory. This means that the couplings of all relevant interactions are set to their critical values. These interactions contain a term linear in λ and a polynomial in ϕ^2 of degree depending on the dimension. Note that, in a discrete set of dimensions, some monomials become just renormalizable. Therefore, we work in generic dimensions. The quantities we calculate are regular in the dimension. The field theory with the action (18.88) can be studied with standard field theory methods. The peculiar form of the λ quadratic term, which is not strictly local, does not create a problem. Similar terms are encountered in statistical systems with long range forces, and renormalization theory still applies. The most direct consequence is that the λ -field is not renormalized because *counter-terms are always local*.

It is convenient to rescale $\phi \mapsto \phi\sqrt{T}$, $\lambda \mapsto v\lambda$. The action becomes

$$\mathcal{S}_v(\phi, \lambda) = \frac{1}{2} \int d^d x \left[(\nabla \phi(x))^2 + v\lambda(x)\phi^2(x) - \lambda(x)(-\nabla^2)^{-\varepsilon/2}\lambda(x) + \text{relevant terms} \right].$$

The renormalized critical action then reads

$$[\mathcal{S}_v]_r = \frac{1}{2} \int d^d x \left[Z_\phi (\nabla \phi(x))^2 + v_r Z_v \lambda \phi^2(x) - \lambda(x)(-\nabla^2)^{-\varepsilon/2}\lambda(x) + \text{relevant terms} \right]. \quad (18.89)$$

It follows that the RG equations for vertex functions of l λ fields and n ϕ fields, in the critical theory, take the form

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta_{v^2}(v) \frac{\partial}{\partial v^2} - \frac{n}{2} \eta(v) \right] \tilde{\Gamma}^{(l,n)} = 0. \quad (18.90)$$

The solution to the RG equations (18.90) can be written as

$$\tilde{\Gamma}^{(l,n)}(\ell p, v, \Lambda) = Z^{-n/2}(\ell) \ell^{d-2l-n(d-2)/2} \tilde{\Gamma}^{(l,n)}(p, v(\ell), \Lambda), \quad (18.91)$$

with the usual definitions,

$$\ell \frac{dv^2}{d\ell} = \beta(v(\ell)), \quad \ell \frac{d \ln Z}{d\ell} = \eta(v(\ell)).$$

The RG functions can then be calculated as a power series in $1/N$. It is easy to verify that v^2 has to be taken of order $1/N$. Therefore, to generate a $1/N$ expansion, one first has to sum the multiple insertions of the one-loop λ two-point function, contributions that form a geometric series. The λ propagator then becomes

$$\tilde{\Delta}_\lambda(p) = -\frac{2p^{4-d}}{b(d)R(v)}, \quad (18.92)$$

where $b(d)$ is the constant (18.50), and where we have defined

$$R(v) = 2/b(d) + Nv^2. \quad (18.93)$$

We are interested in the neighbourhood of the fixed point $v^2 = \infty$. One verifies that the RG function $\eta(v)$ approaches the exponent η obtained by direct calculation, and the RG β -function behaves like v^2 . The flow equation for the coupling constant becomes

$$\ell \frac{dv^2}{d\ell} = \rho v^2, \Rightarrow v^2(\ell) \sim \ell^\rho. \quad (18.94)$$

We note that, to each power of the λ field, a power of v corresponds. It follows that

$$\begin{aligned}\tilde{\Gamma}^{(l,n)}(\ell p, v, \Lambda) &\propto v^l(\ell) \ell^{d-2l-n(d-2+\eta)} \\ &\propto \ell^{d-(2-\rho/2)l-n(d-2+\eta)}.\end{aligned}\quad (18.95)$$

To compare with the result (15.71) obtained from the perturbative RG, one has still to take into account that the functions $\Gamma^{(l,n)}$ defined here are obtained by an additional Legendre transformation with respect to the source of ϕ^2 . Therefore,

$$2 - \rho/2 = d_{\phi^2} = d - 1/\nu. \quad (18.96)$$

18.8 Explicit calculations

We explain how $1/N$ corrections to exponents can be calculated, and then report a few higher order results. The equation of state [184] and the two-point function [185] are also known up to order $1/N$.

18.8.1 Critical exponents at order $1/N$

Most calculations at order $1/N$ rely on the evaluation of the generic integral, defined for all μ and ν by an analytic continuation,

$$\frac{1}{(2\pi)^d} \int \frac{d^d q}{(p+q)^{2\mu} q^{2\nu}} = p^{d-2\mu-2\nu} \frac{\Gamma(\mu + \nu - d/2)\Gamma(d/2 - \mu)\Gamma(d/2 - \nu)}{(4\pi)^{d/2} \Gamma(\mu)\Gamma(\nu)\Gamma(d - \mu - \nu)}. \quad (18.97)$$

For later purpose, it is convenient to set

$$X_1 = \frac{2N_d}{b(d)} = \frac{4\Gamma(d-2)}{\Gamma(d/2)\Gamma(2-d/2)\Gamma^2(d/2-1)} = \frac{4\sin(\pi\varepsilon/2)\Gamma(2-\varepsilon)}{\pi\Gamma(1-\varepsilon/2)\Gamma(2-\varepsilon/2)}. \quad (18.98)$$

To compare with fixed dimension results, note that $X_1 \sim 2(4-d)$ for $d \rightarrow 4$, and $X_1 \sim (d-2)$ for $d \rightarrow 2$.

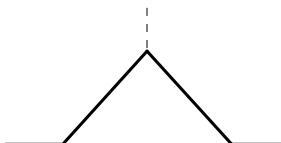


Fig. 18.5 Diagram contributing to $\Gamma_{\sigma\sigma\lambda}^{(3)}$ at order $1/N$ (A_1)

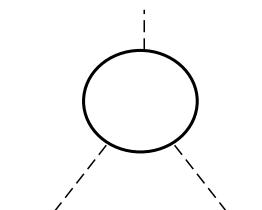


Fig. 18.6 Diagram contributing to $\Gamma_{\sigma\sigma\lambda}^{(3)}$ at order $1/N$ (A_2)

The calculation of the $\langle\phi\phi\rangle$ vertex function at order $1/N$ involves the evaluation of the diagram of Fig. 18.4. To determine the coefficient of $p^2 \ln \Lambda/p$, one can, at one-loop order, replace the λ propagator $q^{-\varepsilon}$ by $q^{2\nu}$ and send the cut-off to infinity. One then uses the result (18.97) with $\mu = 1$. In the limit $2\nu \rightarrow -\varepsilon$ the integral has a pole. The residue of the pole yields the coefficient of $p^2 \ln \Lambda$ and the finite part contains the $p^2 \ln p$ contribution. One finds

$$\tilde{\Gamma}_{\sigma\sigma}^{(2)}(p) = p^2 + \frac{\varepsilon}{4-\varepsilon} \frac{2N_d}{b(d)R(v)} v^2 p^2 \ln(\Lambda/p).$$

Expressing that the function satisfies the RG equation, one obtains the function $\eta(v)$.

The second RG function can be inferred from the divergent parts of the $\langle\phi\phi\lambda\rangle$ vertex function. One finds

$$\tilde{\Gamma}_{\sigma\sigma\lambda}^{(3)} = v + A_1 v^3 R^{-1}(v) \ln \Lambda + A_2 v^5 R^{-2}(v) \ln \Lambda + \text{finite},$$

with

$$A_1 = -\frac{2}{b(d)} N_d = -X_1, \quad A_2 = -\frac{4N}{b^2(d)} (d-3)b(d)N_d = -2N(d-3)X_1,$$

where A_1 and A_2 correspond to the diagrams of Figs. 18.5 and 18.6, respectively.

Applying the RG equation, one finds the relation at order $1/N$,

$$\beta_{v^2}(v) = 2v^2 \eta(v) - 2A_1 v^4 R^{-1}(v) - 2A_2 v^6 R^{-2}(v). \quad (18.99)$$

We thus obtain

$$\eta(v) = \frac{\varepsilon v^2}{4-\varepsilon} X_1 R^{-1}(v), \quad (18.100)$$

$$\beta_{v^2}(v) = \frac{8v^4}{4-\varepsilon} X_1 R^{-1}(v) + 4N(1-\varepsilon)v^6 X_1 R^{-2}(v), \quad (18.101)$$

where the first term in β_{v^2} comes from A_1 and η and the second from A_2 .

From the large v^2 behaviour, one infers the exponents

$$\eta = \frac{\varepsilon}{N(4-\varepsilon)} X_1 + O(1/N^2), \quad (18.102)$$

and

$$\rho = \frac{4(3-\varepsilon)(2-\varepsilon)}{N(4-\varepsilon)} X_1 > 0 \Rightarrow \frac{1}{\nu} = d-2 + \frac{2(3-\varepsilon)(2-\varepsilon)}{N(4-\varepsilon)} X_1 + O(1/N^2). \quad (18.103)$$

The expansion of γ up to order $1/N$ can be obtained from the scaling relation $\gamma = \nu(2-\eta)$:

$$\gamma = \frac{1}{1-\varepsilon/2} \left(1 - \frac{3}{2N} X_1 \right) + O\left(\frac{1}{N^2}\right). \quad (18.104)$$

The exponents ω and $\theta = \omega\nu$, governing the leading corrections to scaling, can also be calculated, for example, from the $\langle\lambda^2\lambda\lambda\rangle$ function. One finds [179]

$$\omega = \varepsilon \left(1 - \frac{2(3-\varepsilon)^2}{(4-\varepsilon)N} X_1 \right) + O\left(\frac{1}{N^2}\right), \quad (18.105)$$

$$\theta = \omega\nu = \frac{\varepsilon}{2-\varepsilon} \left(1 - \frac{2(3-\varepsilon)}{N} X_1 \right) + O\left(\frac{1}{N^2}\right). \quad (18.106)$$

Note that the exponents are regular functions of ε up to $\varepsilon = 2$ and free of renormalon singularities at $\varepsilon = 0$.

18.8.2 Higher order results

The calculations beyond the order $1/N$ are rather technical. The reason is easy to understand; because the effective field theory is renormalizable in all dimensions $2 \leq d \leq 4$, the dimensional regularization, which is so useful in perturbative calculations, no longer works. Therefore, either one keeps a true cut-off, or one introduces more sophisticated regularization schemes. For details, the reader is referred to the literature.

The equation of state and the spin–spin correlation function in zero field are also known at order $1/N$ but, since the expressions are complicated, we again refer the reader to the literature for details.

Generic dimensions. The exponents ν and η are known up to order $1/N^2$ and $1/N^3$, respectively, in arbitrary dimensions, but the expressions are too complicated to be reproduced here [186].

Three-dimensional results. For $d = 3$, The exponent γ is known up to order $1/N^2$ [187, 188]:

$$\gamma = 2 - \frac{24}{N\pi^2} + \frac{64}{N^2\pi^4} \left(\frac{44}{9} - \pi^2 \right) + O\left(\frac{1}{N^3}\right).$$

The expansion of η reduces to [186],

$$\eta = \frac{\eta_1}{N} + \frac{\eta_2}{N^2} + \frac{\eta_3}{N^3} + O\left(\frac{1}{N^4}\right),$$

with

$$\eta_1 = \frac{8}{3\pi^2}, \quad \eta_2 = -\frac{8}{3}\eta_1^2, \quad \eta_3 = \eta_1^3 \left(-\frac{797}{18} - \frac{61}{24}\pi^2 + \frac{27}{8}\psi''(1/2) + \frac{9}{2}\pi^2 \ln 2 \right),$$

$\psi(x)$ being the logarithmic derivative of the Γ function.

Note that the $1/N$ expansion seems to be rapidly divergent and a direct summation of these terms does not provide reliable estimates of critical exponents in three dimensions for directly useful values of N .

The ratio of amplitudes R_χ (equation (16.182)) is also known to order $1/N^2$ [188].

Among other results, the β -function at order $1/N$ in the massive theory renormalized at zero momentum have been reported in Ref. [189], and a calculation of the dimensions of composite operators to order $1/N^2$ in Ref. [190].

19 The non-linear σ -model near two dimensions: Phase structure

In this chapter, we discuss, mostly in dimensions close to 2, the non-linear σ -model [103, 191], a quantum field theory, where the (scalar) field is an N -component vector of fixed length. The model possesses a global, non-linearly realized symmetry $O(N)$ symmetry: under a group transformation, the transformed field is a non-linear function of the field itself. The construction of the non-linear σ -model is based on a special class homogeneous spaces, the symmetric spaces that, as Riemannian manifolds, admit a unique metric.

A simple analysis reveals that, by contrast with a $(\phi^2)^2$ -like field theory with the same symmetry, in the non-linear σ -model, in the tree approximation, the $O(N)$ symmetry is always spontaneously broken: the action describes the interactions of $(N - 1)$ massless fields, the Goldstone modes.

Note that, since the fields are massless, in two dimensions infrared (IR) divergences appear in the perturbative expansion, and an IR regulator is required.

To understand the phase structure beyond leading order, a renormalization group (RG) analysis is necessary. This requires understanding how the model renormalizes. Power counting shows that the model is renormalizable in two dimensions. Then, the field is dimensionless and we face a problem already mentioned in Section 8.3: although the degree of divergence of Feynman diagrams is bounded, an infinite number of counter-terms is generated, because all correlation functions are divergent. However, we prove that, due to the special geometric properties of the model, the coefficients of all counter-terms can be calculated as a function of two of them so that the renormalized theory depends only on a finite number of parameters. The proof is based on a set of Ward–Takahashi identities (WT) satisfied by correlation functions, which is summarized in the form of a quadratic equation satisfied by the generating functional of vertex functions (like in non-Abelian gauge theories, see Chapter 26).

In Section 19.8, we discuss the renormalization of composite operators. In Section 19.9, we indicate how all results can be recovered from a representation of the model where the condition that the field is a vector of fixed length is enforced by a field integration.

In Chapters 14–16, we have argued that universal properties, in the critical domain, of $O(N)$ symmetric lattice spin models can be derived from an effective $(\phi^2)^2$ Euclidean field theory. We have determined a number of critical properties near the upper-critical dimension $d = 4$, using RG methods, within the framework of the $(4 - d)$ expansion. Some general results obtained in this way, whose validity can only be established analytically for $(4 - d)$ infinitesimal, have been confirmed in Chapter 18, for all dimensions $2 < d < 4$, in the large N limit and, more generally, within the framework of the $1/N$ expansion.

One outcome of the large N analysis is the discovery of a remarkable relation between the $(\phi^2)^2$ field theory and the non-linear σ -model in the continuum limit (at least for generic $(\phi^2)^2$ couplings), although from the viewpoint of perturbation theory these models are quite different (see Section 18.6). This property, mysterious at the classical level, emphasizes the essential nature of quantum (or statistical) fluctuations. Large N calculations also predict the behaviour of the critical domain when the dimension approaches 2 continuously: the critical temperature vanishes continuously, and the resulting critical domain is confined to a neighbourhood of zero temperature.

This strongly suggests that the non-linear σ -model, whose perturbative expansion is a low temperature expansion, as a lattice regularization indicates (see Section 19.4.2), is the proper tool to study universal properties of critical phenomena in the neighbourhood of dimension 2, the dimension in which the model is renormalizable.

The special role of two dimensions, the lower-critical dimension, in models with continuous symmetries, has already been pointed out in Chapter 14. In contrast to models with discrete symmetries, due to massless Goldstone modes, in these models, correlation functions decay algebraically for any temperature below T_c , IR instabilities prevent a phase transition in two dimensions. Considerations based on the non-linear σ -model are thus relevant for the N -vector model only for $N \geq 2$. The main analytic tools are RG equations and $\varepsilon = d - 2$ expansion.

In two dimensions, the σ -model also provides an important example of a model having the property of ultraviolet (UV) asymptotic freedom.

IR instability of the massless perturbative phase and asymptotic freedom generalizes to other two-dimensional models: models based on symmetric spaces (Chapter 29) and various fermion self-interacting models with chiral symmetry (Chapter 20).

19.1 The non-linear σ -model: Definition

The field manifold. We consider an N -component (scalar) field $\phi(x)$, which satisfies the $O(N)$ invariant constraint,

$$\phi^2(x) = 1. \quad (19.1)$$

The field ϕ belongs to the sphere S_{N-1} , which can also be identified with the homogeneous (symmetric) space $O(N)/O(N-1)$: indeed, let $\mathbf{g}(x)$ be an $N \times N$ matrix, element of the $O(N)$ group, depending on the coordinate x , and \mathbf{u} the fixed N -component vector of unit length, $\mathbf{u} = (1, 0, \dots, 0)$. The field ϕ can be parametrized as

$$\phi(x) = \mathbf{g}(x)\mathbf{u}, \quad (19.2)$$

since, by definition, an orthogonal transformation leaves the length of a vector invariant.

The little group of \mathbf{u} (or stabilizer), is the subgroup of $O(N)$ that leaves \mathbf{u} invariant. It is isomorphic to $O(N-1)$. If one multiplies $\mathbf{g}(x)$ on the right by any element of the little group of \mathbf{u} , the right-hand side of equation (19.2) is left unchanged. The relation (19.2) thus exhibits an isomorphism between the coset (homogeneous) space $O(N)/O(N-1)$ and the sphere S_{N-1} .

The partition function. The quantization of the model is not straightforward, due to operator-ordering problems [192, 193]. Formally, the partition function can be written as (see also Section 18.6.2)

$$\mathcal{Z} = \int [d\phi(x)] \prod_x \delta(\phi^2(x) - 1) \exp[-\mathcal{S}(\phi)/g], \quad (19.3)$$

where $g > 0$ is the coupling constant and

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x (\nabla \phi(x))^2. \quad (19.4)$$

The action (19.4) is the most general $O(N)$ symmetric action containing at most two derivatives, up to a multiplicative constant (in a theory invariant under space translations). Indeed, due to the constraint (19.1), any symmetric derivative-free term reduces to a constant (and $\phi \cdot \nabla \phi$ vanishes).

In the field integral, the expression of the $O(N)$ invariant integration measure is not really defined, and is purely formal (see also Sections 3.3, 3.5). It can be defined by using, for example, a lattice regularization.

Parametrization and non-linear group representation. To generate a perturbative expansion, we need a parametrization of the field ϕ in terms of independent variables. A convenient parametrization of the sphere (19.1) is

$$\phi(x) = \{\sigma(x), \boldsymbol{\pi}(x)\}, \quad (19.5)$$

in which $\boldsymbol{\pi}(x)$ is an $(N-1)$ -component field, and the field $\sigma(x)$ a function of $\boldsymbol{\pi}(x)$ through equation (19.1). The equation can be solved locally; for example, if $\sigma(x)$ is positive,

$$\sigma(x) = (1 - \boldsymbol{\pi}^2(x))^{1/2}. \quad (19.6)$$

The consequences of the singularity of this parametrization will be discussed later.

We decompose the set of generators of the Lie algebra of $O(N)$ into the set of generators of the Lie algebra of the stabilizer group $O(N-1)$, and the complementary set. The group $O(N-1)$ acts linearly on $\boldsymbol{\pi}(x)$. To the complementary set correspond infinitesimal transformations of the form

$$\delta_\omega \boldsymbol{\pi}_i = \omega_i (1 - \boldsymbol{\pi}^2(x))^{1/2}, \quad (19.7)$$

in which ω_i are constants, infinitesimal parameters of the transformation. The transformation of the σ -field is then a consequence of the transformation (19.7) of the $\boldsymbol{\pi}$ -field:

$$\delta_\omega \sigma(x) \equiv \delta_\omega (1 - \boldsymbol{\pi}^2(x))^{1/2} = -\boldsymbol{\omega} \cdot \boldsymbol{\pi}(x). \quad (19.8)$$

Action and partition function. In terms of the field $\boldsymbol{\pi}(x)$, the action (19.4) can be cast into another geometric form,

$$\mathcal{S}(\boldsymbol{\pi}) = \frac{1}{2} \int d^d x \sum_{i,j} G_{ij}(\boldsymbol{\pi}(x)) \nabla \pi_i(x) \cdot \nabla \pi_j(x), \quad (19.9)$$

in which $G_{ij}(\boldsymbol{\pi})$ is a metric tensor on the sphere:

$$G_{ij} = \delta_{ij} + \frac{\pi_i \pi_j}{1 - \boldsymbol{\pi}^2}. \quad (19.10)$$

In the form (19.9), the action is covariant under a reparametrization of the sphere.

In Section 3.3, we have shown that the quantization of actions of the form (19.9) introduces an additional undefined determinant. From equations (3.26) and (3.27), we infer the representation of the generating functional of $\boldsymbol{\pi}$ -field correlation functions,

$$\mathcal{Z}(\mathbf{J}) = \int \left[\frac{d\boldsymbol{\pi}(x)}{(1 - \boldsymbol{\pi}^2(x))^{1/2}} \right] \exp \left[-\frac{1}{g} \left(\mathcal{S}(\boldsymbol{\pi}) - \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\pi}(x) \right) \right], \quad (19.11)$$

a form consistent with expression (19.3).

From the viewpoint of Riemannian manifolds, a non-trivial measure is needed for geometric reasons (Section 28.3.1): here indeed, we find the $O(N)$ -invariant functional measure $d\boldsymbol{\pi}/\sqrt{1 - \boldsymbol{\pi}^2}$ on the sphere (see Section 3.5).

However, this formal interpretation of the determinant does not eliminate the difficulty. We have to find a method to deal with this undefined contribution to the action:

$$\prod_x (1 - \boldsymbol{\pi}^2(x))^{-1/2} \sim \exp \left[-\frac{1}{2} \delta^d(0)' \int d^d x \ln (1 - \boldsymbol{\pi}^2(x)) \right]. \quad (19.12)$$

As we have explained in Section 3.3, this difficulty is directly related to the problem of operator ordering that appears in the quantization. Indeed, denoting by $\mathbf{q}(x)$ the conjugate momentum of $\boldsymbol{\pi}(x)$, the commutation relation reads

$$[\pi_i(x), q_j(y)] = \hbar \delta^{(d-1)}(x-y) \Rightarrow [\pi_i(x), q_j(x)] = \hbar \delta^{(d-1)}(0)'.$$

19.2 Perturbation theory. Power counting

For $g \rightarrow 0$ and in zero source ($J = 0$), the field integral (19.3) is dominated by saddle points, minima of the classical action (19.4):

$$|\nabla\phi(x)| = 0 \Rightarrow \phi(x) = \phi_0 ,$$

where ϕ_0 is an arbitrary constant unit vector. The action has a continuous set of degenerate and equivalent minima, which are related by $O(N)$ transformations. Each minimum is the starting point of a perturbative expansion. As we have pointed out in Section 13.4, the choice between summing over the contributions of all minima or selecting one particular minimum only, depends on the true physical situation, beyond perturbation theory. In this chapter, we rely on simple perturbative arguments. A more complete analysis can be found in Chapters 14–17, in which the theory of phase transitions is discussed. Here, we first examine the contribution of one saddle point. In the parametrization (19.5), we choose the minimum $\boldsymbol{\pi}(x) = 0$ or $\phi(x) = \mathbf{u}$.

19.2.1 Formal perturbation theory

In the field integral (19.11), the parameter g , from the point of view of classical statistical physics plays the role of the temperature, and from the point of view of quantum physics the role of \hbar . Therefore, an expansion in powers of g is a loop expansion. For g small, the fields $\boldsymbol{\pi}(x)$ that contribute to the field integral are such that

$$|\nabla\boldsymbol{\pi}(x)| \sim \sqrt{g}$$

and, since we expand around $\boldsymbol{\pi}(x) = 0$, the field itself must satisfy

$$|\boldsymbol{\pi}(x)| \sim \sqrt{g} . \quad (19.13)$$

Values of $\boldsymbol{\pi}(x)$ of order 1 give exponentially small contributions to the field integral (of order $\exp(-\text{const.}/g)$), which are negligible at any finite order of perturbation theory.

This has two consequences: the restrictions imposed by the parametrization (19.6) ($\sigma(x) > 0$) are irrelevant in perturbation theory and, in addition, in the field integral, we can freely integrate over $\boldsymbol{\pi}(x)$ from $+\infty$ to $-\infty$, disregarding the constraint $|\boldsymbol{\pi}(x)| \leq 1$.

Perturbation theory, then, again relies on the evaluation of simple Gaussian integrals.

We first discuss formal perturbation theory, setting temporarily aside the question of UV or IR divergences. We rewrite the field integral (19.11) as

$$\mathcal{Z}(\mathbf{J}) = \int [d\boldsymbol{\pi}] \exp \left[- \int d^d x \mathcal{L}(\boldsymbol{\pi}, \mathbf{J})/g \right] , \quad (19.14)$$

where the Lagrangian density is given by

$$\mathcal{L}(\boldsymbol{\pi}, \mathbf{J}) = \frac{1}{2} \left[(\nabla\boldsymbol{\pi}(x))^2 + (\nabla\sigma(x))^2 \right] + \frac{1}{2} g \cdot \delta^d(0)' \ln(1 - \boldsymbol{\pi}^2(x)) - \mathbf{J}(x) \cdot \boldsymbol{\pi}(x) ,$$

with

$$(\nabla\sigma(x))^2 = \frac{(\boldsymbol{\pi}(x) \cdot \nabla\boldsymbol{\pi}(x))^2}{1 - \boldsymbol{\pi}^2(x)} .$$

The measure term has no $1/g$ factor, and only starts contributing at one-loop order.

Since $\boldsymbol{\pi}$ is of order \sqrt{g} , it may be convenient to rescale the field, $\boldsymbol{\pi} \mapsto \boldsymbol{\pi} \sqrt{g}$. After this rescaling, the action density becomes

$$\begin{aligned}\mathcal{L}(\boldsymbol{\pi}, \mathbf{J})/g &= \frac{1}{2} (\nabla \boldsymbol{\pi}(x))^2 + \frac{g}{2} \frac{(\boldsymbol{\pi}(x) \cdot \nabla \boldsymbol{\pi}(x))^2}{1 - g\boldsymbol{\pi}^2(x)} + \frac{1}{2} \delta^d(0)' \ln(1 - g\boldsymbol{\pi}^2(x)) \\ &\quad - \frac{1}{\sqrt{g}} \mathbf{J}(x) \cdot \boldsymbol{\pi}(x).\end{aligned}\tag{19.15}$$

Expression (19.15) shows that the interaction term in the action, once expanded in powers of g , generates an infinite number of different vertices with arbitrary even powers of $\boldsymbol{\pi}$ and two derivatives. Still, it is possible to verify that, at any finite order in perturbation theory and for a given correlation function, only a finite number of vertices contributes. Formally, the measure term yields additional vertices without derivatives.

The propagator $\tilde{\Delta}_{ij}(p)$ of the $\boldsymbol{\pi}$ -field, in the Fourier representation, is

$$\tilde{\Delta}_{ij}(p) = \frac{\delta_{ij}}{p^2}.\tag{19.16}$$

In the tree approximation, the $\boldsymbol{\pi}$ -field is massless. Returning to the analysis of Section 13.4.2, we conclude that, at leading order in perturbation theory, the non-linear σ -model automatically realizes the $O(N)$ symmetry in the phase of spontaneous symmetry breaking, the $\boldsymbol{\pi}$ -field corresponding to the Goldstone modes. The massive partner of the $\boldsymbol{\pi}$ -field in the linear realization, the σ component, has been eliminated by the constraint (19.1). This constraint formally freezes the fluctuations of $\phi^2(x)$ and sends, in the classical limit, the σ mass to infinity.

Note that these specific properties are independent of the special choice (19.5) of parametrization of $\phi(x)$.

19.2.2 Power counting and renormalization

The general analysis is given in Chapter 8. The form of the propagator shows that the dimension $[\boldsymbol{\pi}]$ of the $\boldsymbol{\pi}$ -field is

$$[\boldsymbol{\pi}] = \frac{1}{2} (d - 2).\tag{19.17}$$

Therefore, the dimension of a vertex containing $2n$ $\boldsymbol{\pi}$ -fields is

$$[\partial^2 \boldsymbol{\pi}^{2n}] = n(d - 2) + 2.\tag{19.18}$$

As a consequence

- (i) For $d = 2$, the theory is just renormalizable;
- (ii) for $d > 2$, the theory is not renormalizable.

Therefore, we first study the model in dimension 2. We have already mentioned a peculiarity of this case: although the theory is renormalizable by power counting, any local monomial in the field containing at most two derivatives and an arbitrary power of $\boldsymbol{\pi}$ can *a priori* appear as a counter-term. The symmetry $O(N - 1)$, which is linearly realized, only restricts the counter-terms to be of the general form

$$(\nabla \boldsymbol{\pi} \cdot \boldsymbol{\pi})^2 (\boldsymbol{\pi}^2)^n, \quad (\nabla \boldsymbol{\pi})^2 (\boldsymbol{\pi}^2)^n, \quad (\boldsymbol{\pi}^2)^n.$$

However, the non-linear $O(N)$ symmetry implies that, up to a normalization factor, the unrenormalized action is unique. To understand the structure of the theory after renormalization, we have to investigate the implications of the non-linear $O(N)$ symmetry on the form of the divergences in perturbation theory. We have first to exhibit a regularization scheme that preserves the $O(N)$ symmetry, and then derive a set of WT identities which express the consequences of the symmetry for correlation functions.

19.3 IR divergences

Since, in a massless scalar field theory, the propagator behaves like $1/p^2$, perturbation theory has low momentum (IR) divergences in dimension 2, the dimension in which the non-linear σ -model is renormalizable (see Section 19.11 for explicit calculations). This divergence is directly related to the impossibility of spontaneous breaking of continuous symmetries in local scalar field theories in two dimensions.

To generate a finite perturbation theory, it is necessary to introduce an IR cut-off, for example, by giving a mass to the π -field. Since the absence of mass is a consequence of the spontaneous breaking of the $O(N)$ symmetry in the classical limit, a mass breaks the symmetry explicitly. The study of symmetry breaking mechanisms in Chapter 13 suggests that a convenient method to give a mass consists in adding to the action (19.9) a constant source h for the σ -field (a magnetic field in a model of classical spins):

$$\mathcal{S}(\boldsymbol{\pi}) \longmapsto \mathcal{S}(\boldsymbol{\pi}) - h \int \sigma(x) d^2x, \quad h > 0. \quad (19.19)$$

After this modification, the minimum of the action is no longer degenerate. A maximization of the source term yields a minimum at $\boldsymbol{\pi} = 0$.

If we then expand σ in powers of $\boldsymbol{\pi}$,

$$\sigma = (1 - \boldsymbol{\pi}^2)^{1/2} = 1 - \frac{1}{2}\boldsymbol{\pi}^2 + O((\boldsymbol{\pi}^2)^2), \quad (19.20)$$

and collect the quadratic terms in the action, we find the new $\boldsymbol{\pi}$ -field propagator

$$\tilde{\Delta}_{ij}(p) = \delta_{ij} \frac{g}{p^2 + h}. \quad (19.21)$$

The linear σ term has thus generated a mass $h^{1/2}$ for the $\boldsymbol{\pi}$ -field, together with new derivative-free interactions.

We recall that, in the case of linearly realized symmetries, the breaking term $h\sigma$ is linear in an independent field and, therefore, as we have shown in Chapter 13, generates no new renormalization.

Finite volume. In Chapter 32, we discuss another IR regularization scheme, which does not break the $O(N)$ symmetry. It is based on the property that a symmetry cannot be spontaneously broken in a finite volume: therefore, no Goldstone mode is generated. Technically in a hypercube of linear size L , the momenta after Fourier transformation are quantized. This solves the IR problem, because integrals are replaced by discrete sums. In the special case of periodic boundary conditions, the momenta have the form $\mathbf{p} = (2\pi/L)\mathbf{n}$, with $\mathbf{n} \in \mathbb{Z}^2$. In this case, some care is required to properly handle the zero-momentum mode $\mathbf{p} = 0$, which seems to still lead to divergences. However, it has to be eliminated in favour of an integral over a constant unit vector which represents the sum over all degenerate minima. The other momenta do not lead to IR divergences and can be treated perturbatively.

IR finiteness of $O(N)$ invariant correlation functions in two dimensions. We mention here briefly, without proof, an interesting result whose significance is discussed later with the physics (in the sense of statistical physics) of the non-linear σ -model. We have explained that, in two dimensions, correlation functions are IR divergent, as one verifies on the explicit expression (19.95), an indication that the $O(N)$ symmetry is not broken.

This has required the introduction of an IR cut-off to give a mass to the π -field. Moreover, this additional term, breaking the $O(N)$ symmetry, lifts the degeneracy of the classical minimum of the action and, therefore, eliminates a potential difficulty with perturbation theory: is it necessary to take into account all degenerate minima of the action or can one choose one of them? However, as we point out in Section 19.2, this question is irrelevant for $O(N)$ invariant correlation functions which, therefore, play a special role. Actually, it has been conjectured by Elitzur and proved by David [194], that, order by order in perturbation theory, $O(N)$ invariant correlation functions have, in two dimensions, a finite IR limit, that is, for example, a limit when, in the notation (19.19), the breaking parameter h goes to zero, or any other IR cut-off is removed.

19.4 UV regularization

A regularization of the non-linear σ -model has to preserve the $O(N)$ symmetry. This is less trivial than in the linear case since, as a consequence of the symmetry, the interaction terms in the action (19.9) are related to the quadratic part. Since dimensional regularization is a possible regularization, we *generalize the model to space dimension d* .

19.4.1 Perturbative regularizations

Momentum cut-off regularization. A natural idea is to use the regularization (13.13) that works for linear symmetries. A simple way to implement the regularization, is to start from the description of the model in terms of the ϕ -field, because the action (19.4) is formally a free field action. The action $\mathcal{S}(\phi)$ is replaced by [195]

$$\mathcal{S}_\Lambda(\phi) = \frac{1}{2} \int d^d x \nabla \phi(x) \cdot \prod_{r=1}^{r_{\max}} (1 - \nabla^2 / M_r^2) \nabla \phi(x), \quad (19.22)$$

in which the masses M_r are proportional to the cut-off Λ . Expressing, then, $\phi(x)$ in terms of $\pi(x)$, we discover that the large momentum behaviour of the propagator has improved, but simultaneously new, more singular interactions have been generated. If the propagator behaves like

$$\Delta(p) \propto 1/p^s,$$

then the most singular interaction has s derivatives. Using equation (8.10), which gives the superficial degree of divergence $\delta(\gamma)$ of a diagram γ , and for $d = 2$ becomes

$$\delta(\gamma) = 2L - sI + \sum_\alpha k_\alpha v_\alpha,$$

and eliminating the number of internal lines I through the topological relation (8.11),

$$L = I - \sum_\alpha v_\alpha + 1,$$

one finds

$$\delta(\gamma) = (2 - s)L + s + \sum_\alpha (k_\alpha - s)v_\alpha, \quad (19.23)$$

where L is the number of loops, and k_α the number of derivatives at vertex α .

The worst case is $k_\alpha = s$:

$$\delta(\gamma) \leq (2 - s)L + s.$$

As a consequence, for $L \geq 2$, it is sufficient to take $s \geq 6$ to regularize all diagrams. However, the one-loop diagrams have a behaviour independent of s , and thus cannot be regularized. This property is not completely independent of the other limitation of the momentum cut-off (or Pauli–Villars’s) regularization, that it cannot regularize the undefined measure term. Indeed, we show later that the one-loop divergences generated by the interaction cancel the divergences coming from the measure. These difficulties are related to the problem of quantization and the order of non-commuting operators in local products.

Momentum regularization is mainly useful in the study of non-linear models coupled to chiral fermions.

Dimensional regularization. Dimensional regularization, based on generalizing the model to d space dimensions, with d continuous, preserves the $O(N)$ symmetry of the action. Furthermore, as a consequence of the formal rule

$$\int d^d k = (2\pi)^d \delta^d(0) = 0,$$

which performs a partial renormalization, the measure term does not contribute, as well as the power-law momentum divergences of perturbation theory. The symmetry is maintained only if both type of divergences exactly cancel. This has to be demonstrated in a regularization scheme that deals both with the measure and all other divergences. For this purpose, we use lattice regularization. Once renormalizability is demonstrated, owing to its technical simplicity, one can use dimensional regularization for Feynman diagram calculations.

19.4.2 Lattice regularization and statistical physics

To construct a lattice regularized version of the non-linear σ -model, which preserves the $O(N)$ symmetry, it is again convenient to start from the description of the model in terms of the ϕ -field. We then replace derivatives by finite differences, that is, in the notation of Section 8.7,

$$\frac{\partial \phi(x)}{\partial x_\mu} \mapsto \nabla_\mu^{\text{lat.}} \phi(x) = [\phi(x + an_\mu) - \phi(x)]/a,$$

in which x now belongs to a hypercubic lattice of spacing a , and n_μ is the unit vector in the μ direction.

Finally, to implement condition (19.1), we integrate over $\phi(x)$ with the invariant measure on the sphere. The regularized field integral has the form

$$\mathcal{Z}(\mathbf{J}) = \int \prod_{x \in (a\mathbb{Z})^d} \delta(\phi^2(x) - 1) d\phi(x) \exp[-\mathcal{S}(\phi, \mathbf{J})/g], \quad (19.24)$$

with

$$\mathcal{S}(\phi, \mathbf{J}) = \frac{1}{2} \sum_x [\nabla_\mu^{\text{lat.}} \phi(x)]^2 - \sum_x \mathbf{J}(x) \cdot \phi(x). \quad (19.25)$$

Using the parametrization (19.5), we can express the lattice field $\phi(x)$ in terms of $\boldsymbol{\pi}(x)$. This yields a regularized form of the field integral (19.11). In particular, the measure term now generates well-defined interactions:

$$\frac{1}{2} \delta^d(0) \int d^d x \ln(1 - \boldsymbol{\pi}^2(x)) \mapsto \frac{1}{2} \sum_x \ln(1 - \boldsymbol{\pi}^2(x)).$$

In the expression (19.24), we recognize the partition function of a classical spin lattice model with nearest-neighbour (n.n.) ferromagnetic interactions and in the presence of an external magnetic field $\mathbf{J}(x)$. The coupling constant g plays the role of the temperature. The critical properties (in the sense of phase transitions) of the model are discussed in Chapters 14–17 and 31. Expression (19.24) not only provides a regularization of perturbation theory, but also provides various non-perturbative methods to study the non-linear σ -model. Moreover, it is the only regularization which makes a discussion of the role of the measure in perturbation theory possible (see Sections 19.5.1 and 19.6).

19.5 WT identities and master equation

Having regularized the theory both at short and large distances (large and small momenta), we now derive a set of WT identities expressing the consequence of the $O(N)$ symmetry for correlation functions. We discuss only the part of the $O(N)$ symmetry which acts non-linearly, the consequences of the linear $O(N - 1)$ symmetry having already been discussed in Chapter 13.

19.5.1 WT identities

Notation. For notational convenience, we will use a continuum notation, but a lattice regularization is implied.

We consider the infinitesimal transformation (19.7),

$$\delta_\omega \pi_i(x) = \omega_i \sigma(x) \equiv \omega_i \sqrt{1 - \boldsymbol{\pi}^2(x)}. \quad (19.26)$$

The $O(N)$ symmetric part of the action and the measure are left invariant by such a transformation. Only the source term and the breaking term are affected:

$$\delta_\omega \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\pi}(x) = \int d^d x \boldsymbol{\omega} \cdot \mathbf{J}(x) \sigma(x), \quad \delta_\omega \int d^d x \sigma(x) = - \int d^d x \boldsymbol{\omega} \cdot \boldsymbol{\pi}(x).$$

Both the symmetry breaking term (19.19) and the variation of $\mathbf{J}(x)$ involve the composite operator $\sigma(x) = (1 - \boldsymbol{\pi}^2(x))^{1/2}$. Following the general strategy already explained in the case of quadratic symmetry breaking terms in Section 13.8, we introduce a source $H(x)$ for this operator. Then, since the variation of this new operator, and the variation of the breaking term, under a transformation (19.26) are both proportional to $\boldsymbol{\pi}$ itself, no additional operator is needed. We note that this would not have been the case with a breaking term proportional, for example, to $\int \boldsymbol{\pi}^2(x) d^d x$. For convenience, we introduce the notation

$$[d\rho(\boldsymbol{\pi})] = \prod_x \left[\frac{d\boldsymbol{\pi}(x)}{\sqrt{1 - \boldsymbol{\pi}^2(x)}} \right]. \quad (19.27)$$

To generate the perturbative expansion, we also expand in powers of $H(x) - h$. Thus, we consider the generating functional

$$\mathcal{Z}(\mathbf{J}, H) = \int [d\rho(\boldsymbol{\pi})] \exp [-\mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J})/g], \quad (19.28)$$

with the notation,

$$\mathcal{S}(\boldsymbol{\pi}, H) = \mathcal{S}(\boldsymbol{\pi}) - H(x)\sigma(x), \quad \mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J}) = \mathcal{S}(\boldsymbol{\pi}, H) - \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\pi}(x). \quad (19.29)$$

The generating functional of $\boldsymbol{\pi}$ -field correlation functions with the linear $h\sigma$ breaking term is given by

$$\mathcal{Z}(\mathbf{J}) = \mathcal{Z}(\mathbf{J}, H)|_{H(x)=h}.$$

Performing the infinitesimal transformation (19.26), we obtain the equation

$$0 = \int [d\rho(\boldsymbol{\pi})] \int d^d x [\mathbf{J}(x) \cdot \delta\boldsymbol{\pi}(x) + H(x)\delta\sigma(x)] \exp [-\mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J})/g],$$

or, explicitly,

$$0 = \int [d\rho(\boldsymbol{\pi})] \int d^d x [\sigma(x)\mathbf{J}(x) - H(x)\boldsymbol{\pi}(x)] \exp [-\mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J})/g]. \quad (19.30)$$

We now replace $\boldsymbol{\pi}(x)$ by $g(\delta/\delta\mathbf{J}(x))$ and $\sigma(x)$ by $g(\delta/\delta H(x))$. Equation (19.30) then becomes

$$\int d^d x \left(\mathbf{J}(x) \frac{\delta}{\delta H(x)} - H(x) \frac{\delta}{\delta \mathbf{J}(x)} \right) \mathcal{Z}(\mathbf{J}, H) = 0. \quad (19.31)$$

This first order linear differential equation makes no reference to the non-linear character of the transformation (19.26). It is identical to the equation one obtains in the case of a linear $O(N)$ symmetry, when $\mathbf{J}(x)$ and $H(x)$ are the sources for the independent fields $\boldsymbol{\pi}(x)$ and $\sigma(x)$.

It is clear that $\mathcal{W}(\mathbf{J}, H) = g \ln \mathcal{Z}(\mathbf{J}, H)$ satisfies the same equation:

$$\int d^d x \left(\mathbf{J}(x) \frac{\delta}{\delta H(x)} - H(x) \frac{\delta}{\delta \mathbf{J}(x)} \right) \mathcal{W}(\mathbf{J}, H) = 0. \quad (19.32)$$

To derive the WT identities of the generating functional Γ of vertex (1PI) functions, we perform a Legendre transformation. However, in contrast with the linear case, $\sigma(x)$ is a function of the $\boldsymbol{\pi}(x)$ field here and, therefore, the Legendre transformation applies only to the source $\mathbf{J}(x)$, and not to $H(x)$:

$$\mathcal{W}(\mathbf{J}, H) + \Gamma(\boldsymbol{\pi}, H) = \int d^d x \mathbf{J}(x) \cdot \boldsymbol{\pi}(x), \quad \boldsymbol{\pi}(x) = \frac{\delta \mathcal{W}}{\delta \mathbf{J}(x)}. \quad (19.33)$$

We again use identity (7.65), since $H(x)$ is an external parameter:

$$\frac{\delta \mathcal{W}}{\delta H(x)} \Big|_{\mathbf{J}} = - \frac{\delta \Gamma}{\delta H(x)} \Big|_{\boldsymbol{\pi}}. \quad (19.34)$$

Therefore, the Legendre transform of equation (19.32) is [196]

$$\int d^d x \left(\frac{\delta \Gamma}{\delta \boldsymbol{\pi}(x)} \frac{\delta \Gamma}{\delta H(x)} + H(x)\boldsymbol{\pi}(x) \right) = 0. \quad (19.35)$$

This is the basic equation from which the general form of the counter-terms which render the theory finite can be derived.

Remark. Equation (19.35) is quadratic in the vertex functional Γ . This is an essential difference with the case of linearly realized symmetries. This property is shared with non-Abelian gauge theories (see Section 26.8). Actually, one can show that if one uses the strategy that we just explained, that is, adding sources for all new composite operators generated by the group transformation, then the WT identities satisfied by Γ are at most quadratic.

19.5.2 Master equation

It is easy to verify that the initial action satisfies equation (19.35), either directly, or by performing a loop expansion of equation (19.35), and noting that

$$\Gamma(\boldsymbol{\pi}, H) = \mathcal{S}(\boldsymbol{\pi}, H) + O(g).$$

Conversely, we now assume that the action $\mathcal{S}(\boldsymbol{\pi}, H)$ satisfies the *master equation* [196]

$$\int d^d x \left(\frac{\delta \mathcal{S}}{\delta H(x)} \frac{\delta \mathcal{S}}{\delta \boldsymbol{\pi}(x)} + H(x) \boldsymbol{\pi}(x) \right) = 0. \quad (19.36)$$

We perform an infinitesimal change of variables in the field integral (19.28) of the form

$$\boldsymbol{\pi}(x) = \boldsymbol{\pi}'(x) + \delta_\omega \boldsymbol{\pi}(x), \quad \text{with} \quad \delta_\omega \boldsymbol{\pi}(x) = \frac{\delta \mathcal{S}(\boldsymbol{\pi}', H)}{\delta H(x)} \boldsymbol{\omega}. \quad (19.37)$$

The variations of the action and the source term are (we omit the primes on the dummy variable $\boldsymbol{\pi}$),

$$\delta \left[\exp \left(-\frac{1}{g} \mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J}) \right) \right] = \frac{\boldsymbol{\omega}}{g} \cdot \left[H(x) \frac{\delta}{\delta \mathbf{J}(x)} - \mathbf{J}(x) \frac{\delta}{\delta H(x)} \right] \exp \left(-\frac{1}{g} \mathcal{S}(\boldsymbol{\pi}, H, \mathbf{J}) \right).$$

The invariant measure. We show now that the invariant measure for the transformations (19.37) is

$$\prod_x d\boldsymbol{\pi}(x) \left(\frac{\delta \mathcal{S}}{\delta H(x)} \right)^{-1}. \quad (19.38)$$

First, the change of variables (19.37) generates the Jacobian

$$\mathcal{J} = 1 + \omega_i \int d^d x \frac{\delta^2 \mathcal{S}}{\delta \boldsymbol{\pi}_i(x) \delta H(x)}.$$

Correspondingly, the variation of the measure term is

$$\left(\frac{\delta \mathcal{S}}{\delta H(x)} \right)^{-1} \mapsto \left(\frac{\delta \mathcal{S}}{\delta H(x)} \right)^{-1} \left\{ 1 - \left(\frac{\delta \mathcal{S}}{\delta H(x)} \right)^{-1} \frac{\delta^2 \mathcal{S}}{\delta \boldsymbol{\pi}_i(x) \delta H(x)} \delta \boldsymbol{\pi}_i(x) \right\},$$

which, using the explicit form (19.37), exactly cancels the Jacobian. Finally, note that the initial measure $[d\boldsymbol{\pi}(1 - \boldsymbol{\pi}^2)^{-1/2}]$ has the form (19.38).

Therefore, the master equation (19.36) alone, implies equation (19.31) for $\mathcal{Z}(\mathbf{J}, H)$ and, thus, equation (19.35) for $\Gamma(\boldsymbol{\pi}, H)$.

Finally, in *dimensional regularization* the measure term and the Jacobian vanish identically, and equation (19.36) is also implied.

19.6 Renormalization

We summarize the successive steps of the proof of the renormalizability of the model.

First, we prove the stability of equation (19.36) under renormalization: this means that if the classical action satisfies equation (19.36), then it is possible to renormalize the theory in such a way that the renormalized action still satisfies equation (19.36). We then determine the general solution of equation (19.36). It is important to realize that the equation does not explicitly refer to the transformation (19.26). This explains why the explicit form of the transformation law can be modified by the renormalization, although the geometric structure does not change. Indeed, solving equation (19.36) with the constraints coming from power counting, one finds that only two renormalization constants are needed, a coupling constant and a field renormalization. After renormalization, the model is still $O(N)$ invariant for $h = 0$, but the field ϕ now belongs to a sphere of renormalized radius:

$$\phi^2(x) = \pi^2(x) + \sigma^2(x) = 1/Z .$$

Linearized master equation. We assume that the theory has been regularized. We make a loop expansion, that is, as the explicit form of the action shows, an expansion in powers of g of the generating functional $\Gamma(\pi, H)$ of vertex functions,

$$\Gamma(\pi, H) = \sum_{n=0}^{\infty} \Gamma_n(\pi, H) g^n .$$

We insert the expansion into equation (19.35). The functional Γ_0 is simply the initial action and satisfies by itself equation (19.35). The one-loop functional Γ_1 satisfies

$$\int d^d x \left(\frac{\delta \Gamma_0}{\delta \pi(x)} \frac{\delta \Gamma_1}{\delta H(x)} + \frac{\delta \Gamma_0}{\delta H(x)} \frac{\delta \Gamma_1}{\delta \pi(x)} \right) = 0 . \quad (19.39)$$

This linear partial differential equation for Γ_1 can be written symbolically as

$$\mathcal{D}_i \Gamma_1(\pi, H) = 0 , \quad (19.40)$$

with

$$\mathcal{D}_i = \int d^d x \left(\frac{\delta \Gamma_0}{\delta \pi_i(x)} \frac{\delta}{\delta H(x)} + \frac{\delta \Gamma_0}{\delta H(x)} \frac{\delta}{\delta \pi_i(x)} \right) . \quad (19.41)$$

The commutator $[\mathcal{D}_i, \mathcal{D}_j]$ is given by (in the calculation, the space dependence plays no role and can be omitted),

$$[\mathcal{D}_i, \mathcal{D}_j] = \left(\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial \pi_j \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial \pi_j} \right) \frac{\partial}{\partial H} + \left(\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial H \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial H} \right) \frac{\partial}{\partial \pi_j} - (i \leftrightarrow j) .$$

Differentiating equation (19.36) with respect to π_j and H , respectively, one obtains

$$\begin{aligned} \frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial \pi_j \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial \pi_j} + H \delta_{ij} &= 0 , \\ \frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial H \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial H} + \pi_i &= 0 . \end{aligned}$$

One concludes that

$$[\mathcal{D}_i, \mathcal{D}_j] = \int d^d x \left(\pi_j(x) \frac{\delta}{\delta \pi_i(x)} - \pi_i(x) \frac{\delta}{\delta \pi_j(x)} \right).$$

The commutators $[\mathcal{D}_i, \mathcal{D}_j]$ are the generators of the subgroup $O(N-1)$. We recognize that \mathcal{D}_i is a generator in $O(N)/O(N-1)$ acting on functionals of $\boldsymbol{\pi}$ and H . The commutators applied to an $O(N-1)$ -invariant functional thus vanish, which shows that the system (19.39) is compatible (see also Section 13.1).

Renormalization. We now examine the large cut-off behaviour (or the behaviour when d approaches 2). Equation (19.39) is satisfied for all values of the regularizing parameter. We conclude that the divergent part $\Gamma_1^{\text{div.}}$ of Γ , defined in any minimal subtraction scheme, also satisfies equation (19.39):

$$\mathcal{D}_i \Gamma_1^{\text{div.}} = 0.$$

By adding $-g\Gamma_1^{\text{div.}}(\boldsymbol{\pi}, H)$ to the action $\mathcal{S}(\boldsymbol{\pi}, H)$ of the tree approximation, we render the theory finite at one-loop order. Actually, it is necessary to also add higher order terms to construct the one-loop renormalized action $\mathcal{S}_1(\boldsymbol{\pi}, H)$:

$$\mathcal{S}_1(\boldsymbol{\pi}, H) = \mathcal{S}(\boldsymbol{\pi}, H) - g\Gamma_1^{\text{div.}}(\boldsymbol{\pi}, H) + \sum_2^\infty g^n \delta \mathcal{S}_1^{(N)}(\boldsymbol{\pi}, H).$$

These terms do not contribute to the one-loop order, which is now finite, and are chosen in such a way that $\mathcal{S}_1(\boldsymbol{\pi}, H)$ satisfies the non-linear equation (19.36) exactly. Indeed, at order 0, equation (19.36) is verified, since $\mathcal{S}(\boldsymbol{\pi}, H)$ satisfies it. At order 1, equation (19.36) implies equation (19.39) for $\Gamma_1^{\text{div.}}(\boldsymbol{\pi}, H)$ which is also satisfied. Higher order equations determine the higher order terms $\delta \mathcal{S}_1^{(n)}$.

We now generalize the argument to all orders in the loop expansion. We proceed by induction over the number of loops: we assume that it has been possible to construct an action $\mathcal{S}_{n-1}(\boldsymbol{\pi}, H)$ that satisfies equation (19.36) exactly, and such that $\Gamma_1, \dots, \Gamma_{n-1}$ have been rendered finite. Then, as we have shown in Section 19.5.2, the generating functional $\Gamma(\boldsymbol{\pi}, H)$ renormalized up to order $(n-1)$ also satisfies equation (19.35). We write equation (19.35) symbolically as

$$\Gamma * \Gamma = K. \quad (19.42)$$

The n th order ($n > 0$) in a loop expansion of equation (19.42) then takes the form

$$\sum_{p=0}^n \Gamma_p * \Gamma_{n-p} = 0, \quad (19.43)$$

and, therefore,

$$\Gamma_0 * \Gamma_n + \Gamma_n * \Gamma_0 = - \sum_{p=1}^{n-1} \Gamma_p * \Gamma_{n-p}. \quad (19.44)$$

The induction hypothesis implies that the right-hand side is finite. The divergent part of the equation thus satisfies

$$\Gamma_0 * \Gamma_n^{\text{div.}} + \Gamma_n^{\text{div.}} * \Gamma_0 = 0. \quad (19.45)$$

The form of the equation is independent of n . We then define \mathcal{S}_n , the renormalized action at order n , by

$$\mathcal{S}_n = \mathcal{S}_{n-1} - g^n \Gamma_n^{\text{div.}} + \sum_{n+1}^{\infty} g^p \delta \mathcal{S}_n^{(p)}. \quad (19.46)$$

It follows that

$$\begin{aligned} \mathcal{S}_n * \mathcal{S}_n - K &= (\mathcal{S}_{n-1} - g^n \Gamma_n^{\text{div.}}) * (\mathcal{S}_{n-1} - g^n \Gamma_n^{\text{div.}}) - K + O(g^{n+1}) \\ &= -g^n (\mathcal{S}_{n-1} * \Gamma_n^{\text{div.}} + \Gamma_n^{\text{div.}} * \mathcal{S}_{n-1}) + O(g^{n+1}). \end{aligned} \quad (19.47)$$

Since at this order in the right-hand side, \mathcal{S}_{n-1} can be replaced by Γ_0 , that is, $\mathcal{S}(\boldsymbol{\pi}, H)$, equation (19.45) then implies

$$\mathcal{S}_n * \mathcal{S}_n - K = O(g^{n+1}). \quad (19.48)$$

Hence, \mathcal{S}_n satisfies equation (19.36) at order n . As in the case $n = 1$, we then choose the higher order terms $\delta \mathcal{S}_n^{(p)}$ in such a way that \mathcal{S}_n satisfies equation (19.36) identically.

This concludes the induction. The generating functional of renormalized vertex functions satisfies equation (19.35), while the complete renormalized action satisfies equation (19.36).

The renormalized action is now the general solution of equation (19.36), consistent with $O(N - 1)$ symmetry, locality and power counting.

The measure. In the discussion, we have not mentioned the role of the measure. In the next section, we verify that the field transformation has been renormalized, and thus the measure must have changed accordingly; our discussion was really only valid for dimensional regularization where the measure term vanishes identically. We now extend the derivation to the case of lattice regularization.

We have shown in Section 19.5.1 that the invariant measure for the transformations (19.37) is

$$\prod_x d\boldsymbol{\pi}(x) \left[\frac{\delta \mathcal{S}}{\delta H(x)} \right]^{-1}.$$

We note that the vertices generated by the measure term are not multiplied by a factor $1/g$, in contrast with those coming from the classical action. As a consequence, they always contribute at the next order relative to the vertices coming from the action. For instance, at one-loop order, they contribute under the form of their tree approximation. Therefore, once Γ_n is rendered finite, if we modify the measure term by a divergent term of order g^n , to take into account the n th order renormalization, this will affect $\Gamma_{n+1}, \Gamma_{n+2}, \dots$ which are not yet renormalized, but leave Γ_n unchanged. Therefore, it is possible to introduce the field renormalization into the measure without changing the arguments given above about renormalization.

Actually, the measure term cancels the potential quadratic divergences which could appear at next order, according to power counting. It prevents, at the same time, the generation of a mass by loop corrections.

19.7 The renormalized action: Solution to the master equation

We first determine the renormalized action, the general solution of the master equation (19.36), and then show how the renormalizations appear order by order in perturbation theory.

19.7.1 The solution of the master equation for $d = 2$

The general form of the renormalized action is obtained by solving equation (19.36) [196]:

$$\int d^2x \left(\frac{\delta\mathcal{S}_r}{\delta\boldsymbol{\pi}(x)} \frac{\delta\mathcal{S}_r}{\delta H(x)} + H(x)\boldsymbol{\pi}(x) \right) = 0.$$

Power counting tells us that for $d = 2$, the dimension $[\boldsymbol{\pi}]$ of $\boldsymbol{\pi}$ and $[H]$ of H are, respectively,

$$[\boldsymbol{\pi}] = 0, \quad [H] = 2.$$

Since the action density has dimension 2, the action has the general form,

$$\mathcal{S}_r(\boldsymbol{\pi}, H) = \mathcal{S}_r(\boldsymbol{\pi}) - \int d^2x \sigma_r(\boldsymbol{\pi}(x)) H(x), \quad (19.49)$$

a quadratic term in H having at least dimension 4. Moreover, the coefficient $\sigma_r(\boldsymbol{\pi})$ is dimensionless and thus a derivative-free function of $\boldsymbol{\pi}(x)$, while $\mathcal{S}_r(\boldsymbol{\pi})$ has dimension 2 and contains terms with at most two derivatives.

The coefficient of $H(x)$ in equation (19.36) yields

$$\sigma_r(\boldsymbol{\pi}(x)) \frac{\delta\sigma_r}{\delta\boldsymbol{\pi}(x)} + \boldsymbol{\pi}(x) = 0. \quad (19.50)$$

Since $\sigma_r(\boldsymbol{\pi})$ is derivative-free and $O(N-1)$ -invariant, the general solution of equation (19.50) is simply

$$\sigma_r^2(\boldsymbol{\pi}(x)) + \boldsymbol{\pi}^2(x) = Z^{-1}. \quad (19.51)$$

This shows that $\sigma_r(\boldsymbol{\pi})$ is the renormalized σ -field, and Z is the field renormalization constant. For $H(x) = 0$, the equation reduces to

$$\int d^d x \frac{\delta\mathcal{S}_r(\boldsymbol{\pi})}{\delta\boldsymbol{\pi}(x)} \sigma_r(\boldsymbol{\pi}(x)) = 0. \quad (19.52)$$

The equation implies that $\mathcal{S}_r(\boldsymbol{\pi})$ is invariant under an infinitesimal transformation of the form

$$\delta_\omega \boldsymbol{\pi}(x) = \boldsymbol{\omega} \sigma_r(\boldsymbol{\pi}),$$

or solving equation (19.51):

$$\delta_\omega \boldsymbol{\pi}(x) = \boldsymbol{\omega} (Z^{-1} - \boldsymbol{\pi}^2(x))^{1/2}. \quad (19.53)$$

This is the renormalized form of the non-linear part of the $O(N)$ transformations. Equations (19.51) and (19.52) show that the renormalized functional $\mathcal{S}_r(\boldsymbol{\pi})$ is $O(N)$ -invariant, but the radius of the sphere has been renormalized. Therefore, the renormalized action can be written as

$$\mathcal{S}_r(\boldsymbol{\pi}, H) = \frac{1}{2Z_g} \int \left[(\nabla\boldsymbol{\pi}(x))^2 + (\nabla\sigma_r(x))^2 \right] d^d x - h \int \sigma_r(x) d^d x, \quad (19.54)$$

with the relation

$$\sigma_r(x) = [Z^{-1} - \boldsymbol{\pi}^2(x)]^{1/2}. \quad (19.55)$$

We have proved a rather remarkable result: although the interaction in the non-linear σ -model is not polynomial, the theory can be renormalized with only two renormalization constants [196]. In particular, by giving a mass to the $\boldsymbol{\pi}$ -field through a term of the form $\int \sigma(x) d^d x$, we have introduced no additional renormalization constant. This would not have been the case for a term of the form $\int \boldsymbol{\pi}^2(x) d^d x$, for example.

Let us also stress the similarity with the linear $O(N)$ -invariant $(\phi^2)^2$ field theory in four dimensions—the only differences come from the absence of the renormalization of $(\phi^2)^2$ coupling, which has no equivalent in the non-linear theory, and the structure of the mass renormalization. In particular, no spontaneous mass is generated.

19.7.2 Linearized WT identities

In our discussion of the renormalization of the model we have admitted that, in addition to the necessary counter-terms, additional higher order local terms $\delta\mathcal{S}_n^{(p)}$ could be added to the action, to render \mathcal{S}_n exactly $O(N)$ -invariant. This is a point that can be easily justified here.

Explicit solution of the linearized WT identities. To find the explicit general solution of the equations (19.39) or (19.45) satisfied by the divergent part of Γ , we first discuss the general solution of an equation of the form

$$\int d^d x \left(\frac{\delta \mathcal{S}}{\delta \pi_i(x)} \frac{\delta}{\delta H(x)} + \frac{\delta \mathcal{S}}{\delta H(x)} \frac{\delta}{\delta \pi_i(x)} \right) \mathcal{O}(\boldsymbol{\pi}, H) = 0, \quad (19.56)$$

in which $\mathcal{O}(\boldsymbol{\pi}, H)$ is an arbitrary $O(N-1)$ symmetric local functional, and \mathcal{S} is the initial action,

$$\mathcal{S}(\boldsymbol{\pi}, H) = \int d^d x \left\{ \frac{1}{2} \left[(\nabla \boldsymbol{\pi}(x))^2 + (\nabla \sigma(x))^2 \right] - H(x) \sigma(x) \right\}, \quad (19.57)$$

with $\sigma(x) = (1 - \boldsymbol{\pi}^2(x))^{1/2}$. If we define

$$\alpha(x) = \frac{1}{\sigma(x)} [H(x) + \Delta\sigma(x)], \quad (19.58)$$

equation (19.56) can be written, explicitly, as

$$\int d^d x \left\{ [-\nabla^2 \pi_i(x) + \alpha(x) \pi_i(x)] \frac{\delta}{\delta H(x)} - \sigma(x) \frac{\delta}{\delta \pi_i(x)} \right\} \mathcal{O}(\boldsymbol{\pi}, H) = 0. \quad (19.59)$$

The operator $\alpha(x)$ is an affine function of $H(x)$. We change variables in equation (19.59), $H \mapsto \alpha$, and consider $\mathcal{O}(\boldsymbol{\pi}, H)$ as a functional $\tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)$. A short but careful calculation then leads to the new equation,

$$\int d^d x \sigma(x) \frac{\delta \tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)}{\delta \pi_i(x)} = 0, \quad (19.60)$$

which shows that $\tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)$ is $O(N)$ symmetric at $\alpha(x)$ fixed.

This result has two applications: it completes the proof and, as we will show, it also yields the renormalized form of a general $O(N)$ -invariant local functional.

From power counting, we know that $\Gamma_n^{\text{div.}}$ has dimension 2. Therefore, it has terms of degree 0 and 1 in α . According to the preceding result, it has the form

$$\Gamma_n^{\text{div.}} = \frac{1}{2} a_n \int \left\{ [\nabla \boldsymbol{\pi}(x)]^2 + [\nabla \sigma(x)]^2 \right\} d^d x + \frac{1}{2} b_n \int \alpha(x) d^d x. \quad (19.61)$$

The first term can be absorbed into a coupling constant renormalization:

$$g \mapsto g (1 + g^n a_n). \quad (19.62)$$

We now calculate the variation of the action when the radius of the sphere is renormalized. The variation $\delta\sigma(x)$ of $\sigma(x)$ is

$$\delta\sigma(x) = [1 - \delta Z - \boldsymbol{\pi}^2(x)]^{1/2} - [1 - \boldsymbol{\pi}^2(x)]^{1/2} = -\frac{1}{2} \delta Z / \sigma(x).$$

It follows that

$$\delta \left\{ \int d^d x \left[\frac{1}{2} [\nabla \sigma(x)]^2 - H(x) \sigma(x) \right] \right\} = \frac{1}{2} \delta Z \int \alpha(x) d^d x. \quad (19.63)$$

Therefore, the second term can be absorbed into a field renormalization:

$$\delta Z = b_n g^n. \quad (19.64)$$

This completes the proof of the renormalization of the non-linear σ -model. Nevertheless, for completeness, we outline in Section 19.9 a different derivation, whose basic idea is to return to a linear formulation of the symmetry.

19.8 Renormalization of local functionals

General $O(N)$ -invariant local functionals. From the general solution of equation (19.56), one also infers the structure of renormalized $\phi(x)$ correlation functions with one insertion of an $O(N)$ -invariant local functional of arbitrary dimension. Indeed, to generate such insertions, one can add to the action a source term of the form $\int d^d x K(x) \mathcal{O}(\boldsymbol{\pi}(x))$. Since $\mathcal{O}(\boldsymbol{\pi})$ is $O(N)$ invariant, equation (19.36) holds for the action to which includes this new term has been added. An expansion of equation (19.36) at first order in $K(x)$ leads to equation (19.56) with $\mathcal{S}(\boldsymbol{\pi}, H)$ replaced by the renormalized action. The renormalized operator $\mathcal{O}(\boldsymbol{\pi}, H)$ is thus the most general local functional of the renormalized fields $\boldsymbol{\pi}(x)$ and $\sigma(x)$, of dimension $[\mathcal{O}]$, $O(N)$ invariant at $\alpha(x)$ fixed [196, 197].

Renormalization of dimensionless operators and parametrization of the sphere. Dimensionless operators are derivative-free local functions of the $\boldsymbol{\pi}(x)$ field. A simple extension of previous arguments shows that they should be classified according to irreducible representations of the $O(N)$ group. A new renormalization constant is associated to each different irreducible representation. For example, a mass term $\int d^d x \boldsymbol{\pi}^2(x)$ is a component of the symmetric traceless tensor

$$\int d^d x [\phi_i(x) \phi_j(x) - \frac{1}{N} \delta_{ij} \phi^2(x)],$$

and introduces an additional renormalization constant. It now becomes clear why we have chosen the particular parametrization of the sphere in terms of the $\boldsymbol{\pi}$ field.

Fields $\boldsymbol{\theta}(x)$ corresponding to another $O(N-1)$ symmetric parametrization of the sphere are related to $\boldsymbol{\pi}(x)$ by

$$\boldsymbol{\theta}(x) = \boldsymbol{\pi}(x) \sum_{\ell} f_{\ell} (\boldsymbol{\pi}^2(x))^{\ell}.$$

We rewrite this expansion in terms of the $(i, 0, \dots, 0)$ components of the tensors transforming under irreducible representations of $O(N)$. These components have the form $\boldsymbol{\pi}(x) P_{\ell}^N(\sigma(x))$, in which P_{ℓ}^N are hyperspherical polynomials:

$$\boldsymbol{\theta}(x) = \boldsymbol{\pi}(x) \sum c_{\ell} P_{\ell}^N(\sigma(x)).$$

Each non-vanishing coefficient c_{ℓ} introduces an independent renormalization constant [198]. In particular, in the generic case in which all coefficients c_{ℓ} are present, the renormalization of $\boldsymbol{\theta}(x)$ corresponds to an arbitrary change of parametrization:

$$\boldsymbol{\theta}(x) = \boldsymbol{\theta}_r(x) Z [\boldsymbol{\theta}_r^2(x)].$$

19.9 A linear representation

For models on homogeneous spaces, it is always possible to express the relations between the components of the field in the linear representation by introducing a set of Lagrange multipliers. In the case of the non-linear σ -model, this is especially easy since only one $O(N)$ -invariant additional field, which we denote by $\alpha(x)$, is required to implement the constraint (19.1):

$$\phi^2(x) = 1.$$

The integral representation (19.28) of the generating functional $\mathcal{Z}(\mathbf{J})$ can also be written as (see also Section 18.6.2)

$$\mathcal{Z}(\mathbf{J}) = \int [d\alpha d\phi] \exp \left[-\frac{1}{g} \left(\mathcal{S}(\phi, \alpha) - \int d^d x \mathbf{J}(x) \cdot \phi(x) \right) \right], \quad (19.65)$$

where the integration contour for α is parallel to the imaginary axis, and

$$\mathcal{S}(\phi, \alpha) = \frac{1}{2} \int d^d x \left\{ [\nabla \phi(x)]^2 + \alpha(x) [\phi^2(x) - 1] \right\}. \quad (19.66)$$

Note that, here, $\mathbf{J}(x)$ is an N -component source for the N -component field $\phi(x)$.

We choose an extremum of the potential as the starting point for perturbation theory:

$$\alpha(x) = 0, \quad \sigma(x) \equiv \phi_1(x) = 1, \quad \phi_i(x) = 0 \quad \text{for } i > 1. \quad (19.67)$$

The propagator of the $(N - 1)$ remaining components $\pi(x)$ is still proportional to $1/p^2$. The component $\sigma(x)$ is coupled to $\alpha(x)$. At leading order, the matrix of connected two-point functions is

$$\begin{pmatrix} \langle \sigma\sigma \rangle & \langle \sigma\alpha \rangle \\ \langle \alpha\sigma \rangle & \langle \alpha\alpha \rangle \end{pmatrix}_{\text{c}}(p) = \begin{pmatrix} 0 & g \\ g & gp^2 \end{pmatrix}. \quad (19.68)$$

In this formulation, the $O(N)$ symmetry is realized linearly but, nevertheless, the symmetry in the tree approximation is automatically spontaneously broken. The consequences of such a situation have already been analysed in Section 13.4. Power counting tells us that the dimensions of the fields in dimension 2 are

$$[\phi] = 0, \quad [\alpha] = 2.$$

The general form of the renormalized action $\mathcal{S}_r(\phi, \alpha)$ follows:

$$\begin{aligned} \mathcal{S}_r(\phi, \alpha) = & \frac{1}{2} \int d^d x \left\{ B_1(\phi^2(x)) [\nabla \phi(x)]^2 + B_2(\phi^2(x)) (\phi(x) \cdot \nabla \phi(x))^2 \right. \\ & \left. + B_3(\phi^2(x)) \alpha(x) - B_4(\phi^2(x)) \right\}, \end{aligned} \quad (19.69)$$

in which B_1 , B_2 , B_3 , and B_4 are arbitrary functions and, therefore, an infinite number of renormalization constants are needed.

However, if one is interested only in the ϕ -field correlation functions (and not the α -field), one can integrate explicitly over the α -field. This fixes the value of $\phi^2(x)$ to be a solution of the equation

$$B_3(\phi^2(x)) = 0. \quad (19.70)$$

This equation has a solution order by order in perturbation theory of the form

$$\phi^2(x) = Z^{-1}(g) = 1 + O(g). \quad (19.71)$$

After integration, B_1 , B_2 , and B_3 become pure constants and $\phi \cdot \nabla \phi$ vanishes identically.

Therefore, one recovers the results of the previous sections. The main disadvantage of this formulation is that it introduces an infinite number of renormalization constants as an intermediate step, and that its generalization to other homogeneous spaces is complicated, and not aesthetically appealing.

Finally, the non-linear formulation emphasizes the connection with other geometric models like gauge theories. On the other hand, the linear formulation clarifies the discussion of multiple insertions of general $O(N)$ -invariant operators. In particular, one understands that the insertions of operators of dimensions larger than 2 eventually generate terms of degrees larger than 1 in $\alpha(x)$. The integration over α then no longer implies the strict constraint (19.71).

19.10 $(\phi^2)^2$ field theory in the ordered phase and non-linear σ -model

Before we begin discussing the physics of the σ -model, we review shortly the arguments that, beyond the $1/N$ expansion (Section 18.6), confirm the relation between the non-linear σ -model and the $(\phi^2)^2$ field theory. In particular, we show that, at fixed temperature $T < T_c$, the non-linear σ -model emerges from the analysis of the large distance behaviour in the ordered phase of the $(\phi^2)^2$ field theory (for an early work see Ref. [199]).

19.10.1 Lattice spin models

One possible regularization of the non-linear σ -model is an $O(N)$ symmetric spin model on the lattice, as we have shown in Section 19.4.2. The regularized model belongs to the class studied in Chapters 14–16: the variables are classical spins \mathbf{S}_i of unit length on a lattice of site i , interacting through a short range ferromagnetic $O(N)$ symmetric two-body interaction V_{ij} . In zero field, the partition function can be written as (T is the temperature)

$$\mathcal{Z} = \int \prod_i d\mathbf{S}_i \delta(\mathbf{S}_i^2 - 1) \exp[-\mathcal{H}(\mathbf{S})/T], \quad (19.72)$$

in which the configuration energy \mathcal{H} has the form

$$\mathcal{H}(\mathbf{S})/T = - \sum_{i,j} V_{ij} \mathbf{S}_i \cdot \mathbf{S}_j / T. \quad (19.73)$$

We have shown, by different techniques, summation of the most IR terms at the critical temperature T_c of the mean-field expansion, direct RG analysis, that the universal properties in the critical domain of such models can be described in d dimensions, for $(4-d)$ small, by the corresponding $(\phi^2)^2$ field theory.

19.10.2 The $(\phi^2)^2$ field theory in the ordered phase

Conversely, we now consider the $(\phi^2)^2$ field theory:

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} u (\phi^2(x))^2 \right], \quad (19.74)$$

in the low temperature phase ($r < r_c = r(T_c)$).

At low temperature, that is, T fixed, $T < T_c$, in a system in which a discrete symmetry is spontaneously broken, the connected correlation functions decrease exponentially. The situation is quite different, as we have already pointed out, when the symmetry is continuous, because Goldstone modes are present. We now show that the non-linear σ -model describes the interaction between Goldstone modes.

We change variables in the field integral

$$\mathcal{Z} = \int [d\phi] \exp[-\mathcal{S}(\phi)],$$

setting (a change meaningful, because $\langle \rho \rangle > 0$)

$$\phi(x) = \rho(x)\hat{\phi}(x), \quad \text{with} \quad \hat{\phi}^2(x) = 1. \quad (19.75)$$

The field integral becomes (assuming a lattice or, in the continuum, a dimensional regularization to deal with the ρ measure):

$$\mathcal{Z} = \int [\rho^{N-1}(x)d\rho(x)] [d\hat{\phi}(x)] \exp [-\mathcal{S}(\rho, \hat{\phi})], \quad (19.76)$$

with

$$\mathcal{S}(\rho, \hat{\phi}) = \int d^d x \left\{ \frac{1}{2} \rho^2(x) [\nabla \hat{\phi}(x)]^2 + \frac{1}{2} [\nabla \rho(x)]^2 + \frac{1}{2} r \rho^2(x) + \frac{1}{4!} u \rho^4(x) \right\}. \quad (19.77)$$

The integration over the $\rho(x)$ field generates an effective action $\mathcal{S}_{\text{eff.}}(\hat{\phi})$ for the field $\hat{\phi}$:

$$\exp [-\mathcal{S}_{\text{eff.}}(\hat{\phi})] = \int [\rho^{N-1}(x)d\rho(x)] \exp [-\mathcal{S}(\rho, \hat{\phi})]. \quad (19.78)$$

In the ordered phase, below T_c , the field $\rho(x)$ has a non-zero expectation value and is massive; its dynamics, therefore, is not critical. As long as we explore momenta much smaller than the ρ -mass or distances much larger than the corresponding correlation length, the effective action resulting from the integration over the ρ -field can be expanded in local terms. To evaluate $\mathcal{S}_{\text{eff.}}$ explicitly, we define

$$\mathcal{S}(\rho) = \int d^d x \left\{ \frac{1}{2} [\nabla \rho(x)]^2 + \frac{1}{2} r \rho^2(x) + \frac{1}{4!} u \rho^4(x) \right\}, \quad (19.79)$$

and expand expression (19.78) in powers of $[\nabla \hat{\phi}(x)]^2$. Denoting by $\langle \bullet \rangle$ expectation values corresponding to the action (19.79), one obtains

$$\begin{aligned} \mathcal{S}_{\text{eff.}}(\hat{\phi}) = & \frac{1}{2} \langle \rho^2 \rangle \int d^d x [\nabla \hat{\phi}(x)]^2 - \frac{1}{8} \int d^d x d^d y \langle \rho^2(x) \rho^2(y) \rangle_{\text{conn.}} [\nabla \hat{\phi}(x)]^2 [\nabla \hat{\phi}(y)]^2 \\ & + O(\rho^6). \end{aligned} \quad (19.80)$$

At large distance, the leading contribution comes from the first term, which has only two derivatives. Due to the $O(N)$ symmetry, it has the form of the action of the non-linear σ -model. In the second term, the ρ^2 connected two-point function has to be expanded for large ρ mass, yielding a derivative expansion acting on $\delta^{(d)}(x - y)$. The leading term is proportional to $\int d^d x ([\nabla \hat{\phi}(x)]^2)^2$, a subleading interaction with four derivatives. All other neglected terms have at least six derivatives. In the expansion (19.80) all contributions, except the first one, thus correspond to irrelevant operators.

Explicit calculations involve then shifting the ρ field, $\rho(x) = \langle \rho \rangle + \rho'(x)$, evaluating perturbatively the successive ρ^2 connected correlation functions, and performing a local expansion of each term for large ρ mass.

It is difficult to characterize precisely the domain of validity of the resulting effective theory, because the ρ^2 correlation functions can only be calculated perturbatively, but we expect it to include the low temperature phase with T fixed, $T < T_c$. In this regime, the non-linear σ -model (19.80) completely describes the long distance properties of the $(\phi^2)^2$ field theory. In addition, the coefficient in front of the effective action becomes large at low temperature like in the lattice model. Therefore, from several different point of views, we verify the relation between σ -model and $(\phi^2)^2$ field theory in the continuum.

At higher temperatures, problems arise if the configurations for which ϕ vanishes become important. One example is provided by the $O(2)$ model in two dimensions (see Chapter 31).

19.10.3 Spontaneous symmetry breaking: The role of dimension 2

By three different methods, we have shown that the large distance physics of the $O(N)$ symmetric N -vector model can be described, below T_c , by the non-linear σ -model. However, the non-linear σ model is especially well-suited to describe the ordered phase of the $O(N)$ symmetric vector model, because, within the perturbative expansion, it is automatically in a phase in which the $O(N)$ symmetry is spontaneously broken. It then describes the interactions between the $(N - 1)$ massless Goldstone modes, or spin wave excitations. Moreover, it makes a direct analysis of the role of dimension 2, where IR divergences appear, possible.

(i) In Section 14.3, we have argued that in the N -vector model, for $d > 2$, the $O(N)$ symmetry is spontaneously broken at low temperature. This argument is consistent with the property that, for $d > 2$, the perturbative non-linear σ model, which also predicts spontaneous symmetry breaking, is not IR divergent. However, the perturbative expansion does not describe the expected phase transition at a higher temperature.

(ii) For $d \leq 2$, we know from the Mermin–Wagner theorem (see also the analysis of Section 14.3) that spontaneous symmetry breaking with ordering ($\langle \mathbf{S} \rangle \neq 0$) is impossible in a model with a continuous symmetry and short range forces. This is again consistent with the appearance of IR divergences in perturbation theory. For $d = 2$, perturbation theory only makes sense in the presence of an IR cut-off, which explicitly breaks the symmetry and orders the spins. Therefore, perturbation theory gives no information about the long distance properties of the expected unbroken theory.

To go somewhat beyond perturbation theory, we now introduce field theory RG methods. Therefore, it is necessary to first define the model in the dimension in which it is renormalizable. Because the non-linear σ -model is renormalizable in two dimensions, IR divergences have to be dealt with, and we introduce again an IR cut-off in the form of a linear term in σ (equation (19.19)). We then proceed in a formal analogy with the example of the $(\phi^2)^2$ field theory, that is, study the theory in dimension $d = 2 + \varepsilon$ as a double series expansion in the temperature T and ε . In this way, the perturbative expansion remains renormalizable for $d = 2 + \varepsilon$ and RG equations follow.

19.11 Renormalization, RG equations

We now study the non-linear σ model from the point of view of renormalization and RG [200–202]. To generate perturbation theory, we return to the parametrization of the field ϕ of Section 19.1,

$$\phi(x) = \{\sigma(x), \boldsymbol{\pi}(x)\},$$

and locally eliminate the field $\sigma(x)$ by

$$\sigma(x) = (1 - \boldsymbol{\pi}^2(x))^{1/2}.$$

As we have done for the $(\phi^2)^2$ model, we scale all distances in order to measure momenta in units of the inverse lattice spacing Λ . We then consider the partition function in an external constant field h :

$$\mathcal{Z}(h) = \int \left[(1 - \boldsymbol{\pi}^2(x))^{-1/2} d\boldsymbol{\pi}(x) \right] \exp[-\mathcal{S}(\boldsymbol{\pi}, h)], \quad (19.81)$$

now with

$$\mathcal{S}(\boldsymbol{\pi}, h) = \frac{1}{g} \int d^d x \left\{ \frac{1}{2} \left[(\nabla \boldsymbol{\pi}(x))^2 + (\sigma(x))^2 \right] - h \sqrt{1 - \boldsymbol{\pi}^2(x)} \right\}. \quad (19.82)$$

We introduce a dimensionless coupling constant

$$t = g \Lambda^{d-2}. \quad (19.83)$$

We have shown that the model is renormalizable in two dimensions, and characterized the form of the renormalized action (Sections 19.6, 19.7). We recall that the presence of a non-vanishing external field h is required, for $d \leq 2$, both to select the classical minimum of the action around which to expand perturbation theory, and to provide the theory with an IR cut-off (Section 19.3). The renormalized action can be written as

$$\mathcal{S}_r(\boldsymbol{\pi}_r) = \frac{\mu^{d-2} Z}{2t_r Z_t} \int d^d x \left[(\nabla \boldsymbol{\pi}_r(x))^2 + (\nabla \sigma_r(x))^2 \right] - \frac{\mu^{d-2}}{t_r} h_r \int \sigma_r(x) d^d x, \quad (19.84)$$

in which μ is the renormalization scale, and

$$\sigma_r(x) = [Z^{-1} - \boldsymbol{\pi}_r^2]^{1/2}. \quad (19.85)$$

Note that the renormalization constants can be chosen to be h independent. This is automatically realized in the minimal subtraction scheme.

The relations

$$t = (\Lambda/\mu)^{d-2} Z_t t_r, \quad \boldsymbol{\pi}_r(x) = Z^{-1/2} \boldsymbol{\pi}(x), \quad (19.86)$$

imply

$$h = Z_h h_r, \quad Z_h = Z_t / \sqrt{Z}. \quad (19.87)$$

With our conventions, the coupling constant, which is proportional to the temperature, is dimensionless. The relation between the cut-off dependent and the renormalized vertex functions, in the Fourier representation, is

$$Z^{n/2} (\Lambda/\mu, t) \tilde{\Gamma}^{(n)}(p_i; t, h, \Lambda) = \tilde{\Gamma}_r^{(n)}(p_i; t_r, h_r, \mu). \quad (19.88)$$

$$V^{(4)} = \frac{1}{8} \delta_{i_1 i_2} \delta_{i_3 i_4} [(p_1 + p_2)^2 + h] \quad \frac{1}{2}(N-1)h\Omega_d(\sqrt{h}) \quad p^2\Omega_d(\sqrt{h})$$

Fig. 19.1 One-loop diagrams: the dotted lines do not correspond to propagators, but are used only to represent faithfully the flow of group indices

Differentiating with respect to Λ , *at renormalized parameters fixed*, we obtain the RG equations [201]:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} - \frac{n}{2} \zeta(t) + \rho(t) h \frac{\partial}{\partial h} \right] \tilde{\Gamma}^{(n)}(p_i; t, h, \Lambda) = 0. \quad (19.89)$$

We have assumed that the renormalization constants, and thus the RG functions defined by

$$\Lambda \frac{\partial}{\partial \Lambda} \Big|_{\text{ren. fixed}} \quad t = \beta(t), \quad \Lambda \frac{\partial}{\partial \Lambda} \Big|_{\text{ren. fixed}} \quad (-\ln Z) = \zeta(t), \quad \Lambda \frac{\partial}{\partial \Lambda} \Big|_{\text{ren. fixed}} \quad \ln h = \rho(t),$$

have been chosen h independent. The coefficient of $\partial/\partial h$ can be derived from equation (19.87), which implies (taking the logarithm of both members),

$$0 = h^{-1} \Lambda \frac{\partial}{\partial \Lambda} h + d - 2 - \frac{1}{2} \zeta(t) - \frac{\beta(t)}{t}, \quad (19.90)$$

and, therefore, setting $d = 2 + \varepsilon$ and $\zeta(t) = \varepsilon + \eta(t)$,

$$\rho(t) = \frac{1}{2} (\eta(t) - \varepsilon) + \frac{\beta(t)}{t}. \quad (19.91)$$

Thus,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} - \frac{n}{2} (\eta(t) + \varepsilon) + \left(\frac{1}{2} (\eta(t) - \varepsilon) + \frac{\beta(t)}{t} \right) h \frac{\partial}{\partial h} \right] \tilde{\Gamma}^{(n)}(p_i; t, h, \Lambda) = 0. \quad (19.92)$$

A discussion of correlation functions involving the σ -field, also requires RG equations satisfied by connected correlation functions $\tilde{W}^{(n)}$:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} + \frac{n}{2} (\eta(t) + \varepsilon) + \rho(t) h \frac{\partial}{\partial h} \right] \tilde{W}^{(n)}(p_i; t, h, \Lambda) = 0. \quad (19.93)$$

19.12 RG equations: Solutions (magnetic terminology)

RG functions at one-loop order. The two RG functions can be obtained at one-loop order from a calculation of the two-point vertex function $\tilde{\Gamma}^{(2)}$ (Fig. 19.1):

$$\tilde{\Gamma}^{(2)}(p) = \frac{\Lambda^\varepsilon}{t} (p^2 + h) + \left[p^2 + \frac{1}{2}(N-1)h \right] \Omega_d(\sqrt{h}) + O(t), \quad (19.94)$$

where we have defined the function Ω_d in equation (18.16). Here (omitting the implicit cut-off dependence),

$$\Omega_d(\sqrt{h}) = \frac{1}{(2\pi)^d} \int_{d \rightarrow 2+}^{\Lambda} \frac{d^d q}{q^2 + h} \underset{\Lambda \rightarrow \infty}{=} \Omega_d(0) + \frac{1}{(4\pi)^{d/2}} \Gamma(1-d/2) h^{\varepsilon/2} + O(h\Lambda^{\varepsilon-2}), \quad (19.95)$$

with $\Omega_d(0)\Lambda^{-\varepsilon} \sim -1/(2\pi\varepsilon)$.

Applying equation (19.92) to $\tilde{\Gamma}^{(2)}$, and identifying the coefficients of p^2 and h , we obtain two equations, which determine $\beta(t)$ and $\zeta(t)$ at one-loop order:

$$\beta(t) = \varepsilon t - \frac{(N-2)}{2\pi} t^2 + O(t^3, t^2\varepsilon), \quad (19.96)$$

$$\eta(t) = \zeta(t) - \varepsilon = -\varepsilon + \frac{(N-1)}{2\pi} t + O(t^2, t\varepsilon). \quad (19.97)$$

From the expression of $\beta(t)$ of equation (19.96), we immediately conclude:

for $d \leq 2$ ($\varepsilon \leq 0$), $t = 0$ is an unstable IR fixed point, this IR instability being induced by the vanishing mass of the would-be Goldstone bosons. The spectrum of the theory thus is not given by perturbation theory, and the perturbative assumption of spontaneous symmetry breaking at low temperature is inconsistent. As mentioned before, this result agrees with rigorous arguments. Moreover, since the model depends only on one coupling constant, $t = 0$ is also a UV stable fixed point (the property of large momentum asymptotic freedom, see also Section 23.4.1). Section 19.14 contains a short discussion of the physics in two dimensions for $N > 2$. The special Abelian case $N = 2$ is examined in Chapter 31.

For $d > 2$, that is, $\varepsilon > 0$, $t = 0$ is a stable IR fixed point, the $O(N)$ symmetry is spontaneously broken at low temperature in zero field. The effective coupling constant, which determines the large distance behaviour, approaches the origin for all temperatures $t < t_c$, t_c being the possible first non-trivial zero of $\beta(t)$. Therefore, the large distance properties of the model can be obtained from the low temperature expansion and RG, replacing the perturbative parameters by effective parameters obtained by solving the RG equations.

Finally, we note that, at least for ε positive and small, and $N > 2$, the RG function $\beta(t)$ has a non-trivial zero,

$$\beta(t_c) = 0, \quad \text{for } t_c = \frac{2\pi\varepsilon}{N-2} + O(\varepsilon^2), \quad \text{with} \quad \beta'(t_c) = -\varepsilon + O(\varepsilon^2). \quad (19.98)$$

Since t_c is a UV fixed point, it governs the large momentum behaviour of the renormalized theory. The large momentum behaviour of correlation functions is not given by perturbation theory but by the fixed point. As a consequence, the perturbative result that the theory cannot be rendered finite for $d > 2$ with a finite number of renormalization constants, cannot be trusted.

Moreover, since t_c is an unstable IR fixed point, it has the property of a *critical temperature*.

19.12.1 Critical temperature and exponents, for $d = 2 + \varepsilon$, $N > 2$

At t_c and for $h = 0$, equation (19.92) for $n = 2$ reduces to

$$\left(\Lambda \frac{\partial}{\partial \Lambda} - (\eta(t_c) + \varepsilon) \right) \tilde{\Gamma}^{(2)}(p; t_c, 0, \Lambda) = 0.$$

Since, from a dimensional relation,

$$\tilde{\Gamma}^{(2)}(p; t_c, 0, \Lambda) = \Lambda^{2+\varepsilon} \tilde{\Gamma}^{(2)}(p/\Lambda; t_c, 0, 1),$$

we conclude

$$\tilde{\Gamma}^{(2)}(p; t_c, 0, \Lambda) \propto \Lambda^{2+\varepsilon} (p/\Lambda)^{2-\eta(t_c)}.$$

We recognize the standard exponent η , related to the field renormalization,

$$\eta = \eta(t_c) = \frac{\varepsilon}{N-2} + O(\varepsilon^2) \Rightarrow d_\pi = \frac{1}{2}(d-2+\eta) = \frac{N-1}{2(N-2)}\varepsilon + O(\varepsilon^2). \quad (19.99)$$

For $|t - t_c| \ll 1$ and $h = 0$, the RG equation can be approximated by

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta'(t_c)(t - t_c) \frac{\partial}{\partial t} - (\eta + \varepsilon) \right) \tilde{\Gamma}^{(2)}(p; t, 0, \Lambda) = 0.$$

For $p = 0$, the solution has the scaling form,

$$\tilde{\Gamma}^{(2)}(0; t, 0, \Lambda) \propto \Lambda^{2+\varepsilon} |t - t_c|^{-(2-\eta)/\beta'(t_c)}.$$

We recognize the magnetic susceptibility exponent $\gamma = -(2 - \eta)/\beta'(t_c)$, and from the relation $\gamma = \nu(2 - \eta)$, we infer the correlation length exponent given by

$$\frac{1}{\nu} = -\frac{1}{\beta'(t_c)} = \varepsilon + O(\varepsilon^2). \quad (19.100)$$

19.12.2 Integration of the RG equations: $d > 2$, $t < t_c$

We now discuss the solutions of the RG equations more precisely. We first examine the implications of the RG equations for the large distance behaviour of correlation functions for $d > 2$, where $t = 0$ is an IR fixed point. As usual, we solve equation (19.92) by introducing a scaling parameter λ and looking for a solution of the form

$$\tilde{\Gamma}^{(n)}(p_i, t, h, \Lambda) = Z^{-n/2}(\lambda) \lambda^{-n\varepsilon/2} \tilde{\Gamma}^{(n)}(p_i, t(\lambda), h(\lambda), \lambda\Lambda). \quad (19.101)$$

Compatibility with equation (19.92) implies

$$\ln \lambda = \int_t^{t(\lambda)} \frac{dt'}{\beta(t')}, \quad (19.102a)$$

$$\ln Z(\lambda) = \int_t^{t(\lambda)} dt' \frac{\eta(t')}{\beta(t')}, \quad (19.102b)$$

$$h(\lambda) = \lambda^{2-d} Z^{1/2}(\lambda) \frac{t(\lambda)}{t} h. \quad (19.102c)$$

With our conventions, $\tilde{\Gamma}^{(n)}$ has the dimension d and h dimension 2. Taking into account dimensional analysis, we then rewrite relation (19.101) as

$$\tilde{\Gamma}^{(n)}(p_i, t, h, \Lambda) = Z^{-n/2}(\lambda)(\lambda\Lambda)^d \tilde{\Gamma}^{(n)}(p_i/\Lambda\lambda, t(\lambda), h(\lambda)/(\lambda\Lambda)^2, 1). \quad (19.103)$$

For $h \ll \Lambda^2$, perturbation theory has IR singularities. By choosing λ solution of the equation

$$h(\lambda) = (\lambda\Lambda)^2, \quad (19.104)$$

we ensure that the perturbation expansion in the effective theory at scale λ is no longer IR singular.

It is easy to verify that, at least for t small, $h \ll \Lambda^2$ implies $\lambda \rightarrow 0$. We then introduce three functions of the temperature $M_0(t)$, $\xi(t)$ and $K(t)$ defined by

$$M_0(t) = \exp \left[-\frac{1}{2} \int_0^t \frac{\eta(t') + \varepsilon}{\beta(t')} dt' \right], \quad (19.105)$$

$$\xi(t) = \Lambda^{-1} t^{1/\varepsilon} \exp \left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{\varepsilon t'} \right) dt' \right]. \quad (19.106)$$

$$K(t) = M_0(t) [\Lambda \xi(t)]^{d-2} / t = 1 + O(t). \quad (19.107)$$

The solution of equation (19.104) then takes the form

$$\lambda \sim K^{1/2}(t)h^{1/2}\Lambda^{-1}. \quad (19.108)$$

Because $t = 0$ is an IR fixed point, the scale-dependent temperature $t(\lambda) \rightarrow 0$, and thus the leading terms in the small h and small momenta limit, can be calculated perturbatively. Using equations (19.102a, b), one obtains the behaviours of $t(\lambda)$ and $Z(\lambda)$:

$$t(\lambda) \sim \lambda^{d-2}tK(t)M_0^{-1} \sim t [K(t)]^{d/2} M_0^{-1} h^{(d-2)/2} \Lambda^{2-d}, \quad (19.109)$$

$$Z(\lambda) \sim M_0^2(t). \quad (19.110)$$

It follows that

$$\begin{aligned} \tilde{\Gamma}^{(n)}(p_i, t, h, \Lambda) &\sim M_0^{-n}(t)[K(t)h]^{d/2} \\ &\times \tilde{\Gamma}^{(n)}\left(\frac{p_i}{[K(t)h]^{1/2}}, \frac{t[K(t)]^{d/2}}{M_0(t)}\left(\frac{h}{\Lambda^2}\right)^{(d-2)/2}, 1, 1\right). \end{aligned} \quad (19.111)$$

Actually, it is simple to verify directly, using dimensional analysis in the form

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + 2h \frac{\partial}{\partial h} + p_i \frac{\partial}{\partial p_i}\right) \tilde{\Gamma}^{(n)} = d \tilde{\Gamma}^{(n)},$$

that the expression (19.111) is the general solution of equation (19.92).

Similarly, one finds

$$\begin{aligned} \widetilde{W}^{(n)}(p_i; t, h, \Lambda) &= [K(t)h]^{(1-n)d/2} M_0^n(t) \\ &\times \widetilde{W}^{(n)}\left(\frac{p_i}{[K(t)h]^{1/2}}, \frac{t[K(t)]^{d/2}}{M_0(t)}\left(\frac{h}{\Lambda^2}\right)^{(d-2)/2}, 1, 1\right). \end{aligned} \quad (19.112)$$

The coexistence curve. This result can be applied to the determination of the singularities near the coexistence curve, that is, at t fixed below the critical temperature when the magnetic field h goes to zero. With our normalization, the magnetization satisfies

$$M(t, h, \Lambda) \equiv \langle \sigma(x) \rangle = W^{(1)}(t, h, \Lambda) = M_0(t)M(t(\lambda), 1, 1). \quad (19.113)$$

At one-loop order in a field, it is given by

$$M = 1 - \frac{N-1}{2} \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d q}{q^2 + h} \Lambda^{-\varepsilon} t + O(t^2).$$

Thus, using the relation (19.113), one finds

$$M(t, h, \Lambda = 1) = M_0(t) - \frac{N-1}{2} t [K(t)]^{d/2} h^{(d-2)/2} \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} + O(h, h^{d-2}).$$

This result shows that $M_0(t)$ is the *spontaneous magnetization*. The singularity $h^{(d-2)/2}$ gives an interpretation to the logarithm found in equation (16.150) and to the singularity for $x = -1$ found in equation (18.41).

19.12.3 Scaling of correlation functions, the critical domain

It is convenient to rewrite equation (19.111) as

$$\tilde{\Gamma}^{(n)}(p_i; t, h, \Lambda) = \xi^{-d}(t) M_0^{-n}(t) F^{(n)}(p_i \xi(t), h/h_0(t)), \quad (19.114)$$

with

$$h_0(t) = t M_0^{-1}(t) \xi^{-d}(t) \Lambda^{2-d}. \quad (19.115)$$

For the connected correlation functions, the same analysis leads to the form

$$\widetilde{W}^{(n)}(p_i; t, h, \Lambda) = \xi^{d(n-1)}(t) M_0^n(t) G^{(n)}(p_i \xi(t), h/h_0(t)). \quad (19.116)$$

The induced magnetization is

$$M = W^{(1)}(t, h, \Lambda) = M_0(t) G^{(1)}(h/h_0(t)). \quad (19.117)$$

Inversion of this relation yields the scaling form of the *equation of state*:

$$h = h_0(t) f(M/M_0(t)), \quad (19.118)$$

and the vertex functions can thus be written in terms of the magnetization as

$$\tilde{\Gamma}^{(n)}(p_i, t, M, \Lambda) = \xi^{-d}(t) M_0^{-n}(t) F^{(n)}(p_i \xi(t), M/M_0(t)). \quad (19.119)$$

Equations (19.118, 19.119) are consistent with equations (16.64, 16.73): the appearance of two different functions $\xi(t)$ and $M_0(t)$ is consistent with the existence of two independent critical exponents ν, β , in the $(\phi^2)^2$ field theory. They extend, in the large distance limit, the scaling form of correlation functions, valid in the critical region, to all temperatures below t_c . However, there is one important difference between the RG equations of the $(\phi^2)^2$ theory and of the σ -model: the $(\phi^2)^2$ theory depends on two coupling constants, the coefficient of ϕ^2 which plays the role of the temperature, and the coefficient of $(\phi^2)^2$ that has no equivalent here. The correlation functions of the continuum $(\phi^2)^2$ theory have the exact scaling form (19.119) only at the IR fixed point. In contrast, in the case of the σ -model, it has been possible to eliminate all corrections to scaling corresponding to irrelevant operators order by order in perturbation theory. Therefore, we conclude that, at leading order in the critical domain, the correlation functions of the $O(N)$ non-linear model are identical to the correlation functions of the $(\phi^2)^2$ field theory at the IR fixed point. This conclusion is supported by the analysis of the scaling behaviour performed within the $1/N$ expansion (see equation (18.68)). There, the identity between correlation functions of both field theories has been proven (in the sense of a $1/N$ expansion) under only one condition, that the $(\phi^2)^2$ coupling is generic (*i.e.* not too small) in the cut-off scale.

The critical domain: critical exponents. We now determine the behaviour of functions for $t \rightarrow t_c$ (for $N > 2$). The function $\xi(t)$ (equation (19.106)) diverges as

$$\xi(t) \sim \Lambda^{-1} (t_c - t)^{1/\beta'(t_c)}. \quad (19.120)$$

Comparing with the scaling form (16.35), we recover that the correlation length exponent ν is given by $\nu = -1/\beta'(t_c)$ (equation (19.100)). For d close to 2, the exponent ν thus behaves like $\nu \sim 1/\varepsilon$. The function $M_0(t)$ vanishes at t_c :

$$\ln M_0(t) = -\frac{1}{2} \frac{\eta(t_c) + \varepsilon}{\beta'(t_c)} \ln(t_c - t) + \text{const.} \quad (19.121)$$

This yields the exponent β , a result consistent with the expression (19.99) for η through the scaling relation $\beta = \frac{1}{2}\nu(d - 2 + \eta)$. We verify in Section 19.13 the results at two-loop order [201],

$$\frac{1}{\nu} = \varepsilon + \frac{1}{N-2}\varepsilon^2 + O(\varepsilon^3), \quad \eta = \frac{\varepsilon}{N-2} - \frac{N-1}{(N-2)^2}\varepsilon^2 + O(\varepsilon^3). \quad (19.122)$$

Finally, one verifies that the singularity of $\tilde{\Gamma}^{(n)}$ coming from the factor $\xi^{-d}M_0^{-n}$ indeed agrees, near t_c , with the result of equation (16.37).

The nature of the correlation length $\xi(t)$. Equation (19.114) shows that $\xi(t)$ has, in zero field, the nature of a correlation length. For $t < t_c$ fixed, the length scale $\xi(t)$ is of order $1/\Lambda$, which is the microscopic scale. In this regime, large distance physics is governed by massless Goldstone modes (spin wave excitations), and can thus be inferred from the perturbative low temperature expansion, since $t = 0$ is the IR fixed point. However, when t approaches t_c , $\xi(t)$ becomes much larger than the microscopic scale. Then distances exist that are large with respect to the microscopic scale, but small with respect to $\xi(t)$, in which correlation functions have a different universal behaviour, the critical behaviour. The length $\xi(t)$ thus becomes a *crossover scale* between two different continuum behaviours, a long distance perturbative behaviour dominated by Goldstone modes, and a long (but shorter) distance critical behaviour. In this situation, we can construct continuum correlation functions consistent on all scales, the critical behaviour being also the large momentum behaviour in the renormalized field theory.

Remarks. $t > t_c$, N large. Perturbative calculations cannot be performed for $t > t_c$, but RG equations have a continuation that remains valid, at least in the framework of the large N expansion (see Section 18.6.1). Moreover, the large N analysis indicates that RG equations in the form (19.92) are indeed valid for dimensions d larger than 2, but smaller than 4.

General comments. Within the framework of the low temperature expansion, we have been able to describe, for theories with a continuous symmetry, not only the complete structure of the low temperature phase, and this was expected, but, in the non-Abelian case $N > 2$, also the critical behaviour near two dimensions.

What is somewhat surprising in this result is that perturbation series is only sensitive to the local structure of the sphere $\mathbf{S}^2 = 1$, while the restoration of symmetry involves the sphere globally. This explains the peculiarity of the Abelian case $N = 2$, because locally a circle cannot be distinguished from a non-compact straight line. By contrast, for $N > 2$, the sphere has a local characteristic curvature. Still different regular compact manifolds may have the same local metric and, therefore, the same perturbation theory. They all have the same low temperature physics. However, previous results concerning the critical behaviour are physically relevant only if they are still valid when ε is not infinitesimal and t approaches t_c , a condition that cannot be checked directly.

In particular, the low temperature expansion misses, in general, terms decreasing like $\exp(-\text{const.}/t)$ which may, in some cases, be essential for the physics. Finally, we note that, at least, we have found a direct relation between the $(\phi^2)^2$ and the σ -model through the large N expansion (see Section 18.6.1). This gives us some confidence that the previous considerations are valid for the $O(N)$ model, at least for N sufficiently large. On the other hand, the physics for $N = 2$ is not well reproduced (see Chapter 31). Cardy and Hamber have speculated about the RG flow for N close to 2 and dimension d close to 2, incorporating phenomenologically the Kosterlitz–Thouless transition in their analysis.

19.13 Results beyond one-loop order

For explicit calculations of the RG functions beyond one-loop, it is convenient to use the renormalized action (19.84), dimensional regularization and minimal subtraction. As an exercise, we calculate explicitly the RG functions at two-loop order.

Two-loop calculation. The field renormalization constant can be easily derived from the free energy or the σ -field expectation value. We can also obtain both RG functions from the two-point vertex function. We need the one-loop function with the renormalization constants:

$$\tilde{\Gamma}^{(2)}(p) = Z \left[\frac{\mu^\varepsilon}{tZ_t} (p^2 + hZ_h) + [p^2 + \frac{1}{2}(N-1)hZ_h] I(hZ_h) \right] + O(t) \quad (19.123)$$

(we have omitted the subscript r on all quantities). Here (equations (18.16, 19.95)),

$$I(h) = \Omega_d(\sqrt{h}) = N_d \left(\frac{1}{2-d} - \frac{1}{2} \ln h \right) + O(\varepsilon), \quad (19.124)$$

where N_d is the usual loop factor:

$$N_d = \frac{2}{(4\pi)^{d/2} \Gamma(d/2)} = \frac{1}{2\pi} + O(\varepsilon).$$

Expanding Z , Z_t , Z_h , we first determine the renormalization constants at one-loop order. At one-loop order we recover, as expected, the expressions (19.96, 19.97).

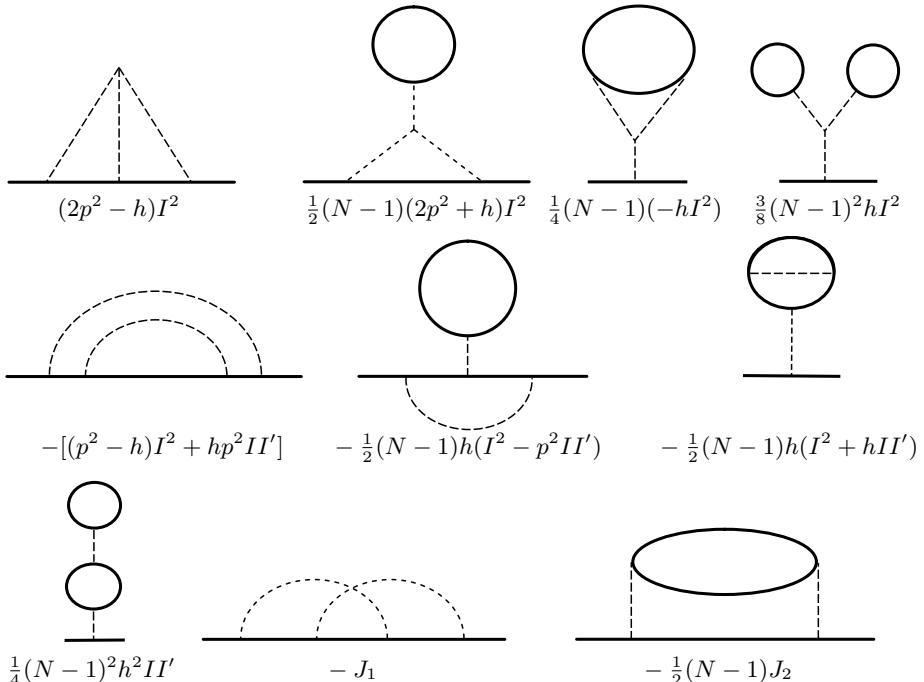


Fig. 19.2 Two-loop diagrams: faithful representation and contribution

Fig. 19.2 displays the two-loop diagrams with their natural weight factors. The notation I' means derivative with respect to h .

We now also need the six-point vertex,

$$V^{(6)} = \frac{1}{16}[(p_1 + p_2)^2 + h]\delta_{i_1 i_2}\delta_{i_3 i_4}\delta_{i_5 i_6}.$$

It generates four faithful two-loop diagrams, displayed in the first line of Fig. 19.2.

We now consider the diagrams involving two four-point vertices. One topology corresponds to the diagrams of line 2 and the first of line 3. Then, there remain two more complicated diagrams of the same topology, J_1, J_2 displayed at the right of line 3:

$$\begin{aligned} J_1 &= \int \frac{d^d q_1 d^d q_2}{(2\pi)^{2d}} \frac{[(p+q_1)^2+h][(p+q_2)^2+h]}{(q_1^2+h)(q_2^2+h)[(p+q_1+q_2)^2+h]}, \\ J_2 &= \int \frac{d^d q_1 d^d q_2}{(2\pi)^{2d}} \frac{[(p+q_1)^2+h]^2}{(q_1^2+h)(q_2^2+h)[(p+q_1+q_2)^2+h]}. \end{aligned}$$

Although these contributions can no longer be reduced to one-loop diagrams, the divergent part can be extracted without too much work. One finds

$$J_1 = 2(1 - 1/d)p^2 I^2(h) + O(1), \quad J_2 = [(4/d - 1)p^2 - h]I^2(h) + O(1).$$

It follows that the sum of the two-loop divergent terms can be written as

$$\frac{N_d^2}{\varepsilon^2} \left[\left(\left(\frac{1}{8}N^2 - \frac{1}{2}N + \frac{3}{8} \right) \varepsilon + \frac{3}{8}N^2 - N + \frac{5}{8} \right) h + \left(\left(-\frac{7}{4} + \frac{3}{4}N \right) \varepsilon - \frac{1}{2} + \frac{1}{2}N \right) p^2 \right].$$

After some simple algebra, in the minimal subtraction scheme, one finds

$$Z = 1 + (N-1)(\tilde{t}/\varepsilon) + (N-1)(N-\frac{3}{2})(\tilde{t}/\varepsilon)^2 + O(\tilde{t}^3). \quad (19.125)$$

$$Z_t = 1 + (N-2)(\tilde{t}/\varepsilon) + (N-2)(N-2+\frac{1}{2}\varepsilon)(\tilde{t}/\varepsilon)^2 + O(\tilde{t}^3), \quad (19.126)$$

where we have introduced the rescaled coupling constant t :

$$\tilde{t} = tN_d. \quad (19.127)$$

The RG functions in the renormalized theory follow from

$$\beta(t) = \varepsilon t \left(1 + t \frac{d \ln Z_t}{dt} \right)^{-1}, \quad \zeta(t) = \beta(t) \frac{d \ln Z}{dt}. \quad (19.128)$$

The calculations can be generalized to some type of long range forces [173].

Higher orders. We now give the two RG functions and the critical exponents at the order presently available [203], that is, four loops, without presenting the details of the calculation. At this order $\tilde{\beta}$, the corresponding β -function, reads

$$\tilde{\beta}(\tilde{t}) = \varepsilon \tilde{t} - (N-2)\tilde{t}^2 \left[1 + \tilde{t} + \frac{1}{4}(N+2)\tilde{t}^2 + b\tilde{t}^3 \right] + O(\tilde{t}^6), \quad (19.129)$$

in which the constant b is,

$$b = -\frac{1}{12}(N^2 - 22N + 34) + \frac{3}{2}\zeta(3)(N-3). \quad (19.130)$$

The numerical constant $\zeta(3)$ is a value of Riemann's ζ function (not be confused with the function $\zeta(t)$ of equation (19.97), $\zeta(3) = 1.2020569\dots$

The anomalous dimensions of the composite operator of spin l (see Section 19.8) is given by

$$\zeta_l(\tilde{t}) = a\tilde{t} \left\{ 1 + (N-2)\tilde{t}^2 \left[\frac{3}{4} + \left(\frac{1}{3}(5-N) + \frac{1}{2}(1+a)\zeta(3) \right) \tilde{t} \right] \right\} + O(\tilde{t}^5), \quad (19.131)$$

with

$$a = -\frac{1}{2}(N+l-2)l. \quad (19.132)$$

The case $l=1$ corresponds to the field itself. The function $\zeta(t)$ given at one-loop order by equation (19.97) is related to ζ_1 by

$$\zeta(t) = -2\zeta_1(t).$$

From these expressions, the values of the critical exponents η and ν follow. In terms of

$$\tilde{\varepsilon} = \varepsilon/(N-2), \quad (19.133)$$

the expansions read

$$\begin{aligned} \eta &= \tilde{\varepsilon} + (N-1)\tilde{\varepsilon}^2 \left\{ -1 + \frac{N}{2}\tilde{\varepsilon} + \left[-b + (N-2) \left(\frac{2-N}{3} + \frac{3-N}{4}\zeta(3) \right) \right] \tilde{\varepsilon}^2 \right\} + O(\tilde{\varepsilon}^5), \\ \nu(d-2) &= 1 - \tilde{\varepsilon} + \frac{(4-N)}{2}\tilde{\varepsilon}^2 + \frac{1}{4} [N^2 - 10N + 18 + 18(3-N)\zeta(3)] \tilde{\varepsilon}^3 + O(\tilde{\varepsilon}^4). \end{aligned}$$

19.14 The dimension 2: Asymptotic freedom

The RG function $\beta(t)$ reduces to

$$\beta(t) = -\frac{(N-2)}{2\pi}t^2 + O(t^3). \quad (19.134)$$

The non-linear σ -model for $N > 2$ is the simplest example of a so-called *asymptotically free* field theory (UV free), since the first coefficient of the β -function is negative, in contrast with the ϕ^4 field theory. Therefore, in the whole range where the function β remains negative, the effective interaction decreases at short distance, as equation (19.102a), for $\lambda \rightarrow \infty$, shows. The large momentum behaviour of correlation functions is then entirely calculable from perturbation theory and RG arguments.

By contrast, the theory is IR unstable and thus, in zero field h , the spectrum of the theory is not perturbative. Contrary to perturbative indications, it consists of N massive degenerate states, since the $O(N)$ symmetry is not broken. Asymptotic freedom and the non-perturbative character of the spectrum are also properties of quantum chromodynamics in four dimensions, the theory that describes strong interactions in the Standard Model of particle physics (see Chapters 23 and 24).

The definition (19.106) of the correlation length $\xi(t)$ (or inverse physical mass $m(t)$), in zero field, has to be replaced by

$$\xi(t) = \Lambda^{-1} \exp \left[\int_{t_0}^t \frac{dt'}{\beta(t')} \right], \quad (19.135)$$

where t_0 corresponds to the cut-off scale. Due to the negative sign of the β function, the correlation length becomes large only for $t \rightarrow 0$. Using the small t expansion of the β function, given by equation (19.129) (up to order t^3 , the bare and renormalized β function have the same expansion, see Section 9.10.1), one can evaluate the integral and finds

$$\xi^{-1}(t) = m(t) \propto \Lambda t^{-1/(N-2)} e^{-2\pi/[(N-2)t]} (1 + O(t)), \quad (19.136)$$

the proportionality factor depending on an explicit microscopic model.

Finally, the scaling forms (19.114, 19.116) imply that the perturbative expansion at fixed magnetic field is valid, at low momenta or large distances, and for $h/h_0(t)$ large.

20 Gross–Neveu–Yukawa and Gross–Neveu models

In this chapter, as an instructive exercise, we first consider a model (a simplified form of the Nambu–Jona-Lasinio model) that can be defined in continuous dimensions, the Gross–Neveu–Yukawa (GNY) model [204], which involves N Dirac fermions and one scalar field. The model has a continuous $U(N)$ symmetry, and a discrete chiral symmetry that prevents the addition of a fermion mass term to the action. For a specific value of a coefficient of the action, the model undergoes a continuous phase transition. It illustrates the mechanism of spontaneous symmetry breaking, leading to *spontaneous fermion mass generation*, like in the Standard Model of particle physics.

Therefore, in four dimensions, the GNY can be considered as a toy model to represent the interactions between the top quark and the Higgs boson, the heaviest particles of the Standard Model of fundamental interactions, when the gauge fields are omitted.

The model is renormalizable in four dimensions, and its renormalization group (RG) properties can be studied in $d = 4$ and $d = 4 - \varepsilon$ dimensions. At fixed dimension, for $d < 4$, the perturbative expansion at the transition is infrared (IR) divergent.

A model of self-interacting fermions with the same symmetries and fermion content, the Gross–Neveu (GN) model [178, 205], has been widely studied. In perturbation theory, for $d > 2$, it describes only a phase with massless fermions but, in $d = 2 + \varepsilon$ dimensions, RG indicates that, at a critical value of the coupling constant, the model experiences a phase transition. In two dimensions, it is renormalizable and exhibits the phenomenon of *asymptotic freedom* (see Sections 23.4.1 and 24.4). The massless phase becomes infrared (IR) unstable and the spectrum non-perturbative. There is strong evidence that the phase corresponds to spontaneous symmetry breaking and fermion mass generation.

The large N expansion then shows that the GN and GNY models have the same large scale properties to all orders in $1/N$. Therefore, the two models can be studied by methods similar to those already employed for the $(\phi^2)^2$ field theory and the non-linear σ -model, based on RG equations near four and two dimensions, and large N expansion.

20.1 The GNY model: Spontaneous mass generation

We consider the GNY model, a field theory with one scalar field $\sigma(x)$ and N massless Dirac fermion fields $\psi^i(x)$ and $\bar{\psi}^i(x)$, and the Euclidean action in d dimensions,

$$\mathcal{S}(\bar{\psi}, \psi, \sigma) = \int d^d x \left[-\bar{\psi}(x) \cdot (\not{\partial} + g\sigma(x)) \psi(x) + \frac{1}{2} (\nabla\sigma(x))^2 + V(\sigma(x)) \right], \quad (20.1)$$

with

$$V(\sigma) = \frac{1}{2} r\sigma^2 + \frac{\lambda}{4!}\sigma^4, \quad \lambda > 0.$$

It has a $U(N)$ fermion symmetry corresponding to the transformations,

$$\psi(x) \mapsto \mathbf{U}\psi(x), \quad \bar{\psi}(x) \mapsto \bar{\psi}(x)\mathbf{U}^\dagger, \text{ with } \mathbf{U}\mathbf{U}^\dagger = \mathbf{1},$$

and, in addition, in all dimensions, a discrete symmetry which forbids a fermion mass term and odd powers of σ in σ self-interactions.

In even dimensions, the symmetry can be identified with the discrete chiral \mathbb{Z}_2 symmetry (γ_S is defined by equation (12.3), for $d = 4$, $\gamma_S \equiv \gamma_5$)

$$\psi(x) \mapsto \gamma_S \psi(x), \quad \bar{\psi}(x) \mapsto -\bar{\psi}(x) \gamma_S, \quad \sigma(x) \mapsto -\sigma(x). \quad (20.2)$$

In odd dimensions, a fermion mass term breaks space parity. In generic dimensions, the discrete symmetry can be expressed as $\sigma(x) \mapsto -\sigma(x)$ and

$$\mathbf{x} = \{x_1, x_2, \dots, x_d\} \mapsto \tilde{\mathbf{x}} = \{-x_1, x_2, \dots, x_d\}, \quad \begin{cases} \psi(x) \mapsto \gamma_1 \psi(\tilde{x}), \\ \bar{\psi}(x) \mapsto -\bar{\psi}(\tilde{x}) \gamma_1. \end{cases} \quad (20.3)$$

The phase transition: Tree order analysis. At leading order, the system undergoes a phase transition at $r = r_c = 0$. For $r = m^2 > r_c$, the vacuum expectation value $\langle \sigma \rangle$ of σ vanishes, the fermions remain massless and the \mathbb{Z}_2 symmetry remains unbroken. By contrast, for $r < r_c$, the minimum of the σ potential is given by

$$V'(\sigma) = 0 \Rightarrow \langle \sigma \rangle = \pm \sqrt{-6r/\lambda} \Rightarrow m_\sigma^2 = V''(\sigma) = -2r > 0,$$

which shows that the \mathbb{Z}_2 symmetry is *spontaneously broken*. This does not generate Goldstone bosons, because the symmetry is discrete. Then, returning to the action (20.1), replacing the field σ by its expectation value, we note that the *fermions become massive* with a mass (g can always be chosen positive)

$$m_\psi = g |\langle \sigma \rangle| = g \sqrt{-6r/\lambda}, \quad \text{and thus} \quad \frac{m_\psi}{m_\sigma} = g \sqrt{3/\lambda}. \quad (20.4)$$

The physics of the model is characterized by *spontaneous \mathbb{Z}_2 symmetry breaking* and *fermion mass generation*.

The σ expectation value gives a mass to the fermions, a mechanism reminiscent of the Standard Model of weak-electromagnetic interactions (see Section 23.1), where all fundamental particles become massive through the Higgs-field expectation value.

Gaussian renormalization. We assume some momentum regularization with a cut-off Λ , inverse of the microscopic scale. The Gaussian (or ultraviolet (UV)) dimensions $[\psi]$ for fermions, given by equation (12.53), and $[\sigma]$ by equation (8.7) for the scalar field, are

$$[\psi] = [\bar{\psi}] = \frac{1}{2}(d-1), \quad [\sigma] = \frac{1}{2}(d-2).$$

We perform a Gaussian renormalization in such a way that the coupling constants become dimensionless. This leads to

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi, \sigma) = & \int d^d x \left[-\bar{\psi}(x) \cdot \left(\not{\partial} + g \Lambda^{(4-d)/2} \sigma(x) \right) \psi(x) \right. \\ & \left. + \frac{1}{2} (\nabla \sigma(x))^2 + \frac{1}{2} r \sigma^2(x) + \Lambda^{4-d} \lambda \sigma^4(x) \right]. \end{aligned} \quad (20.5)$$

We have not rescaled r , since anyway r has to be *fine tuned* to stay close to the critical value r_c , like in the $\phi_{d=4}^4$ field theory (see Section 9.11). The value r_c is not universal, but modified by interactions.

20.1.1 Renormalization and RG near four dimensions

To study the model beyond the tree approximation, we discuss its renormalization and derive RG equations at and near four dimensions.

In four dimensions, the Gaussian dimensions of the fields take the values

$$[\sigma] = 1, \quad [\psi] = [\bar{\psi}] = 3/2,$$

and the GNY model is renormalizable by power counting.

Dimensional regularization and minimal subtraction. Because, at leading order, the RG functions are identical in the microscopic (or bare) theory, corresponding to the action (20.5), and in the renormalized field theory defined by dimensional regularization and minimal subtraction (Sections 10.2.1 and 10.4), we choose the latter framework for practical calculations. Since we calculate only the RG functions, we restrict the calculations to the critical (massless) theory $r = r_c$. In dimensional regularization, the *fine-tuning problem is hidden*, because dimensional regularization automatically performs a partial renormalization, which here implies $r_c = 0$ to all orders.

The critical renormalized action, in a form suitable for dimensional regularization, can thus be written as ($\varepsilon = 4 - d$),

$$\begin{aligned} S_r(\bar{\psi}, \psi, \sigma) = & \int d^d x \left\{ -\bar{\psi}(x) \left[Z_\psi \partial + g_r Z_g \mu^{\varepsilon/2} \sigma(x) \right] \psi(x) + \frac{1}{2} Z_\sigma (\nabla \sigma(x))^2 \right. \\ & \left. + \frac{1}{4!} \mu^\varepsilon \lambda_r Z_\lambda \sigma^4(x) \right\}, \end{aligned} \quad (20.6)$$

where ψ, σ are renormalized fields, μ the physical mass scale, g_r and λ_r dimensionless renormalized coupling constants, and Z_ψ, Z_σ, Z_g , and Z_ψ renormalization constants. The relation between bare (or microscopic) and renormalized coupling constants, denoted g, λ and g_r, λ_r , respectively, then is

$$g = \mu^{\varepsilon/2} g_r Z_g / (Z_\psi Z_\sigma^{1/2}), \quad \lambda = \mu^\varepsilon \lambda_r Z_\lambda / Z_\sigma^2. \quad (20.7)$$

Notation. In the following, we denote by $\tilde{\Gamma}^{(l,n)}$ the vertex functions corresponding to l fermion fields and n scalar fields, in the Fourier representation.

20.1.2 One-loop calculations

In the calculations, we fix $\text{tr } \mathbf{1} = 4$ in the space of γ matrices.

We have given in Section 12.5 (equation (12.50)) the generating functional of vertex functions $\Gamma(\sigma, \psi, \bar{\psi})$ at one-loop order for an action that contains the action (20.1) for $N = 1$, as a special example. Here, one finds

$$\begin{aligned} \Gamma_{1\text{ loop}} = & -N \text{tr} \ln \left[1 + g_r \mu^{\varepsilon/2} \partial^{-1} \sigma(x) \right] + \frac{1}{2} \text{tr} \ln \left[1 - \frac{1}{2} \lambda_r \mu^\varepsilon (\nabla^2)^{-1} \sigma^2(x) \right. \\ & \left. - 2g_r^2 \mu^\varepsilon (\nabla^2)^{-1} \bar{\psi}(x) \cdot \left(\partial + g_r \mu^{\varepsilon/2} \sigma(x) \right)^{-1} \psi(x) \right]. \end{aligned} \quad (20.8)$$

An expansion in powers of σ, ψ , and $\bar{\psi}$ generates all vertex functions at one-loop order. We omit all contributions involving only the σ^4 interaction, which have already been evaluated in Section 10.4.

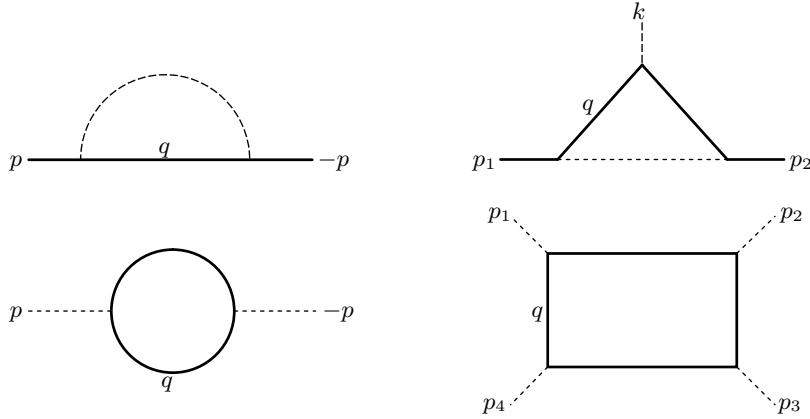


Fig. 20.1 Feynman diagrams: fermions and bosons correspond to continuous and dotted lines, respectively

Then, the renormalized vertex functions are given by

$$\tilde{\Gamma}_r^{(2,0)}(p) = iZ_\psi \not{p} + \frac{g_r^2 \mu^\varepsilon}{(2\pi)^d} \int \frac{d^d q}{(p-q)^2} \frac{iq}{q^2}, \quad (20.9)$$

$$\tilde{\Gamma}_r^{(2,1)}(p_1, p_2; k) = g_r Z_g \mu^{\varepsilon/2} - \frac{g_r^3 \mu^{3\varepsilon/2}}{(2\pi)^d} \int \frac{d^d q}{(q+p_2)^2} \frac{q(q-k)}{q^2(q-k)^2}, \quad (20.10)$$

$$\tilde{\Gamma}_r^{(0,2)}(p) \Big|_{\lambda=0} = p^2 Z_\sigma - \frac{N g_r^2 \mu^\varepsilon}{(2\pi)^d} \int d^d q \frac{\text{tr } q(q+\not{p})}{q^2(p+q)^2}, \quad (20.11)$$

$$\begin{aligned} \tilde{\Gamma}_r^{(0,4)}(p_1, p_2, p_3, p_4) \Big|_{\lambda=0} &= N g_r^4 \mu^{2\varepsilon} \int \frac{d^d q}{(2\pi)^d} \frac{\text{tr } [q(q+\not{p}_1)(q+\not{p}_1+\not{p}_2)(q-\not{p}_4)]}{q^2(q+p_1)^2(q+p_1+p_2)^2(q-p_4)^2} \\ &+ 5 \text{ diagrams corresponding to permutations of } \{p_2, p_3, p_4\}. \end{aligned} \quad (20.12)$$

Fig. 20.1 displays the corresponding Feynman diagrams.

We need the function (10.52),

$$B_d(p) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{q^2(p+q)^2},$$

calculated in Section 10.5 (equation (10.53)),

$$B_d(p) = -\frac{N_d}{2} \frac{\pi}{\sin(\pi d/2)} \frac{\Gamma(d/2-1)\Gamma(d/2)}{\Gamma(d-2)} p^{d-4} = \frac{N_d}{\varepsilon} \left(1 + \left(\frac{1}{2} - \ln p\right)\varepsilon\right) + O(\varepsilon), \quad (20.13)$$

where N_d is the loop factor (10.12), $N_d = 2/(4\pi)^{d/2}\Gamma(d/2)$.

20.1.3 The two-point functions

We first calculate completely the contributions to the two-point vertex functions.

Fermions. We need the integral

$$I_1(p) = \frac{1}{(2\pi)^d} \int \frac{d^d q}{(p-q)^2} \frac{iq}{q^2}. \quad (20.14)$$

Rotational invariance implies that $I_1(p)$ is proportional to \not{p} and, therefore,

$$I_1(p) = i\not{p}A(p), \quad \text{with} \quad A(p) = -\frac{i}{4p^2} \text{tr } \not{p}I_1(p) = \frac{1}{p^2} \frac{1}{(2\pi)^d} \int \frac{d^d q p \cdot q}{(p-q)^2 q^2}.$$

Thus, no fermion mass is generated. Then, using

$$p \cdot q = -\frac{1}{2}(p-q)^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2,$$

we decompose $A(p)$ into the sum of three integrals. In dimensional regularization the first and the last integral vanish. Thus, from equation (20.13),

$$A(p) = \frac{1}{2} \frac{1}{(2\pi)^d} \int \frac{d^d q}{(p-q)^2 q^2} = \frac{1}{2} B_d(p).$$

One infers

$$I_1(p) = -i\not{p}p^{d-4} N_d \frac{\pi}{4 \sin(d/2)} \frac{\Gamma(d/2)\Gamma(d/2-1)}{\Gamma(d-2)}. \quad (20.15)$$

Expanding at first order in $\varepsilon = 4 - d$, one obtains

$$\tilde{\Gamma}_r^{2,0}(p) = i\not{p}Z_\psi + g_r^2 \mu^\varepsilon I_1(p) = i\not{p} \left[1 + \frac{g_r^2}{16\pi^2} \left(\frac{1}{2} - \ln(p/\mu) \right) \right] + O(\varepsilon), \quad (20.16)$$

with

$$Z_\psi = 1 - g_r^2 \frac{N_d}{2\varepsilon} + O(\text{two loops}). \quad (20.17)$$

Bosons. We need the integral

$$I_2(p) = -\frac{N}{(2\pi)^d} \int d^d q \text{ tr } \frac{\not{q}(\not{q} + i\not{p})}{q^2(p+q)^2}. \quad (20.18)$$

After evaluation of the trace ($\text{tr } \mathbf{1} = 4$), the integral becomes

$$I_2(p) = -\frac{4N}{(2\pi)^d} \int d^d q \frac{q^2 + p \cdot q}{q^2(p+q)^2}. \quad (20.19)$$

Then, using

$$q^2 + p \cdot q = \frac{1}{2} [q^2 + (p+q)^2] - \frac{1}{2}p^2,$$

we decompose $I_2(p)$ into a sum of three integrals. Two vanish in dimensional regularization and, therefore,

$$I_2(p) = \frac{4N}{(2\pi)^d} \frac{p^2}{2} \int \frac{d^d q}{q^2(p+q)^2} = 2Np^2 B_d(p). \quad (20.20)$$

One finds

$$\Gamma_r^{(0,2)}(p) = Z_\sigma p^2 + \mu^\varepsilon g_r^2 I_2(p) = p^2 + \frac{N}{4\pi^2} g_r^2 p^2 \left[\frac{1}{2} - \ln(p/\mu) \right] + O(\varepsilon),$$

with

$$Z_\sigma = 1 - 2N \frac{N_d}{\varepsilon} g_r^2 + O(\text{two loops}). \quad (20.21)$$

20.1.4 Three and four-point functions

All one-loop diagrams can be calculated by similar methods. We evaluate here only their divergent part.

For the vertex function $\tilde{\Gamma}_r^{(2,1)}(p_1, p_2; k)$, it is sufficient to calculate the diagram for $k = 0$. Thus, we need

$$I_3(p_2) = - \int \frac{d^d q}{(2\pi)^d} \frac{(q)^2}{(q + p_2)^2 (q^2)^2} = -B_d(p_2)$$

and, therefore,

$$Z_g = 1 + g_r^2 \frac{N_d}{\varepsilon}. \quad (20.22)$$

Similarly, for the σ four-point function $\tilde{\Gamma}_r^{(0,4)}$, we need the diagram only for $p_1 = p_3 = 0$ and thus $p_4 = -p_2$. This reduces the calculation to

$$I_4(p_2) = \frac{24N}{(2\pi)^d} \int \frac{d^d q}{q^2 (q + p_2)^2} = 24N B_d(p_2).$$

Adding the pure boson contribution (see equation (10.59)), one obtains

$$\lambda_r Z_\lambda = \lambda_r + \left(\frac{3}{2} \lambda_r^2 - 24N g_r^4 \right) \frac{N_d}{\varepsilon}. \quad (20.23)$$

Bare coupling constants. Using the expressions (20.7), we infer the expansions of g and λ :

$$g^2 = \mu^\varepsilon \frac{g_r^2 Z_g^2}{Z_\psi^2 Z_\sigma} = g_r^2 \left(1 + (2N + 3) g_r^2 \frac{N_d}{\varepsilon} \right) + O(2 \text{ loops}), \quad (20.24)$$

$$\lambda = \mu^\varepsilon \frac{\lambda_r Z_\lambda}{Z_\sigma^2} = \mu^\varepsilon \left[\lambda_r + \left(\frac{3}{2} \lambda_r^2 + 4N \lambda_r g_r^2 - 24N g_r^4 \right) \frac{N_d}{\varepsilon} \right] + O(2 \text{ loops}). \quad (20.25)$$

20.2 RG equations near four dimensions

The model (20.1) is trivial above four dimensions, renormalizable in four dimensions and can thus be studied near dimension 4 by RG techniques. For the critical theory, four renormalization constants are required, corresponding to the two field renormalizations and the two coupling constants. The RG equations thus involve four RG functions. The renormalized vertex functions $\tilde{\Gamma}_r^{(l,n)}$, for l fermion and n scalar fields, then satisfy

$$\left(\mu \frac{\partial}{\partial \mu} + \beta_{g^2} \frac{\partial}{\partial g_r^2} + \beta_\lambda \frac{\partial}{\partial \lambda_r} - \frac{1}{2} l \eta_\psi - \frac{1}{2} n \eta_\sigma \right) \tilde{\Gamma}_r^{(l,n)}(p_i, \mu, g_r, \lambda_r) = 0. \quad (20.26)$$

20.2.1 Renormalization group functions

In d dimensions, bare and renormalized coupling constants are related by an equation of the form

$$\lambda_i = \mu^\varepsilon f_i(\lambda_r), \quad (20.27)$$

in which λ_i represents symbolically the bare coupling constants (g^2, λ) and λ_i the renormalized coupling constants (g_r^2, λ_r) .

The coupling constant RG functions $\beta_i(\lambda_r)$ are then given by

$$\beta_i(\lambda_r) = \mu \frac{\partial}{\partial \mu} \Big|_\lambda \lambda_{r,i}, \quad (20.28)$$

and thus, using equation (20.27),

$$0 = \mu \frac{\partial}{\partial \mu} \lambda_i \Big|_\lambda = \mu^\varepsilon \left(\varepsilon f_i(\lambda_r) + \sum_j \beta_j(\lambda_r) \frac{\partial f_i}{\partial \lambda_{r,j}} \right). \quad (20.29)$$

A short calculation yields

$$\beta_{g^2} = -\varepsilon g_r^2 + N_d(2N+3)g_r^4 + O(\text{two loops}), \quad (20.30)$$

$$\beta_\lambda = -\varepsilon \lambda_r + N_d \left(\frac{3}{2} \lambda_r^2 + 4N \lambda_r g_r^2 - 24N g_r^4 \right) + O(\text{two loops}). \quad (20.31)$$

The field renormalization RG functions at the same order are:

$$\eta_\sigma = 2NN_d g_r^2, \quad \eta_\psi = \frac{1}{2} N_d g_r^2. \quad (20.32)$$

20.2.2 Bare RG equations: Triviality and mass ratio

The microscopic (or bare) vertex functions at scale Λ , generated by the action (20.5), satisfy RG equations formally analogous to equations (20.26),

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta_{g^2} \frac{\partial}{\partial g^2} + \beta_\lambda \frac{\partial}{\partial \lambda} - \frac{1}{2} l \eta_\psi - \frac{1}{2} n \eta_\sigma \right) \tilde{\Gamma}^{(l,n)}(p_i, \Lambda, g, \lambda) = 0, \quad (20.33)$$

where the RG functions are different, but coincide at one-loop order:

$$\beta_{g^2} = -\varepsilon g^2 + N_d(2N+3)g^4 + O(\text{two loops}), \quad (20.34)$$

$$\beta_\lambda = -\varepsilon \lambda + N_d \left(\frac{3}{2} \lambda^2 + 4N \lambda g^2 - 24N g^4 \right) + O(\text{two loops}). \quad (20.35)$$

The effective couplings at a scale Λe^{-t} satisfy the flow equations

$$\frac{dg^2}{dt} = -\beta_{g^2}(g^2, \lambda), \quad \frac{d\lambda}{dt} = -\beta_\lambda(g^2, \lambda).$$

In four dimensions ($\varepsilon = 0$), the origin $\lambda = g^2 = 0$ is IR stable. Indeed, if the initial couplings at scale Λ are such that the expansions (20.34, 20.35) are, at least qualitatively usable, then equation (20.34) implies that, for $t \rightarrow +\infty$, $g(t)$ goes to 0, and then equation (20.35) implies that $\lambda(t)$ also goes to 0. As a consequence, at fixed physical scale $\mu = e^{-t} \Lambda$, if one insists taking the infinite cut-off limit, then the effective couplings go to 0.

This is called the *triviality issue* (see Section 9.12).

In the spirit of effective field theories, one only demands $\mu \ll \Lambda$. Then, solving the flow equations, one finds asymptotically,

$$g^2(\Lambda/\mu) \sim \frac{8\pi^2}{(2N+3)\ln(\Lambda/\mu)}, \quad \lambda(\Lambda/\mu) \sim \frac{8\pi^2 R_N}{3(2N+3)\ln(\Lambda/\mu)},$$

with

$$R_N = -(2N-3) + \sqrt{4N^2 + 132N + 9}. \quad (20.36)$$

The effective coupling constants at physical scale (the renormalized coupling constants) are expected to be small, and asymptotically independent from the initial bare couplings.

Mass ratio. Since the effective couplings at physical scale are small, perturbation theory can be used to calculate observables. For example, one can evaluate the ratio between the masses of the scalar and fermion fields in the broken phase, using equation (20.4). One obtains

$$\frac{m_\sigma^2}{m_\psi^2} \sim \frac{\lambda(\mu/\Lambda)}{3g^2(\mu/\Lambda)} = \frac{1}{9}R_N, \quad (20.37)$$

which is a fixed number, independent of the values of the initial couplings at the microscopic scale, while in the tree approximation the ratio is arbitrary.

Of course, if the ratio Λ/μ is not extremely large, or the bare couplings are ‘unnaturally’ small, the renormalized couplings at scale μ may be different, and the ratio may be modified.

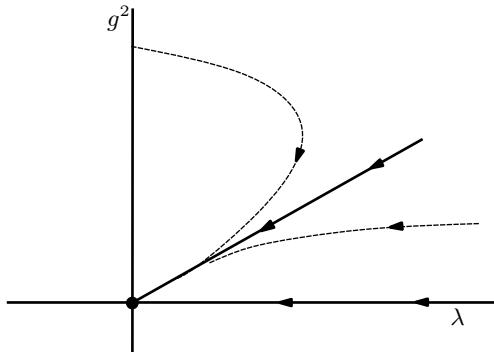


Fig. 20.2 RG flow from cut-off scale to physical scale: full lines correspond to fixed lines

20.2.3 The general RG flow at one-loop order, in four dimensions

Two-dimensional flows can be easily studied, because, quite generally, RG trajectories can only meet at fixed points, here $g = \lambda = 0$. Setting $t = \ln(\Lambda/\mu) + t_0$, with $t_0 > 0$, and

$$K_N = \frac{\sqrt{4N^2 + 132N + 9}}{(2N+3)}, \quad (20.38)$$

one can write the solutions of the one-loop differential equations explicitly as

$$g^2(t) = \frac{8\pi^2}{(2N+3)t},$$

$$\frac{\lambda(t)}{g^2(t)} = 1 - \frac{2N}{3} + \frac{(2N+3)K_N}{3} \frac{1 + \theta e^{-K_N \ln(t)}}{1 - \theta e^{-K_N \ln(t)}},$$

where t_0 and θ are integration constants. One fixed line corresponds to $\theta = 0$ and then, $\lambda = R_N g^2/3$, and another one to $g = 0$ (recovered for $t_0 \rightarrow +\infty$).

These fixed RG trajectories cannot be crossed. By contrast, for $\theta > 0$, the line $\lambda = 0$ can be formally crossed, and the RG trajectories then enter a non-physical region (see Fig. 20.2). Close to the line, in the framework of effective field theories, new interaction terms like dimension 6 operators, which have been neglected in the RG analysis, have to be taken into account and new physics is required. Generic situations correspond to $g^2, \lambda = O(1)$, and the RG trajectories converge towards the fixed line $\lambda = R_N g^2/3$ for $\Lambda/\mu \gg 1$.

20.2.4 Dimension $d = 4 - \varepsilon$

In $d = 4 - \varepsilon$ dimension, for $\varepsilon > 0$ small, one finds a non-trivial fixed point. At order ε ,

$$g_*^2 = \frac{8\pi^2}{2N+3}\varepsilon, \quad \lambda_* = 8\pi^2 \frac{R_N}{3(2N+3)}\varepsilon. \quad (20.39)$$

The matrix of derivatives of the β -functions has two eigenvalues ω_1, ω_2 , (K_N is defined by equation (20.38))

$$\omega_1 = \varepsilon, \quad \omega_2 = 2K_N\varepsilon, \quad (20.40)$$

and thus the fixed point is IR stable. The first eigenvalue is always the smallest.

Using the expressions (20.32), which give

$$\eta_\psi = \frac{2}{2N+3}\varepsilon, \quad \eta_\sigma = \frac{2N}{2N+3}\varepsilon,$$

one obtains the dimensions d_ψ and d_σ of the fields

$$d_\psi = \frac{3}{2} - \frac{N+1}{(2N+3)}\varepsilon, \quad d_\sigma = 1 - \frac{3}{2(2N+3)}\varepsilon. \quad (20.41)$$

A one-loop order calculation of the $\sigma^2\sigma\sigma$ vertex function yields the RG function η_r corresponding to the mass operator,

$$\eta_r = -\frac{\lambda}{16\pi^2} - \eta_\sigma,$$

and thus the exponent ν ,

$$\frac{1}{\nu} = 2 + \eta_r = 2 - \frac{R_N\varepsilon}{6(2N+3)} - \frac{2N\varepsilon}{2N+3} = 2 - \frac{\varepsilon}{6} \left(\frac{10N+3}{2N+3} + K_N \right). \quad (20.42)$$

Finally, the ratio of masses (20.4) at the fixed point keeps its four-dimensional value (20.37) at leading order.

For $d = 4$ and $d = 4 - \varepsilon$, the existence of an IR-stable fixed point has the same consequence: if one assumes that the σ expectation value is much smaller than the cut-off, and that the coupling constants are generic at the cut-off scale, then *the ratio of fermion and scalar masses is fixed*.

20.2.5 RG functions beyond one-loop order

The RG functions are known up to three-loop order [206]. For example, the contributions, at two-loop order, to the RG β functions of the renormalized theory in the $\overline{\text{MS}}$ scheme are,

$$\begin{aligned}\beta_{g^2} &: N_d^2 \left[\frac{1}{24} \lambda_r^2 g_r^2 - \lambda_r g_r^4 - \left(6N + \frac{9}{8} \right) g_r^6 \right], \\ \beta_\lambda &: N_d^2 \left(96N g_r^6 + 7N \lambda_r g_r^4 - 3N g_r^2 \lambda_r^2 - \frac{17}{12} \lambda_r^3 \right).\end{aligned}$$

At the same order,

$$\begin{aligned}\eta_\psi &= \frac{1}{2} N_d g_r^2 - \frac{1}{16} N_d^2 (1 + 12N) g_r^4 \\ \eta_\sigma &= 2NN_d g_r^2 + N_d^2 \left(\frac{1}{24} \lambda_r^2 - \frac{5}{2} g_r^4 \right).\end{aligned}$$

20.3 The GNY model in the large N limit

Notation. In the sections that follow, we denote by \tilde{N} the total number of fermion components:

$$\tilde{N} = N \operatorname{tr} \mathbf{1}, \quad (20.43)$$

where $\mathbf{1}$ is the unit matrix in the representation of γ matrices (*e.g.* $\tilde{N} = 4N$ for $d = 4$).

We assume a momentum regularization and denote the momentum cut-off, which is a representative of the inverse microscopic scale, by Λ .

We start from the action (20.5) of the GNY model. We then integrate over $(N - 1)$ fermion fields. For convenience, we rescale $\Lambda^{(4-d)/2} g\sigma$ into σ , and then find the large N action,

$$\begin{aligned}\mathcal{S}_N(\bar{\psi}, \psi, \sigma) &= \int d^d x \left\{ -\bar{\psi}(x)(\not{\partial} + \sigma(x))\psi(x) \right. \\ &\quad \left. + \frac{\Lambda^{d-4}}{2g^2} \left[(\nabla\sigma(x))^2 + r\sigma^2(x) + \frac{\lambda}{12g^2}\sigma^4(x) \right] \right\} - (N - 1) \operatorname{tr} \ln(\not{\partial} + \sigma). \quad (20.44)\end{aligned}$$

To take the large N limit, we assume σ finite and $g^2, \lambda = O(1/N)$.

The corresponding action density $\mathcal{E}(M)$ for constant field $\sigma(x) = M$ and vanishing fermion fields reads,

$$\begin{aligned}\mathcal{E}(M) &= \Lambda^{d-4} \left(\frac{r}{2g^2} M^2 + \frac{\lambda}{4!g^4} M^4 \right) - N \operatorname{tr} \ln(\not{\partial} + M) \\ &= \Lambda^{d-4} \left(\frac{r}{2g^2} M^2 + \frac{\lambda}{4!g^4} M^4 \right) - \frac{\tilde{N}}{2} \int^\Lambda \frac{d^d q}{(2\pi)^d} \ln(q^2 + M^2).\end{aligned} \quad (20.45)$$

The expectation value of σ for N large is given by the *gap* equation:

$$\mathcal{E}'(M)\Lambda^{4-d} = \frac{r}{g^2} M + \frac{\lambda}{6g^4} M^3 - \tilde{N}\Lambda^{4-d} M \Omega_d(M) = 0, \quad (20.46)$$

where Ω_d is a one-loop diagram (see equations (18.16) and (18.28)), which has the expansion,

$$\begin{aligned}\Omega_d(M) &= \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d p}{p^2 + M^2} \underset{d>2}{=} \Omega_d(0) - C(d) M^{d-2} + a(d) M^2 \Lambda^{d-4} \\ &\quad + O(M^d/\Lambda^2, M^4 \Lambda^{d-6}).\end{aligned} \quad (20.47)$$

The second derivative is

$$\mathcal{E}''(M)\Lambda^{4-d} = \frac{r}{g^2} + \frac{\lambda}{2g^4}M^2 + \tilde{N}\Lambda^{4-d} \int^\Lambda \frac{d^d q}{(2\pi)^d} \frac{M^2 - q^2}{(q^2 + M^2)^2}.$$

The solution $M = 0$ is stable provided

$$\mathcal{E}''(0) > 0 \Leftrightarrow r/g^2 > \tilde{N}\Lambda^{4-d}\Omega_d(0).$$

By contrast, the non-trivial solution to the gap equation exists only for

$$r/g^2 > \tilde{N}\Lambda^{4-d}\Omega_d(0),$$

but then it is stable. We conclude that the critical temperature or critical value r_c of r is given by

$$r_c = g^2\tilde{N}\Lambda^{4-d}\Omega_d(0), \quad (20.48)$$

which shows that the fermions favour the chiral transition. In particular, when d approaches 2, we observe that $r_c \rightarrow +\infty$, which implies that the chiral symmetry is always broken in two dimensions. Using equation (20.48), we set

$$\tau = \Lambda^{d-4}(r - r_c)/g^2, \quad (20.49)$$

where, in the statistical framework, τ plays the role of the deviation from the critical temperature. The equation for the non-trivial solution can be written as

$$\tau + \Lambda^{d-4}\lambda M^2/6g^4 - \tilde{N}(\Omega_d(M) - \Omega_d(0)) = 0.$$

For $d < 4$, keeping only the leading terms for $\tau \rightarrow 0$, we obtain the scaling behaviour,

$$M \sim (-\tau/\tilde{N}C)^{1/(d-2)}. \quad (20.50)$$

At leading order, since the fermion mass $m_\psi = M$, it immediately follows that the exponent ν is also given by

$$\nu \sim \beta \sim 1/(d-2) \Rightarrow \eta_\sigma = 4-d. \quad (20.51)$$

For $N \rightarrow \infty$, at leading order, ν has the same value as in the non-linear σ -model.

At leading order in the scaling limit, the thermodynamic potential $\mathcal{E}(M)$ (or action density) then becomes

$$\Lambda^{4-d}\mathcal{E}(M) = \frac{1}{2}\tau M^2 + (\tilde{N}/d)C(d)|M|^d. \quad (20.52)$$

Although, in terms of the σ -field, the model has a simple Ising-like symmetry, the scaling equation of state for large N is quite different.

From the large N action, one concludes that, at this order, $\eta_\psi = 0$. One also infers the σ -propagator at leading order. Quite generally, using the saddle point equation, one finds for the inverse σ -propagator in the massive phase, in the Fourier representation,

$$\begin{aligned} \tilde{\Delta}_\sigma^{-1}(p) &= \Lambda^{d-4} \left(\frac{p^2}{g^2} + \frac{\lambda}{3g^4}M^2 \right) \\ &+ \frac{\tilde{N}}{2(2\pi)^d} (p^2 + 4M^2) \int^\Lambda \frac{d^d q}{(q^2 + M^2)[(p+q)^2 + M^2]}. \end{aligned} \quad (20.53)$$

In the scaling limit $p, M \rightarrow 0$, the integral yields the leading contribution. Neglecting corrections to scaling, one finds that the propagator vanishes for $p^2 = -4M^2$, which is the $\bar{\psi}\psi$ threshold. Thus, in this limit, $m_\sigma = 2m_\psi$ in all dimensions.

At the transition, the σ propagator reduces to

$$\tilde{\Delta}_\sigma \sim \frac{2}{\tilde{N}b(d)p^{d-2}}, \quad (20.54)$$

with (equation (18.50))

$$b(d) = -\frac{\pi}{\sin(\pi d/2)} \frac{\Gamma^2(d/2)}{\Gamma(d-1)} N_d. \quad (20.55)$$

The result is consistent with the value of η_σ in equation (20.51).

The behaviour of the propagator at the critical point, $\tilde{\Delta}_\sigma(p) \propto p^{2-d}$ implies for the field σ the power counting dimension $[\sigma]$ in the large N expansion, for $2 \leq d \leq 4$:

$$[\sigma] = 1. \quad (20.56)$$

Corrections to scaling and IR fixed point. The IR fixed point is determined by demanding the cancellation of the leading corrections to scaling. We consider the gap equation (20.46) where $\Omega(M)$ expanded as in equation (20.47). The leading correction to scaling is proportional to $(a(\varepsilon) \sim 1/8\pi^2\varepsilon)$

$$M^3 \left(\frac{\lambda}{6g^4} - \tilde{N}a(d)\Lambda^{d-4} \right).$$

Demanding the cancellation of the coefficient of M^3 , one obtains the relation between λ and g^2 ($\tilde{N} = 4N$),

$$g_*^4 = \frac{\lambda_* \Lambda^{4-d}}{6\tilde{N}a(d)} = \frac{\lambda_* \varepsilon \pi^2}{3N} \Lambda^{4-d} + O(\varepsilon^2), \quad (20.57)$$

a result consistent with the results (20.39) of the ε -expansion for N large.

In the same way, it is possible to calculate the leading correction to the σ -propagator (20.53). Demanding the cancellation of the leading correction, one obtains

$$\frac{p^2}{g_*^2} + \frac{\lambda_*}{3g_*^4} M^2 - \frac{1}{2} \tilde{N} (p^2 + 4M^2) \Lambda^{d-4} a(d) = 0.$$

From the coefficient of M^2 one recovers the relation (20.57). The cancellation the coefficient of p^2 yields

$$g_*^2 = \frac{2\Lambda^{4-d}}{\tilde{N}a(d)} = \frac{4\pi^2\varepsilon}{N} + O(\varepsilon^2) \Rightarrow \lambda^* = \frac{24\Lambda^{d-4}}{\tilde{N}a(d)}.$$

a result again consistent with the expressions (20.39).

20.3.1 GN and GNY models

The terms $\nabla\sigma)^2$ and σ^4 of the large N action have an IR dimension 4 and, therefore, for $d < 4$ are irrelevant (and for $d = 4$ marginally irrelevant) in the critical region. We recognize a situation already encountered in the $(\phi^2)^2$ field theory in the large N limit. In the critical region, it is then possible to omit them, and one finds the action

$$\mathcal{S}_N(\bar{\psi}, \psi, \sigma) = \int d^d x \left[-\bar{\psi}(x) \cdot (\not{\partial} + \sigma(x)) \psi(x) + \Lambda^{d-4} \frac{m^2}{2g^2} \sigma^2(x) \right]. \quad (20.58)$$

After integration over the σ field, the action reduces to

$$\mathcal{S}_N(\bar{\psi}, \psi) = - \int d^d x \left[\bar{\psi}(x) \cdot \not{\partial} \psi(x) + \frac{\Lambda^{4-d}}{2m^2} g^2 (\bar{\psi}(x) \cdot \psi(x))^2 \right],$$

which is the action of the GN model. The GN and GNY models are thus equivalent for the universal large distance physics [204]. Within the large N expansion, at least, in the GN model, unlike in the GNY model, the σ particle appears not as a fundamental particle, but as a $\bar{\psi}\psi$ bound state, at threshold in the large N limit [204, 211].

Conversely, it would seem that the GN model depends on a smaller number of parameters than its renormalizable extension. Again, this problem is only interesting in four dimensions where corrections to scaling, that is, to free field theory, are important. However, if we examine the divergences of the term $\text{tr} \ln(\not{\partial} + \sigma)$ in the effective action (20.44) relevant for the large N limit, we find a local polynomial in σ of the form

$$\int d^4 x \left[A\sigma^2(x) + B(\nabla\sigma(x))^2 + C\sigma^4(x) \right].$$

Therefore, the value of the determinant can be modified by a local polynomial of this form by changing the way the cut-off is implemented: additional parameters, as in the case of the non-linear σ -model, are hidden in the cut-off procedure. Near two dimensions, these operators can be identified with $(\bar{\psi}\psi)^2$, $[\partial_\mu(\bar{\psi}\psi)]^2$, $(\bar{\psi}\psi)^4$. It is clear that, by changing the cut-off procedure, one modifies the amplitude of higher dimension operators. In the IR limit, these bare operators have a component on all lower dimensional renormalized operators.

Finally, note that we could have added to the GNY model an explicit breaking term linear in the σ field, which becomes a fermion mass term in the GN model, and which would have played the role of the magnetic field for a ferromagnet.

20.4 The large N expansion

Using the large N dimension of fields and power counting arguments, one can then prove that the $1/N$ expansion is renormalizable, with arguments quite similar to those presented in Section 18.7.

Alternative theory. To prove that the large N expansion is renormalizable, one proceeds as in the case of the $O(N)$ -symmetric scalar theory in Section 18.7. One starts from a critical action with an additional term quadratic in σ , which generates the large N σ -propagator already in perturbation theory,

$$\mathcal{S}(\psi, \bar{\psi}, \sigma) = \int d^d x \left[-\bar{\psi}(x)(\not{\partial} + \sigma(x))\psi(x) + \frac{1}{2v^2} \sigma(x)(-\nabla^2)^{d/2-1} \sigma(x) \right]. \quad (20.59)$$

The initial theory is recovered in the limit $v \rightarrow \infty$. One then rescales σ in $v\sigma$. The model is renormalizable without σ field renormalization, because divergences generate only local counter-terms:

$$\mathcal{S}_r(\psi, \bar{\psi}, \sigma) = \int d^d x \left[-Z_\psi \bar{\psi}(x) (\not{D} + v_r Z_v \sigma(x)) \psi + \frac{1}{2} \sigma(x) (-\nabla^2)^{d/2-1} \sigma(x) \right]. \quad (20.60)$$

RG equations follow:

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta_{v^2}(v) \frac{\partial}{\partial v^2} - \frac{n}{2} \eta_\psi(v) \right) \tilde{\Gamma}^{(l,n)} = 0. \quad (20.61)$$

Again the large N expansion is obtained by first summing the bubble contributions to the σ -propagator. We define

$$R(v) = \frac{2}{b(d)} + Nv^2.$$

Then, the σ propagator for large N reads,

$$\langle \sigma \sigma \rangle = \frac{2}{b(d)R(v)p^{d-2}}. \quad (20.62)$$

The solution to the RG equations can be written as,

$$\tilde{\Gamma}^{(l,n)}(\ell p, v, \Lambda) = Z^{-n/2}(\ell) \ell^{d-l-n(d-2)/2} \tilde{\Gamma}^{(l,n)}(p, v(\ell), \Lambda), \quad (20.63)$$

with the usual definitions,

$$\ell \frac{dv^2}{d\ell} = \beta(v(\ell)), \quad \ell \frac{d \ln Z}{d\ell} = \eta_\psi(v(\ell)).$$

We are interested in the neighbourhood of the fixed point $v^2 = \infty$. Then, the RG function $\eta(v)$ converges towards the exponent η . The flow equation for the coupling constant becomes,

$$\ell \frac{dv^2}{d\ell} = \rho v^2, \Rightarrow v^2(\ell) \sim \ell^\rho.$$

Again, we note that a vertex function with l σ fields becomes proportional to v^l . Therefore,

$$\tilde{\Gamma}^{(l,n)}(\ell p, v, \Lambda) \propto \ell^{d-(1-\rho/2)l-n(d-2+\eta_\psi)/2}. \quad (20.64)$$

We conclude that

$$d_\sigma = \frac{1}{2}(d-2+\eta_\sigma) = 1 - \frac{1}{2}\rho \Leftrightarrow \eta_\sigma = 4-d-\rho. \quad (20.65)$$

RG functions at order $1/N$. A new generic integral is useful here:

$$\frac{1}{(2\pi)^d} \int \frac{d^d q (\not{p} + \not{q})}{(p+q)^{2\mu} q^{2\nu}} = \not{p} p^{d-2\mu-2\nu} \frac{\Gamma(\mu+\nu-d/2)\Gamma(d/2-\mu+1)\Gamma(d/2-\nu)}{(4\pi)^{d/2}\Gamma(\mu)\Gamma(\nu)\Gamma(d-\mu-\nu+1)}. \quad (20.66)$$

We first calculate the $1/N$ contribution to the fermion two-point function at the critical point,

$$\tilde{\Gamma}_{\psi\bar{\psi}}^{(2)}(p) = i\not{p} + \frac{2iv^2}{b(d)R(v)(2\pi)^d} \int^\Lambda \frac{d^d q (\not{p} + \not{q})}{q^{d-2}(p+q)^2}.$$

We need the coefficient of $\not{p} \ln \Lambda / p$. Since we work only at one-loop order, we again replace the σ propagator $1/q^{d-2}$ by $1/q^{2\nu}$, and send the cut-off to infinity. The residue of the pole at $2\nu = d - 2$ gives the coefficient of the term $\not{p} \ln \Lambda$, and the finite part the $\not{p} \ln p$ contribution. We find

$$\tilde{\Gamma}_{\bar{\psi}\psi}^{(2)}(p) = i\not{p} + \frac{2iv^2}{b(d)R(v)} N_d \left(\frac{d-2}{d} \right) \not{p} \ln(\Lambda/p), \quad (20.67)$$

where $N_d = 2/(4\pi)^{d/2} \Gamma(d/2)$ is the loop factor. Expressing that the $\langle \bar{\psi}\psi \rangle$ function satisfies RG equations, we immediately obtain the RG function

$$\eta_\psi(v) = \frac{v^2}{R(v)} \frac{(d-2)}{d} X_1, \quad (20.68)$$

where X_1 is given by equation (18.98).

We then calculate the function $\langle \sigma \bar{\psi}\psi \rangle$ at order $1/N$:

$$\tilde{\Gamma}_{\sigma\bar{\psi}\psi}^{(3)}(p) = v + A_1 D^{-1}(v) v^3 \ln \Lambda,$$

with

$$A_1 = -\frac{2}{b(d)} N_d = -X_1.$$

The last diagram vanishes for symmetry reasons.

The β -function follows:

$$\beta_{v^2}(v) = \frac{4(d-1)v^4}{d} X_1 D^{-1}(v), \quad (20.69)$$

and thus

$$\rho = \frac{8(d-1)N_d}{db(d)N} = \frac{4(d-1)}{dN} X_1.$$

The exponents η_ψ and η_σ at order $1/N$, and thus the corresponding dimensions of fields d_ψ, d_σ , follow:

$$\eta_\psi = \frac{(d-2)}{d} \frac{X_1}{N} = \frac{(d-2)^2}{d} \frac{\Gamma(d-1)}{\Gamma^3(d/2)\Gamma(2-d/2)N}, \quad (20.70)$$

$$2d_\psi = d - 1 - \frac{2(d-2)}{d} \frac{X_1}{N}. \quad (20.71)$$

For $d = 4 - \varepsilon$, we find $\eta_\psi \sim \varepsilon/N$, a result consistent with equation (20.32), for N large. For $d = 2 + \varepsilon$, by contrast, one finds $\eta_\psi \sim \varepsilon^2/2N$, a result consistent with equation (20.82). The dimension d_σ of the field σ is

$$d_\sigma = \frac{1}{2}(d-2+\eta_\sigma) = 1 - \frac{2(d-1)}{dN} X_1 + O(1/N^2). \quad (20.72)$$

A similar evaluation of the $\langle \sigma^2 \sigma \sigma \rangle$ function makes it possible to determine the exponent ν to order $1/N$:

$$\frac{1}{\nu} = d - 2 - \frac{2(d-1)(d-2)}{dN} X_1. \quad (20.73)$$

At present, all exponents are known to order $1/N^2$, except η_ψ which is known to order $1/N^3$ [207]. Moreover, the GNY leading correction exponent ω is known at order $1/N$ [208].

20.5 The GN model

The GN model [205] is described in terms of a $U(N)$ symmetric action for a set of N massless, self-interacting, Dirac fermions $\{\psi^i, \bar{\psi}^i\}$:

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^d x \left[\bar{\psi}(x) \cdot \not{\partial} \psi(x) + \frac{1}{2} G (\bar{\psi}(x) \cdot \not{\psi}(x))^2 \right].$$

The GN model has the same symmetries as the GNY model, in particular, the discrete symmetry (20.2, 20.3), which prevents the addition of a fermion mass term.

Like the GNY model, the model illustrates the physics of spontaneous fermion mass generation and, in even dimensions, chiral symmetry breaking. However, it is renormalizable and asymptotically free in two dimensions.

Like the non-linear σ -model, the perturbative GN model describes only one phase. The main difference is that the role of the spontaneously broken and the explicitly symmetric phase are exchanged. Quite generally, indeed, it is the massless phase that is unstable in low dimensions. Since the symmetry breaking mechanism is non-perturbative, it is instructive to compare the GN and GNY models.

20.5.1 RG equations near and in two dimensions

The GN model is renormalizable in two dimensions and, in perturbation theory, only describes the massless symmetric phase. Perturbative calculations in two dimensions can be made with an IR cut-off of the form of a mass term $\mathcal{M}\bar{\psi}\psi$, which softly breaks the chiral symmetry. It is possible to use dimensional regularization in practical calculations. Note that, in two dimensions, the symmetry group is really $O(2N)$ ($\tilde{N} = 2N$ for $d = 2$), as one verifies after some relabelling of the fields. Therefore, the $(\bar{\psi}\psi)^2$ interaction is multiplicatively renormalized. In generic dimensions $d > 2$, the situation is more complicated, because the algebra of γ matrices is infinite dimensional, and an infinite number of four-fermion interactions mix under renormalization. The following coupling thus has the interpretation of a coupling constant that parametrizes the RG flow that joins the Gaussian fixed point to the non-trivial fixed point. This remark is important from the point of view of explicit calculations in a $(d - 2)$ expansion but, because the problem does not show up at leading order and does not affect the analysis here, we disregard this subtlety.

It is convenient to introduce the dimensionless coupling constant here,

$$u = G\Lambda^{2-d}. \quad (20.74)$$

As a function of the cut-off Λ , the bare vertex functions satisfy the RG equations:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(u) \frac{\partial}{\partial u} - \frac{n}{2} \eta_\psi(u) - \eta_{\mathcal{M}}(u) \mathcal{M} \frac{\partial}{\partial \mathcal{M}} \right] \tilde{\Gamma}^{(n)}(p_i; u, \mathcal{M}, \Lambda) = 0. \quad (20.75)$$

A direct calculation of the β -function in $d = 2 + \varepsilon$ dimension yields [209, 210],

$$\beta(u) = \varepsilon u - (\tilde{N} - 2) \frac{u^2}{2\pi} + (\tilde{N} - 2) \frac{u^3}{4\pi^2} + \frac{(\tilde{N} - 2)(\tilde{N} - 7)}{32\pi^3} u^4 + O(u^5), \quad (20.76)$$

Note that $\tilde{N} = 2N$ for $d = 2$.

The special case $\tilde{N} = 2$, for which the β -function vanishes identically in two dimensions, corresponds to the Thirring model (because for $\tilde{N} = 2$ $(\bar{\psi}\gamma_\mu\psi)^2 = -2(\bar{\psi}\psi)^2$). The latter model is equivalent to the sine-Gordon or the $O(2)$ vector model (Chapter 31).

Finally, the field and mass RG functions, at the order presently available, are

$$\begin{aligned}\eta_\psi(u) &= \frac{\tilde{N}-1}{8\pi^2}u^2 - \frac{(\tilde{N}-1)(\tilde{N}-2)}{32\pi^3}u^3 + \frac{(\tilde{N}-1)(\tilde{N}^2-7\tilde{N}+7)}{128\pi^4}u^4, \\ \eta_{\mathcal{M}}(u) &= \frac{\tilde{N}-1}{2\pi}u - \frac{\tilde{N}-1}{8\pi^2}u^2 - \frac{(2\tilde{N}-3)(\tilde{N}-2)}{32\pi^3}u^3 + O(u^4).\end{aligned}\quad (20.77)$$

As in the example of the non-linear σ -model, the solution of the RG equations (20.75) involves a length scale ξ of the type of a correlation length, which is an RG invariant,

$$\xi^{-1}(u) \equiv \Lambda(u) \propto \Lambda \exp \left[- \int^u \frac{du'}{\beta(u')} \right]. \quad (20.78)$$

20.5.2 Discussion

Two dimensions. For $d = 2$, the model is asymptotically free. Correspondingly, in the chiral theory ($\mathcal{M} = 0$), the spectrum is non-perturbative, and many arguments lead to the conclusion that the chiral symmetry is always broken, and a fermion mass generated. From the statistical point of view, this corresponds to the existence of a gap in the spectrum of fermion excitations (as in a superfluid or superconductor). All masses are proportional to the mass parameter $\Lambda(u)$ which is an RG invariant. Its dependence in the coupling constant is given by equation (20.78):

$$\Lambda(u) \propto \Lambda u^{1/(\tilde{N}-2)} e^{-2\pi/(\tilde{N}-2)u} (1 + O(u)). \quad (20.79)$$

We see that the continuum limit, which is reached when the masses are small compared to the cut-off, corresponds to $u = O(1/\ln(\Lambda/m)) \ll 1$.

S-matrix considerations have then led to the conjecture that, for N finite, the spectrum is [212]

$$m_n = \Lambda(u) \frac{\tilde{N}-2}{\pi} \sin \left(\frac{n\pi}{\tilde{N}-2} \right), \quad n = 1, 2, \dots < \tilde{N}, \quad \tilde{N} > 4,$$

To each mass value corresponds a representation of the $O(\tilde{N})$ group. The nature of the representation leads to the conclusion that n odd corresponds to fermions, and n even to bosons.

This result is consistent with the spectrum for N large evaluated by semi-classical methods in Section 20.3. In particular, the ratio of the masses of the fundamental fermion and the lowest lying boson is

$$\frac{m_\sigma}{m_\psi} = 2 \cos \left(\frac{\pi}{\tilde{N}-2} \right) = 2 + O(1/N^2). \quad (20.80)$$

Note that the two first values of \tilde{N} are special, the model $\tilde{N} = 4$ is conjectured to be equivalent to two decoupled sine-Gordon models.

Dimension $d = 2 + \varepsilon$. As for the non-linear σ -model, asymptotic freedom for $d = 2$ implies, for $d - 2 = \varepsilon > 0$, the existence of a non-trivial UV fixed point, given by

$$u_c = \frac{2\pi}{\tilde{N}-2} \varepsilon \left(1 + \frac{\varepsilon}{\tilde{N}-2} \right) + O(\varepsilon^3).$$

The value u_c of the coupling constant is also the critical value for the transition between a phase in which the chiral symmetry is spontaneously broken and a massless small u phase.

At the fixed point, one obtains the correlation length exponent ν :

$$\nu^{-1} = -\beta'(u_c) = \varepsilon - \frac{\varepsilon^2}{\tilde{N} - 2} + O(\varepsilon^3). \quad (20.81)$$

The fermion field dimension $[\psi]$ is given by

$$2[\psi] = d - 1 + \eta_\psi(u_c) = 1 + \varepsilon + \frac{\tilde{N} - 1}{2(\tilde{N} - 2)^2} \varepsilon^2 + O(\varepsilon^3). \quad (20.82)$$

The dimension of the composite field $\sigma = \bar{\psi}\psi$ is given by

$$[\sigma] = d - 1 - \eta_{\mathcal{M}}(u_c) = 1 - \frac{\varepsilon}{\tilde{N} - 2}.$$

As for the σ -model, the existence of a non-trivial UV fixed point implies that, above two dimensions, the large momentum behaviour is not given by perturbation theory, and this explains why the perturbative result that a renormalized model cannot be defined in higher dimensions cannot be trusted. In the massless phase, the length scale (20.78) for $|u - u_c|$ small is a crossover scale between the free behaviour at largest distances and the critical behaviour, which corresponds to the large momentum behaviour of the renormalized theory.

However, to investigate whether the results of the ε expansion make sense beyond an infinitesimal neighbourhood of dimension 2, other methods are required, like the $1/N$ expansion which is discussed in Section 20.3.

21 Abelian gauge theories: The framework of quantum electrodynamics (QED)

With this chapter, we begin the study of a new class of geometric models, which has played a central role in the theory of fundamental interactions for more than a century: gauge theories. They are characterized by new physics and new formal properties.

We devote this first chapter to a simple and physically important example, the *Abelian gauge theory*, whose physical realization is QED. However, since many textbooks [213, 215, 214] deal extensively with QED, we focus here mainly on the more formal properties of Abelian gauge theories. However, it is worth recalling that QED is the best tested theory. For example, a comparison between the experimental and theoretical values of the anomalous magnetic moment $g - 2$ of the electron, calculated up to order α^5 ($\alpha = e^2/4\pi\hbar c$ is the fine structure constant), yields [216]

$$1/\alpha = 137.035999149(33).$$

The set-up of the chapter is the following: we begin with simple considerations about the free-massive vector field. Its quantization does not immediately follow from the quantization of the scalar field, and thus requires a specific discussion. We show that coupling to the matter field leads to field theories that are renormalizable in four dimensions only if the vector field is coupled to a conserved current. In the latter case, a massless vector limit can be defined. In this limit, the corresponding field theories become *gauge invariant*. We then discuss specific properties of gauge-invariant theories, like the necessity of *gauge fixing*, and mention the infrared (IR) problem of physical observables. We also quantize gauge theories starting directly from first principles. The formal equivalence between different gauges is established.

In Section 21.8, regularization methods are presented, which overcome the new difficulties one encounters in gauge theories. The Abelian gauge symmetry, broken by gauge-fixing terms, leads to a set of Ward–Takahashi (WT) identities which are used to prove the renormalizability of the quantum field theory (QFT). The gauge-dependence of correlation functions in a set of covariant gauges is determined. Renormalization group (RG) equations follow, and we calculate the RG β -function at leading order.

As an introduction to the *Standard Model (SM) of particle physics*, described in Chapter 22, we study the Abelian Landau–Ginzburg–Higgs model, which also leads, in three dimensions, to a macroscopic description of superconductors in a magnetic field. Finally, we comment about the stochastic quantization of gauge theories.

The appendix contains the derivation of the Casimir effect, more remarks about gauge dependence, and the calculation of one-loop divergences using Schwinger’s representation.

21.1 The free massive vector field: Quantization

Notation. In this Section and in Section 21.5, expressions are written for real time. *Summation over repeated lower and upper Greek indices* is then implied. Roman indices denote space components. Space-time coordinates in d dimensions are denoted by $x^\mu \equiv \{t = x_0, x_i\}$, with $\mu = 0, \dots, d-1$, and $i = 1, \dots, d-1$. The Minkowski metric has the signature $\{+, -, -, \dots\}$. The vector field A_μ has components $\{A_0, A_i\}$, the first component being the time component. Everywhere, we set $\hbar = c = 1$.

21.1.1 The classical action: Degrees of freedom

Notation. In the following sections time and space derivatives are indicated by

$$\frac{\partial A_i}{\partial t} \equiv \dot{A}_i, \quad \nabla_x \equiv (0, \partial_1, \dots, \partial_{d-1}), \text{ with the notation } \partial_i \equiv \frac{\partial}{\partial x_i}.$$

The action. The local $O(1, d - 1)$ and translation invariant classical action for a free massive vector field A_μ can be written as ($\mu, \nu = 0, \dots, d - 1$),

$$\mathcal{A}(\mathbf{A}) = - \int dt d^{d-1}x \left[\frac{1}{4} F_{\mu\nu}(t, x) F^{\mu\nu}(t, x) - \frac{1}{2} m^2 A_\mu(t, x) A^\mu(t, x) \right], \quad (21.1)$$

with ($\partial_\mu \equiv \partial/\partial x^\mu$)

$$F_{\mu\nu}(t, x) = \partial_\mu A_\nu(t, x) - \partial_\nu A_\mu(t, x). \quad (21.2)$$

One may wonder about the peculiar form of the derivative term, but it is simple to verify that the additional covariant term one could think of adding, namely $\partial_\mu A^\mu \partial_\nu A^\nu$, depending on its sign corresponds either to a A_0 field with a negative metric, or to an unbounded potential.

Dynamical degrees of freedom. The action (21.1) has one peculiar property: the time component A_0 of the vector field has no conjugate momentum, since the action does not depend on the time derivative \dot{A}_0 . Actually, the action can be rewritten as

$$\begin{aligned} \mathcal{A}(\mathbf{A}) = & \int dt d^{d-1}x \left\{ \frac{1}{2} \sum_{i,j} \left[\dot{A}_i(t, x) \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla_x^2 - m^2} \right) \dot{A}_j(t, x) - \frac{1}{4} F_{ij}^2(t, x) \right] \right. \\ & - \frac{1}{2} m^2 \sum_i A_i^2(x) + \frac{1}{2} \left((-\nabla_x^2 + m^2) A_0(t, x) + \nabla_x \cdot \dot{\mathbf{A}}(t, x) \right) (-\nabla_x^2 + m^2)^{-1} \\ & \times \left. \left((-\nabla_x^2 + m^2) A_0(t, x) + \nabla_x \cdot \dot{\mathbf{A}}(t, x) \right) \right\}. \end{aligned} \quad (21.3)$$

Therefore, A_0 is not a dynamical degree of freedom, and the corresponding field equation,

$$\frac{\delta \mathcal{A}}{\delta A_0(t, x)} = (-\nabla_x^2 + m^2) A_0(t, x) + \nabla_x \cdot \dot{\mathbf{A}}(t, x) = 0, \quad (21.4)$$

is a constraint equation, which can be used to eliminate A_0 from the action. This feature reflects the property that a massive vector field has only $(d - 1)$ physical degrees of freedom (massive vector fields transform, in the rest frame, under representations of the $SO(d - 1)$ group). The Lagrangian density then reduces to

$$\mathcal{L}(\mathbf{A}) = \sum_{i,j} \left[\frac{1}{2} \dot{A}_i(t, x) \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla_x^2 - m^2} \right) \dot{A}_j(t, x) - \frac{1}{4} F_{ij}^2(t, x) \right] - \frac{1}{2} m^2 \sum_i A_i^2(t, x). \quad (21.5)$$

We denote by E_i (because it becomes the electric field in the massless limit) the momentum conjugated to A_i ,

$$E_i(t, x) = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = \dot{A}_i(t, x) - \partial_i (\nabla_x^2 - m^2)^{-1} \nabla_x \cdot \dot{\mathbf{A}}(t, x).$$

A Legendre transformation yields the corresponding Hamiltonian density,

$$\mathcal{H}(\mathbf{E}, \mathbf{A}) = \sum_{i,j} \left[\frac{1}{2} E_i(x) (\delta_{ij} - \partial_i \partial_j / m^2) E_j(x) + \frac{1}{4} F_{ij}^2(x) \right] + \frac{1}{2} m^2 \sum_i A_i^2(x). \quad (21.6)$$

The differential operator $-\partial_i \partial_j$ being non-negative, the Hamiltonian is positive.

The action in the Hamiltonian formalism is

$$\mathcal{A}(E, A) = \int dt d^{d-1}x \left[\sum_i E_i(t, x) \dot{A}_i(t, x) - \mathcal{H}(E, A) \right].$$

21.1.2 Field integral quantization

The field integral quantization from now on is standard. The evolution operator is given by (see Chapter 3),

$$\mathcal{U} = \int [dA dE] \exp [i\mathcal{A}(E, A)/\hbar].$$

(In what follows we always set $\hbar = 1$.) Since the action is quadratic in E , the corresponding Gaussian integration can be performed. In the exponential, the Lagrangian (21.5) is recovered (see Section 3.3). A factor $\det^{-1/2}(\delta_{ij} - \partial_i \partial_j / m^2)$ is also generated which is field independent, and thus can be absorbed into the normalization.

However, the reduced Lagrangian (21.5) has unwanted properties: it is not local and not $O(1, d - 1)$ space-time invariant. We then note that the dependence in A_0 of the action (21.3) is quadratic. Therefore, we can proceed in the following way: we substitute in the field integral representation of the partition function the initial Lagrangian. We then perform the Gaussian integral over the time component. As we know, this is equivalent to solving the corresponding equation of motion, and we thus recover the reduced Lagrangian. Finally, the determinant resulting from the integration is again field independent (see Section A22 for the more interesting non-Abelian case).

This shows that, if we had ignored all these subtleties about quantization, we would have used the physics-acceptable covariant Lagrangian (21.1).

21.2 The Euclidean free action. The two-point function

We now proceed by analytic continuation to imaginary time, setting $x_0 = -ix_d$ and $A_0 = -iA_d$. The Euclidean action reads (here $\mathbf{A} \equiv (A_1, \dots, A_d)$),

$$\mathcal{S}(\mathbf{A}) = \int d^d x \left[\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) + \frac{1}{2} m^2 \mathbf{A}^2(x) \right], \quad (21.7)$$

with ($\partial_\mu \equiv \partial/\partial x_\mu$)

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x), \quad (21.8)$$

Connected two-point function. The generating functional $\mathcal{Z}(\mathbf{J})$ of \mathbf{A} -field correlation functions is given by

$$\mathcal{Z}(\mathbf{J}) = \int [dA] \exp \left[-\mathcal{S}(\mathbf{A}) + \int d^d x \mathbf{J}(x) \cdot \mathbf{A}(x) \right]. \quad (21.9)$$

The corresponding field equations are (here $\nabla \equiv (\partial_{x_1}, \dots, \partial_{x_d})$)

$$(-\nabla^2 + m^2) A_\mu(x) + \partial_\mu \nabla \cdot \mathbf{A}(x) = J_\mu(x). \quad (21.10)$$

Writing the solution as

$$A_\mu(x) = \int d^d y \sum_\nu \Delta_{\mu\nu}(x - y) J_\nu(y),$$

after integration over \mathbf{A} , one obtains

$$\mathcal{W}(\mathbf{J}) = \ln \mathcal{Z}(\mathbf{J}) = \frac{1}{2} \int d^d x d^d y \sum_{\mu, \nu} J_\mu(x) \Delta_{\mu\nu}(x - y) J_\nu(y). \quad (21.11)$$

The two-point connected correlation function or vector propagator $\Delta_{\mu\nu}(x)$ has the Fourier representation,

$$\Delta_{\mu\nu}(x) = \frac{1}{(2\pi)^d} \left(\delta_{\mu\nu} - \frac{\partial_\mu \partial_\nu}{\nabla^2} \right) \int d^d k \frac{e^{ikx}}{k^2 + m^2} = \frac{1}{(2\pi)^d} \int d^d k e^{ikx} \tilde{\Delta}_{\mu\nu}(k), \quad (21.12)$$

with

$$\tilde{\Delta}_{\mu\nu}(k) = \frac{\delta_{\mu\nu} + k_\mu k_\nu / m^2}{k^2 + m^2}. \quad (21.13)$$

At the pole $k_\mu = i m e_\mu$ with $\sum_\mu e_\mu^2 = 1$ (the mass-shell), the numerator $\delta_{\mu\nu} - e_\mu e_\nu$ of the propagator is a projector transverse to the vector k , propagating $(d-1)$ components belonging to the fundamental representation of the $O(d-1)$ subgroup of $O(d)$ which leaves the vector k_μ invariant.

For $|k_\mu| \rightarrow +\infty$, $\tilde{\Delta}_{\mu\mu}$ goes to a constant. Power counting then implies that field theories involving massive vector fields with the action (21.7), coupled to matter, are renormalizable only in dimensions $d \leq 2$. A directly related disease (because the propagator is a homogeneous function of k and m) is the divergence of the massless limit.

Coupling to a conserved current. Considering the first expression in equations (21.12), we note that if the source $J_\mu(x)$ has the form of a conserved current, that is, satisfies

$$\nabla \cdot \mathbf{J}(x) = 0, \quad (21.14)$$

the propagator can be replaced by

$$\Delta_{\mu\nu}(x) = \frac{\delta_{\mu\nu}}{(2\pi)^d} \int d^d k \frac{e^{ikx}}{k^2 + m^2}. \quad (21.15)$$

The propagator (21.15) behaves like the propagator of a scalar particle. Then, both problems of large momentum behaviour and massless limit are solved. One may now wonder why we have not used such a propagator at once: the reason is that it propagates, in addition to a vector field, a scalar particle with negative metric (like the regulator fields of Section 8.4.2). This is better illustrated by a short calculation.

21.2.1 More general propagators, interpretation

We add to the action (21.7), a term of the form of a regulator (non-physical) field action (with χ imaginary):

$$\mathcal{S}(\mathbf{A}, \chi) = \mathcal{S}(\mathbf{A}) - \frac{1}{2} \int d^d x \left[(\nabla \chi(x))^2 + M^2 \chi^2(x) \right], \text{ with } M > 0. \quad (21.16)$$

In the absence of a source for χ , the generating functional

$$\mathcal{Z}(\mathbf{J}) = \int [dA] [\chi] \exp \left[-\mathcal{S}(\mathbf{A}, \chi) + \int d^d x \mathbf{J}(x) \cdot \mathbf{A}(x) \right], \quad (21.17)$$

after integration over χ , becomes proportional to the functional (21.9).

Instead, we change variables, $\mathbf{A} \mapsto \mathbf{A}'$, in the integral (21.17), setting

$$\mathbf{A}(x) = \mathbf{A}'(x) + \frac{1}{m} \nabla \chi(x). \quad (21.18)$$

The change, called *gauge transformation*, leaves $F_{\mu\nu}$ invariant. If the source satisfies the conservation equation (21.14), the source term is not modified. Only the vector field mass term is affected:

$$\frac{1}{2} m^2 \mathbf{A}^2(x) = \frac{1}{2} m^2 \mathbf{A}'^2(x) + m \mathbf{A}'(x) \cdot \nabla \chi(x) + \frac{1}{2} (\nabla \chi(x))^2 \quad (21.19)$$

and, therefore,

$$\mathcal{S}(\mathbf{A}', \chi) = \mathcal{S}(\mathbf{A}') + \int d^d x (m \mathbf{A}(x) \cdot \nabla \chi(x) - \frac{1}{2} M^2 \chi^2(x)). \quad (21.20)$$

We then integrate over $\chi(x)$, and find the new action (omitting now the prime on \mathbf{A}),

$$\mathcal{S}_\xi(\mathbf{A}) = \mathcal{S}(\mathbf{A}) + \frac{1}{2\xi} \int (\nabla \cdot \mathbf{A}(x))^2 d^d x, \quad \text{with } \xi = \frac{M^2}{m^2}. \quad (21.21)$$

The addition to the action (21.7) of the last term is called *gauge fixing*. By varying ξ from 0 to $+\infty$, one describes a set of covariant *gauges*: $\xi = 0$ corresponds to Landau's gauge, $\xi = 1$ is Feynman's gauge. For $\xi = \infty$ (the unitary gauge), one recovers the initial *unitary* QFT.

A short calculation yields the corresponding propagator (in the Fourier representation),

$$[\tilde{\Delta}_\xi]_{\mu\nu}(k) = \frac{\delta_{\mu\nu}}{k^2 + m^2} + \frac{(\xi - 1) k_\mu k_\nu}{(k^2 + m^2)(k^2 + \xi m^2)} = \frac{\delta_{\mu\nu} + k_\mu k_\nu/m^2}{k^2 + m^2} - \frac{k_\mu k_\nu/m^2}{k^2 + \xi m^2}. \quad (21.22)$$

For all finite values of ξ , the propagator behaves at large momentum like a scalar field propagator. However, it has a non-physical pole for $k^2 = -\xi m^2$ (and $\xi \neq 1$), but which does not contribute on the mass-shell limit $k^2 \rightarrow -m^2$.

21.3 Coupling to matter

We conclude from the analysis of Section 21.1 that vector fields coupled to conserved currents, and thus associated to continuous symmetries, are candidates for the construction of theories renormalizable in four dimensions.

Fermion matter. We now construct an explicit, and physically relevant, interacting example (for Euclidean fermions see Chapter 12).

We start from the Euclidean action for a free Dirac-fermion field with mass M (like the electron–positron field),

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (\not{\partial} + M) \psi(x), \quad (21.23)$$

and want to add an $O(d)$ -invariant coupling to a vector field.

As we have already noted in Section 12.3.3, the action (21.23) for a free Dirac fermion field has a $U(1)$ symmetry, associated with the conservation of the fermion number, corresponding to the transformations

$$\psi(x) = e^{i\theta} \psi'(x), \quad \bar{\psi}(x) = e^{-i\theta} \bar{\psi}'(x), \quad \theta = \text{const.} \quad (21.24)$$

To this symmetry corresponds a current whose expression is obtained by calculating the variation of the action under a space-dependent group transformation (see Section A13.1). If θ is space dependent, the variation of the action is

$$\delta \mathcal{S}(\bar{\psi}, \psi) = -i \int d^d x \bar{\psi}(x) \not{\partial} \theta(x) \psi(x), \quad (21.25)$$

and thus the corresponding conserved current $J_\mu(x)$ is

$$J_\mu(x) = -i \bar{\psi}(x) \gamma_\mu \psi(x).$$

The only $O(d)$ -symmetric interaction which, from the point of view of power counting, is renormalizable in four dimensions, is proportional to

$$\int d^d x \bar{\psi}(x) \not{A}(x) \psi(x).$$

It has exactly the form of a vector field linearly coupled to the conserved current $J_\mu(x)$. The action of a Dirac fermion of mass M coupled to the vector field then takes the form,

$$\mathcal{S}_F(\mathbf{A}, \bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (\not{\partial} + M + ie\not{A}(x)) \psi(x), \quad (21.26)$$

where the parameter e is the current–vector field coupling constant or charge. The complete action is obtained by adding the fermion action to the free vector action (21.7):

$$\mathcal{S}(\mathbf{A}, \bar{\psi}, \psi) = \mathcal{S}(\mathbf{A}) + \mathcal{S}_F(\mathbf{A}, \bar{\psi}, \psi). \quad (21.27)$$

Change of gauges. Following the method of Section 21.2.1, we perform a change variables in the field integral of the form (21.18), setting (this implies also rescaling χ into $m\chi/e$)

$$\mathbf{A}(x) = -\frac{1}{e} \nabla \chi(x) + \mathbf{A}'(x). \quad (21.28)$$

The induced variation of $\mathcal{S}_F(\mathbf{A}, \bar{\psi}, \psi)$ can then be cancelled by a change of the fermion fields of the form of the transformation (21.24) (also called *gauge transformation*) with,

$$\psi(x) = e^{i\chi(x)} \psi'(x), \quad \bar{\psi}(x) = e^{-i\chi(x)} \bar{\psi}'(x). \quad (21.29)$$

Therefore, the algebraic transformations of Section 21.2.1 that relate a unitary non-renormalizable action to an action non-unitary, but renormalizable by power counting with the propagator (21.22), remain justified.

The action of a massive vector coupled with fermion matter in a *renormalizable covariant gauge* then can be written as

$$\mathcal{S}_\xi(\mathbf{A}, \bar{\psi}, \psi) = \mathcal{S}_\xi(\mathbf{A}) - \int d^d x \bar{\psi}(x) (\not{D} + ie\mathcal{A}(x) + M) \psi(x), \quad (21.30)$$

where $\mathcal{S}_\xi(\mathbf{A})$ is the action (21.21). The equivalence between the unitary and renormalizable representations is valid for the action and $F_{\mu\nu}$ -field correlation functions. We discuss the problem of matter correlation functions in Section 21.16.

Coupling to scalar fields. We start from the $U(1)$ invariant action

$$\mathcal{S}_B(\phi) = \int d^d x \left[|\nabla \phi(x)|^2 + U(|\phi(x)|^2) \right], \quad (21.31)$$

in which $\phi(x)$ is a complex scalar field, and replace the derivative ∂_μ by a covariant derivative. The explicit form of the covariant derivative depends on the charge assigned to the field $\phi(x)$. If we assume that, in a *gauge transformation*, ϕ transforms like

$$\phi(x) \mapsto e^{iq_B \chi(x)} \phi(x), \quad (21.32)$$

where $e_B = eq_B$ is the gauge-scalar coupling, after addition of the scalar action,

$$\mathcal{S}_B(A, \phi) = \int d^d x \left[\sum_\mu |(\partial_\mu + ie_B A_\mu(x)) \phi(x)|^2 + U(|\phi(x)|^2) \right], \quad (21.33)$$

the algebraic transformations of Section 21.2.1 that relate a unitary non-renormalizable action to an action non-unitary, but renormalizable by power counting with the propagator (21.22), again remain justified.

Note that, for $d = 4$, in the action (21.33) three independent interaction terms, two linear and one quadratic in the vector field, are renormalizable, but the invariance under gauge transformations relates them.

Finally, in the limit of massless vector field, a geometric interpretation of the vector field and the matter actions will be given in Section 21.7.

Higher spins. One may wonder why the strategy which has led to a renormalizable theory of vector particles does not work for higher spins. Let us take the example of the symmetric traceless rank two tensor. It must be coupled to a conserved current which is also a rank two tensor. Only the energy-momentum tensor $T_{\mu\nu}$ has the required property (see Section A13.2). But $T_{\mu\nu}$ has at least dimension 4 for $d = 4$ and once coupled to a field of at least dimension 1 generates a non-renormalizable interaction. A similar argument applies to a spin 3/2 field coupled to the supersymmetry current.

21.4 The massless limit: Gauge invariance

We now examine the special properties of the massless vector QFT. In the massless limit, the action (21.7) coupled to a conserved current, reduces to

$$\mathcal{S}(A, J) = \int d^d x \left[\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) - \mathbf{J}(x) \cdot \mathbf{A}(x) \right], \quad (21.34)$$

has a *gauge symmetry*: it is exactly invariant under the *gauge transformation* (21.28),

$$A_\mu(x) \mapsto A_\mu(x) - \frac{1}{e} \partial_\mu \chi(x).$$

At each point of space, gauge transformations form a group, isomorphic to the additive, Abelian (*i.e.*, commutative) group of real numbers. Such a group, whose parameters are space dependent, is called a *gauge group*, the vector field a *gauge field*, or connection, and the corresponding QFT is called a (here Abelian) *gauge theory*.

The propagator. The corresponding propagator (21.22) has the zero mass limit,

$$[\tilde{\Delta}_\xi]_{\mu\nu}(k) = \frac{\delta_{\mu\nu}}{k^2} + (\xi - 1) \frac{k_\mu k_\nu}{(k^2)^2}. \quad (21.35)$$

However, for values of $\xi \neq 1$ (Feynman's gauge), the term proportional to $1/k^4$ may generate IR divergences in interacting theories for dimensions $d \leq 4$.

21.4.1 Massless vector field and matter

Fermion matter. In the massless vector field limit $m = 0$, the action (21.27) reduces to,

$$\mathcal{S}(\mathbf{A}, \bar{\psi}, \psi) = \int d^d x \left[\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) - \bar{\psi}(x) (\not{D} + M) \psi(x) \right]. \quad (21.36)$$

where $\not{D} = \sum_\mu D_\mu \gamma_\mu$, and $D_\mu = \partial_\mu + ieA_\mu$ is a *covariant derivative*. The action has still a gauge symmetry, but the gauge group is reduced to the *U(1) local group* (*i.e.*, with a space-dependent parameter), the group of *gauge transformations* (21.28) and (21.29).

From the physics viewpoint, the QFT describes the interaction of charged fermions with electric charge e , like electron–positron, with the electromagnetic field (associated to the photon), the basis of QED. By trying to build a QFT renormalizable in four dimensions, involving a vector field, we have been naturally led to introduce a new geometric structure, an Abelian gauge theory, a quantum version of Maxwell's electromagnetism.

Charged scalars. With the addition of scalar matter coupled with a different covariant derivative, like in the action (21.33), and transforming as in equation (21.32), the total action is gauge invariant, and $U(1)$ remains the gauge group if q_B is a rational number.

21.4.2 The massless vector field as a limit

Geometric and physics considerations single out the massless vector field: the QFT is exactly gauge invariant, and it describes the physics of QED, because the photon seems to be massless. Therefore, we could have restricted our discussion to the massless case, as we do for the non-Abelian gauge theory (Chapter 22). However, considering the massless theory as a limit of the massive theory provides us with a resolution to several difficulties.

In Section 21.1, we have found a difficulty in taking the massless limit of the propagator. In a gauge-invariant theory, this problem persists beyond perturbation theory.

Indeed, gauge symmetry implies that the action does not depend on one combination of the dynamical variables, which can be removed by a gauge transformation. As a consequence, the field integral cannot be defined because the volume of the gauge group, which can be factorized, is infinite (this statement has a precise meaning only in the framework of lattice regularization).

This difficulty reflects the property that in classical electrodynamics only the electromagnetic tensor $F_{\mu\nu}$ is physical. The vector field A_μ is a mathematical entity that enables deriving the classical field equations from a local covariant action. It contains redundant degrees of freedom, whose evolution is not determined by the field equations.

In Section 21.5, we quantize the gauge-invariant action starting from first principles. The procedure is less straightforward and leads to non-covariant gauges. By giving a mass to the gauge field A_μ , we make $(d - 1)$ field components dynamical. The symmetry properties of the action make the algebraic transformations explained in Section 21.2.1 possible (if one calculates only gauge-invariant observables, see Section 21.10). We can eventually take the zero mass limit because, in the process, the gauge-dependent part of the gauge field has acquired a dynamics: we have ‘fixed’ the gauge.

Finally, the mass of the vector field provides the theory with a natural IR cut-off which somewhat simplifies the analysis of IR problems. Therefore, in what follows, we mostly work with the massive theory, the gauge-invariant theory appearing as a limiting case.

Physical observables and IR problems. In the gauge-invariant limit, in which the mass of the vector field vanishes, correlation functions calculated in Feynman’s gauge have a finite limit. In Section 21.10, we derive a relation (equation (21.97)) that exhibits the gauge-dependence of the bare two-point function. The corresponding relations for other correlation functions is described in Section A21.2. They imply that bare correlation functions have a limit in all gauges. The renormalized correlation functions have then also a limit, provided one has been careful not to choose IR divergent renormalization constants (as it is usual in massless theories).

However, the important question is to understand whether physical, that is, gauge independent observables, are IR finite. The expectation values of gauge-invariant operators (local polynomials in the fields) certainly have limits. By contrast, we do not expect scattering amplitudes to be IR finite. Indeed, we have shown in Section 5.3.3, using the eikonal approximation, that, already in non-relativistic quantum mechanics, IR divergences appear in the phase of the scattering amplitude, because the Coulomb force decays too slowly at large distance. These divergences survive in the relativistic theory. Moreover, additional relativistic IR divergences appear which cancel only if one adds to the scattering amplitude, the amplitude for producing any number of additional low momentum vector (gauge) particles. This is physically acceptable, because particles which have momenta smaller than the uncertainty in the momenta measurements cannot be detected. We refer to the literature for detailed discussions [220].

Quantization of charge. If the charges of all fields are rationally related, there exists a unit charge to which all charges are proportional. Otherwise, one can define states with arbitrary small charges. However, perturbative properties do not seem to depend on charge *quantization*, in particular, WT identities. By contrast, the only known non-perturbative regularization, based on a lattice approximation, involves group elements in the form of parallel transporters and, thus, seems to require charge quantization (Section 21.8). This non-physical (?) question remains open.

21.5 Massless vector field, gauge invariance, and quantization

Notation. In this section, where we discuss directly quantization of the massless theory in different gauges, we return to the *real-time formalism*.

Although we have been able to quantize the massless vector field coupled to a conserved current as a limit of a massive vector field, it is useful to contemplate the difficulties one encounters when one tries to quantize the massless theory starting from first principles.

Moreover, as we show in Section A22, in the case of non-Abelian gauge symmetries, the massless limit is not continuous. Therefore, we show how, starting directly from the classical field equations of a *gauge-invariant theory* (massless vector field coupled to a conserved current), it is possible to recover the field integral representation of the generating functional of correlation functions.

The problem can be solved by several different strategies and we present two of them, corresponding to different *gauge fixings*, the *Coulomb's gauge* and the *temporal gauge*.

21.5.1 Coulomb's gauge

We again consider the real time action (21.34). We note that

$$\frac{1}{4} \int d^{d-1}x \sum_{i,j} F_{ij}^2(t, x) = -\frac{1}{2} \int d^{d-1}x \sum_{i,j} A_i(t, x) \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla_x^2} \right) \nabla_x^2 A_j(t, x). \quad (21.37)$$

We first proceed like in the massive case and eliminate the field time component from the action (see equation (21.5)). Taking into account current conservation (note $J_0 = iJ_d$), we then obtain the integral of a reduced Lagrangian density:

$$\begin{aligned} \mathcal{L}(\mathbf{A}) \\ = \frac{1}{2} \sum_{i,j} \left[\dot{A}_i(t, x) \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla_x^2} \right) \dot{A}_j(t, x) - \frac{1}{4} F_{ij}^2(t, x) + J_i(t, x) \left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla_x^2} \right) A_j(t, x) \right]. \end{aligned}$$

Using the identity (21.37), we note that, in contrast with the massive case, the action depends only on $\sum_j (\delta_{ij} - \partial_i \partial_j / \nabla_x^2) A_j(t, x)$.

After Fourier transformation, this implies that the action does not depend on the component of $\tilde{A}_i(t, \hat{k})$ along \hat{k} , the space component of the momentum \mathbf{k} . Therefore, a massless vector field has only $(d-2)$ physical components, a well-known property. We then expand the vector $\tilde{A}_i(t, \hat{k})$ on a transverse basis $e_i^a(\hat{k})$, denoting by \mathfrak{A}_a the corresponding $(d-2)$ components (the polarizations):

$$\hat{k} \cdot \mathbf{e}^a(\hat{k}) = 0, \quad \mathbf{e}^a(\hat{k}) \cdot \mathbf{e}^b(\hat{k}) = \delta_{ab}, \quad \tilde{\mathbf{A}}(t, \hat{k}) = \frac{\hat{k} \cdot \tilde{\mathbf{A}}}{\hat{k}^2} \hat{k} + \sum_{a=1}^{d-2} \mathbf{e}^a(\hat{k}) \mathfrak{A}_a(\mathbf{k}),$$

and by J_a the corresponding sources. The Lagrangian density in these variables becomes

$$\mathcal{L}(\mathfrak{A}_a) = \sum_{a=1}^{d-2} [\partial_\mu \mathfrak{A}_a(t, x) \partial^\mu \mathfrak{A}_a(t, x) + J_a(t, x) \mathfrak{A}_a(t, x)].$$

The quantization then is straightforward. One eventually obtains a field integral over the fields \mathfrak{A}_a . However, the corresponding action, once expressed in terms of the initial current J_μ , is not local. To recover a local action, one reintroduces the components A_i of the gauge field but multiplies the integrand by $\delta(\nabla_x \cdot \mathbf{A})$.

The last step, that is, returning to an integral involving the time component, is the same as in the massive case. The final result is the generating functional in *Coulomb's gauge*:

$$\mathcal{Z}_{\text{Coul.}}(\mathbf{J}) = \int [dA_\mu(t, x)\delta(\nabla_x \cdot \mathbf{A}(t, x))] e^{i\mathcal{A}(\mathbf{A}, \mathbf{J})}. \quad (21.38)$$

In the Abelian gauge theory, Coulomb's gauge has a simple physical interpretation: only physical degrees of freedom propagate, but it leads to non-relativistic covariant calculations, and this is a serious drawback.

Indeed, the *Euclidean* gauge field propagator or two-point function in the Fourier representation ($\mathbf{k} \equiv \{k_d, \mathbf{k}_\perp\}$) reads

$$\widetilde{W}_{dd}^{(2)}(k) = \frac{1}{\mathbf{k}_\perp^2}, \quad \widetilde{W}_{id}(k) = 0, \quad \widetilde{W}_{ij}^{(2)}(k) = \frac{1}{\mathbf{k}^2} \left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}_\perp^2} \right). \quad (21.39)$$

Moreover, the time component does not decrease in the large time direction. Therefore, with this propagator, the QFT is not explicitly renormalizable by power counting.

We still have to prove that this gauge is equivalent to the covariant gauges introduced in Section 21.1, but we postpone this point, and describe first another quantization scheme.

21.5.2 The temporal (or Weyl's) gauge

In the non-Abelian case, the quantization in Coulomb's gauge is complicate and, therefore, we now describe another, more easily generalizable, gauge: the temporal gauge.

The field equations in *real time* t , which correspond to the action (21.34), are

$$\partial_\mu F^{\mu\nu}(t, x) = J^\nu(t, x), \quad x \in \mathbb{R}^{d-1}. \quad (21.40)$$

The method relies on the observation that the gauge transformed of any solution of equation (21.40) is again a solution. The set of all solutions can thus be described by restricting the gauge field to a *gauge section*, considering, for example, only the solutions satisfying [217]

$$A_0(t, x) = 0, \quad (21.41)$$

in which A_0 is the time component of the field \mathbf{A} : this defines the temporal gauge (the analogous gauge condition $A_i(t, x) = 0$, defines the axial gauge [218]).

We then rewrite equation (21.40), separating time and space components, and taking into account the condition (21.41):

$$\nabla_x \cdot \dot{\mathbf{A}}(t, x) = J_0(t, x), \quad (21.42)$$

$$\ddot{A}_i(t, x) - \sum_j \partial_j F_{ji}(t, x) = J_i(t, x), \quad (21.43)$$

in which the indices i and j run only from 1 to $(d - 1)$. By construction, the two equations are compatible if the current J^μ is conserved. Equation (21.43) is simply the field equation that can be derived from the classical Lagrangian density

$$\mathcal{L}(\mathbf{A}) = \sum_i \left[\frac{1}{2} \dot{A}_i^2(t, x) + J_i(t, x) A_i(t, x) \right] - \sum_{i,j} \frac{1}{4} F_{ij}^2(t, x). \quad (21.44)$$

The conjugated momentum of the field $A_i(t, x)$,

$$E_i(t, x) = \dot{A}_i(t, x), \quad (21.45)$$

is the electric field.

The expression of the Hamiltonian density follows:

$$\mathcal{H}(\mathbf{E}, \mathbf{A}) = \sum_i \left(\frac{1}{2} E_i^2(x) - J_i(t, x) A_i(x) \right) + \frac{1}{4} \sum_{i,j} F_{ij}^2(x). \quad (21.46)$$

Expressing the partition function $\mathcal{Z}(\mathbf{J})$ as a field integral in the Hamiltonian formalism, and integrating over the conjugate momenta E_i , one finds

$$\mathcal{Z}(\mathbf{J}) = \int [dA_i] e^{i\mathcal{A}(A_i, J_i)} = \int [dA_\mu] \prod_{t,x} \delta(A_0(t, x)) e^{i\mathcal{A}(A, J)}, \quad (21.47)$$

where $\mathcal{A}(A_i, J_i)$ is the integral of the Lagrangian (21.44) and, thus, the covariant action $\mathcal{A}(A, J)$ of equation (21.34) taken for $A_0 = 0$.

In the temporal gauge $A_0 = 0$, the action has still a gauge invariance corresponding to time-independent gauge transformations,

$$A_i(t, x) \mapsto A_i(t, x) + \partial_i \theta(x).$$

This is now a symmetry of the QFT. In the Hamiltonian formalism, the generator of the symmetry can be expressed in terms of the quantized conjugated momenta \hat{E}_i as

$$\frac{1}{i} \sum_i \partial_i \frac{\delta}{\delta A_i(x)} = \sum_i \partial_i \hat{E}_i(x), \quad (21.48)$$

and this operator commutes with the quantum Hamiltonian.

Quantum Gauss's law. We still have to implement the constraint (21.42), which is *Gauss's law*. After quantization, it becomes a constraint on the space of physical states $\Psi(A)$. The condition (21.42) takes the form

$$\frac{1}{i} \sum_i \partial_i \frac{\delta}{\delta A_i(x)} \Psi(A) = J_0(t, x) \Psi(A). \quad (21.49)$$

The left-hand side of equation (21.49) is the generator (21.48) acting on Ψ . In the absence of an external source ($J_0(t, x) = 0$), the physical states must be gauge invariant. This condition is consistent with quantum evolution, due to the remaining gauge symmetry.

For a general external source, the condition (21.49) defines how the admissible states transform. The compatibility between equations (21.42) and (21.43), implies, after quantization, the commutation of the operator $\sum_i \partial_i \hat{E}_i - J_0$ with the Hamiltonian.

For later purpose, we exhibit a state satisfying the condition (21.49) in the case of two opposite static charges:

$$J_0(t, x) = e \left[\delta^{(d-1)}(x - x_2) - \delta^{(d-1)}(x - x_1) \right], \quad J_i(t, x) = 0. \quad (21.50)$$

The state,

$$\Psi(\mathbf{A}) = \exp \left[-ie \oint_C \sum_i A_i(s) ds_i \right], \quad (21.51)$$

in which C is an arbitrary differentiable path joining x_1 to x_2 , indeed verifies

$$\begin{aligned} \frac{1}{i} \sum_i \partial_i \frac{\delta}{\delta A_i(x)} \Psi(A) &= \frac{1}{i} \frac{\delta}{\delta \theta(x)} \Psi(A - \nabla_x \theta) \Big|_{\theta=0} \\ &= e \left[\delta^{(d-1)}(x - x_2) - \delta^{(d-1)}(x - x_1) \right] \Psi(A), \end{aligned}$$

a result that is consistent with equations (21.49) and (21.50).

The representation (21.51) has the form of a parallel transporter corresponding to time-independent gauge transformations. This representation, as well as its non-Abelian generalization, are useful in Chapter 25 devoted to lattice gauge theories, in the discussion of the *confinement* problem.

21.6 Equivalence with covariant quantization

We have obtained different field integral representations of the same gauge-invariant QFT. We now show, using the *Euclidean formulation*, that, for calculating expectation values of gauge-invariant observables, they are formally equivalent to the $O(d)$ covariant representations discussed in Section 21.1. More generally, we show here the equivalence between the temporal gauge and a class of gauges characterized by a gauge condition of the form

$$\mathbf{n}(\nabla) \cdot \mathbf{A}(x) = \nu(x), \quad (21.52)$$

where the vector \mathbf{n} is a constant or a differential operator, and $\nu(x)$ an arbitrary external field (this contains also the axial gauge [218]). By setting the external field $\nu(x)$ to zero, one enforces the strict gauge condition $\mathbf{n}(\nabla) \cdot \mathbf{A}(x) = 0$, but by integrating over it with a Gaussian weight one can generate actions of the form (21.21). This covers all the examples we have met so far. The arguments easily generalize to other gauges.

We start from the continuation of the action (21.34) to imaginary time,

$$\mathcal{S}(\mathbf{A}, \mathbf{J}) = \int d^d x \left[\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) - \mathbf{J}(x) \cdot \mathbf{A}(x) \right]. \quad (21.53)$$

In the Euclidean formulation, the expression (21.47) becomes

$$\mathcal{Z}(\mathbf{J}) = \int [dA] \prod_x \delta(A_d(x)) e^{-\mathcal{S}(\mathbf{A}, \mathbf{J})}. \quad (21.54)$$

We insert the expression

$$\int [d\chi(x)] \prod_x \delta[\mathbf{n}(\nabla) \cdot (\nabla\chi(x) + \mathbf{A}(x)) - \nu(x)] = \text{const.}, \quad (21.55)$$

where $\nu(x)$ is an arbitrary external field, inside expression (21.54)

$$\mathcal{Z}(\mathbf{J}) \propto \int [dA d\chi] \prod_x \delta(A_d(x)) \delta[\mathbf{n}(\nabla) \cdot (\nabla\chi(x) + \mathbf{A}(x)) - \nu(x)] e^{-\mathcal{S}(\mathbf{A}, \mathbf{J})}. \quad (21.56)$$

We change of variables, $\mathbf{A} \mapsto \mathbf{A}'$, setting (a gauge transformation)

$$\mathbf{A}(x) = \mathbf{A}'(x) - \nabla\chi(x).$$

Since the current \mathbf{J} is conserved, only the δ -functions are affected ($\partial_d \equiv \partial/\partial x_d$):

$$\delta(A_d) \delta[\mathbf{n}(\nabla) \cdot (\nabla\chi(x) + \mathbf{A}(x)) - \nu(x)] \mapsto \delta(A_d(x) - \partial_d\chi(x)) \delta[\mathbf{n}(\nabla) \cdot \mathbf{A}(x) - \nu(x)].$$

The integration over χ can again be performed:

$$\int [d\chi] \prod_x \delta(A_d(x) - \partial_d\chi(x)) = \text{const.}, \quad (21.57)$$

and, therefore,

$$\mathcal{Z}(\mathbf{J}) = \int [dA] \delta[\mathbf{n}(\nabla) \cdot \mathbf{A}(x) - \nu(x)] e^{-\mathcal{S}(\mathbf{A}, \mathbf{J})}. \quad (21.58)$$

Since, by construction, the result does not depend on $\nu(x)$, one can either set $\nu(x)$ to 0, or integrate over $\nu(x)$ with, for example, the Gaussian measure

$$[d\rho(\nu)] = [d\nu(x)] \exp\left[-\frac{1}{2} \int d^d x \nu^2(x)\right]. \quad (21.59)$$

One then obtains

$$\mathcal{Z}(\mathbf{J}) = \int [dA] \exp [-\mathcal{S}(\mathbf{A}, \mathbf{J}) - \mathcal{S}_{\text{gauge}}(\mathbf{A})], \quad (21.60)$$

with

$$\mathcal{S}_{\text{gauge}}(\mathbf{A}) = \frac{1}{2} \int d^d x [\mathbf{n}(\nabla) \cdot \mathbf{A}(x)]^2. \quad (21.61)$$

Specializing to $\mathbf{n} = \xi^{-1/2} \nabla$, we note that we have demonstrated the equivalence between *Weyl's or the temporal gauge* $A_d = 0$ and the covariant gauges (21.21). By contrast, if we choose $n_d = 0$ and $\mathbf{n}_\perp \equiv \partial_\perp$, and set $\boldsymbol{\nu} = 0$, we find Coulomb's gauge.

The propagator. From the calculation of the field integral (21.60) for an arbitrary source J_μ , one infers the gauge field propagator (and connected two-point function), in the Fourier representation

$$\widetilde{W}_{\mu\nu}^{(2)}(k) = \frac{1}{k^2} \left[\delta_{\mu\nu} - \frac{(k_\mu n_\nu + k_\nu n_\mu)}{k \cdot n} + \frac{(n^2 + k^2) k_\mu k_\nu}{(k \cdot n)^2} \right]. \quad (21.62)$$

In the temporal gauge, it becomes

$$\widetilde{W}_{ij}^{(2)} = \frac{1}{k^2} \left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}_\perp^2} \right) + \frac{1}{k_d^2} \frac{k_i k_j}{\mathbf{k}_\perp^2} = \frac{1}{k^2} \left(\delta_{ij} + \frac{k_i k_j}{k_d^2} \right), \quad (21.63)$$

in which \mathbf{k}_\perp is the ‘space’ part of \mathbf{k} , that is, its projection on \mathbb{R}^{d-1} . This propagator, as in the case of Coulomb's gauge, has a large momentum behaviour which is not uniform and thus, in contrast with the covariant gauges, leads to theories that are not explicitly renormalizable in four dimensions. Moreover, its longitudinal part has a double pole at $k_d = 0$, which also requires some regularization.

Remark. The strict gauge condition is recovered in the limit $|n_\mu| \rightarrow \infty$, which exists for the propagator but not for the action. A way of writing the limiting action explicitly is to introduce an additional field $\lambda(x)$, which implements the gauge condition as a functional δ -function in Fourier representation,

$$\mathcal{S}_{\text{gauge}}(A) = \int d^d x \lambda(x) \mathbf{n}(\nabla) \cdot \mathbf{A}(x), \quad \text{with } \lambda(x) \text{ imaginary.} \quad (21.64)$$

21.6.1 Interpretation: The Faddeev–Popov quantization

The result we have obtained has a geometric interpretation, which can be justified rigorously only in the lattice approximation on a finite lattice (see Section 21.8). The problem of a gauge-invariant QFT is that locality requires an action with redundant degrees of freedom or, equivalently, that the local gauge-invariant action does not provide a dynamics to the gauge degrees of freedom. Therefore, we supply them with a stochastic dynamics in the sense of Chapter 35. We express the gauge field A_μ in terms of a gauge field B_μ , projection of A_μ on a gauge section (*i.e* satisfying a gauge condition), and a gauge transformation χ :

$$\mathbf{A}(x) = \mathbf{B}(x) + \nabla \chi(x). \quad (21.65)$$

Given the section, we assume that this decomposition is unique. Gauge invariance implies that the gauge action depends only on B_μ and specifies its dynamics.

To $\chi(x)$ we impose, for example,

$$\nabla^2 \chi(x) + \nabla \cdot \mathbf{B}(x) - \nu(x) = 0, \quad (21.66)$$

in which $\nu(x)$ is a stochastic field with a given probability distribution.

We introduce the constraint in the field integral (see Section 26.1) [219]. This also generates the determinant of the functional derivative of equation (21.66) with respect to the field $\chi(x)$. Here, this operator is just ∇^2 and, therefore, the determinant is a field-independent constant, which disappears in the normalization of the field integral.

The field integral becomes

$$\mathcal{Z} = \int [dB d\chi] \delta [\nabla^2 \chi(x) + \nabla \cdot \mathbf{B}(x) - \nu(x)] \exp [-\mathcal{S}(\mathbf{B})]. \quad (21.67)$$

The functional measure $[dB d\chi]$ is the decomposition of the flat measure $[dA]$ into a product of measures on \mathbf{B} and χ . The action $\mathcal{S}(\mathbf{B})$ is the gauge-invariant action $\mathcal{S}(\mathbf{A})$ in which equation (21.65) has been used. We now recognize that the whole expression can be rewritten in terms of A_μ as

$$\mathcal{Z} = \int [d] \delta [\nabla \cdot \mathbf{A}(x) - \nu(x)] \exp [-\mathcal{S}(\mathbf{A})]. \quad (21.68)$$

Moreover, since the result of the field integration does not depend on the dynamics of $\chi(x)$, the result does not depend on the field $\nu(x)$ either and one can integrate over $\nu(x)$ with, for example, the Gaussian measure (21.59).

21.7 Gauge symmetry and parallel transport

The gauge symmetry of the fermion action has a geometric interpretation, which is worth noting. It can be seen as a resulting from the replacement of the derivative ∂_μ of the free fermion theory by the *covariant derivative* $\partial_\mu + ieA_\mu$, which makes the cancellation of the transformation of the gauge field with the contribution of the derivative possible. We are thus reminded of the concepts of covariant derivative, affine connection, curvature and parallel transport introduced for Riemannian manifolds (see Chapter 28). In particular, we note the similarity with the rotations of the local frame considered in Section 28.5.1. The difference is that the vectors that here are parallel-transported do not belong to the space tangential to the manifold (which is flat), but are vectors for Lie group transformations.

Parallel transport. We want to construct gauge-invariant quantities. For quantities product of fields at the same point, like $\bar{\psi}(x)\psi(x)$, global invariance implies gauge invariance. This is no longer the case for products of fields at different points or involving derivatives. The problem can then be solved by introducing parallel transport.

The parallel transporter $U(C_{xy})$, where C_{xy} is a continuous piecewise differentiable, curve oriented from end-points y to x (see Section 28.2 for details), is an element of the $U(1)$ group. Since the group is Abelian, it can be expressed explicitly in terms of a line integral involving the connection or gauge field $A_\mu(x)$ (unlike in the non-Abelian case, see Section 22.1) as

$$U(C_{xy}) = \exp \left[-ie \oint_{C_{xy}} \sum_\mu A_\mu(s) ds_\mu \right]. \quad (21.69)$$

A connection is expected to have three indices, one, here μ , which refers to the manifold, and two which refer to the group space. Since the group $U(1)$ corresponds to multiplication by complex numbers, they are omitted.

One verifies that the transformation of $U(C_{xy})$ induced by a gauge transformation (21.28), with gauge function $\theta(x)$, of the connection A_μ , is given by

$$U(C_{xy}) = U'(C_{xy}) \exp \left[i \oint_{C_{xy}} \sum_\mu \partial_\mu \theta(s) ds_\mu \right] = U'(C_{xy}) \exp [i(\theta(x) - \theta(y))].$$

Thus, if a field $\psi(x)$ transforms as (equation (21.29)) $\psi(x) \mapsto e^{i\theta(x)} \psi(x)$, the product $U(C_{xy})\psi(y)$ transforms like

$$U(C_{xy})\psi(y) \mapsto e^{i\theta(x)} U(C_{xy})\psi(y).$$

In the limit $y \rightarrow x$, setting $y_\mu = x_\mu + dx_\mu$, the product $U(C_{xy})\psi(y)$ has the expansion,

$$U(C_{xy})\psi(y) - \psi(x) = \sum_\mu dx_\mu [\partial_\mu \psi(x) + ieA_\mu(x)] + o(dx_\mu) \equiv \sum_\mu dx_\mu D_\mu \psi(x),$$

where we have defined the *covariant derivative*

$$D_\mu = \partial_\mu + ieA_\mu(x), \quad (21.70)$$

and we define, for later purpose, $\not{D} \equiv \sum_\mu \gamma_\mu D_\mu$.

It follows that the curvature tensor (which is the electromagnetic field in QED) is

$$F_{\mu\nu}(x) = \frac{1}{ie} [D_\mu, D_\nu] = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x). \quad (21.71)$$

The parallel transporter around a closed curve is gauge invariant. According to Stokes's theorem, for a simple closed curve, $U(C)$ can be expressed in terms of a surface integral over a surface D that has the curve C as a boundary:

$$U(C) = \exp \left[-\frac{ie}{2} \int_D \sum_{\mu,\nu} dx_\mu \wedge dx_\nu F_{\mu\nu}(x) \right], \quad \text{with } \partial D = C. \quad (21.72)$$

The expression relates the curvature tensor $F_{\mu\nu}$ to parallel transport around a closed curve. This observation will be useful for constructing lattice gauge theories (see Section 21.8.2 and Chapter 25).

Differential forms. In equation (21.72), we have used the notation of differential forms: dx_μ is an element of an exterior algebra, $\sum_\mu A_\mu(x)dx_\mu$ represents a one-form, $dx_\mu \wedge dx_\nu F_{\mu\nu}(x)$ a two-form, with \wedge denoting the wedge product (see also Section 28.1.3).

21.8 Perturbation theory: Regularization

For most of the section, we consider an action for massive QED with fermions in a covariant gauge, of the form (21.30). With the propagator (21.22), power counting is the same as for a scalar field. Therefore, one can construct interacting renormalizable theories, from the viewpoint of power counting, for dimensions $d \leq 4$. Here, we focus on dimension 4. However, because gauge invariance is essential for the physical consistency of the theory, it must be preserved by the regularization and renormalization. Therefore, in a first step, gauge-invariant regularizations are required.

Dimensional regularization. Dimensional regularization is defined in Section 10.1. It is well-suited to perturbative calculations in QED. Examples are given in Section 21.12.1. However, in even dimensions, it leads to problems in the case of chiral gauge theories, because γ_S has no proper dimensional continuation (see Section 23.6).

21.8.1 Momentum cut-off regularization

In this chapter, the special problems generated by chiral fermions are not met, and we can calculate with dimensional regularization. However, for later purpose, it is instructive to also discuss momentum cut-off or Pauli–Villars’s regularization, especially in the case of fermion matter.

The problem of matter in presence of a gauge field can be decomposed into two steps: first matter in an external gauge field and then the integration over the gauge field.

Charged fermions in a gauge background. From the point of view of momentum regularization, a new problem arises in gauge theories: gauge invariance implies that only covariant derivatives can be used. The regularized action (21.26) in a gauge background then takes the form ($\not{D} \equiv \sum_\mu \gamma_\mu D_\mu$)

$$\mathcal{S}_{F,\text{reg.}}(\bar{\psi}, \psi, \mathbf{A}) = \int d^d x \bar{\psi}(x) (M + \not{D}) \prod_{i=1}^s (1 - \not{D}^2/M_i^2) \psi(x), \quad \frac{M_i}{\Lambda} = O(1), \quad (21.73)$$

which replaces expression (12.54). Up to this point, the regularization, unlike dimensional or lattice regularizations, preserves a possible chiral symmetry for $M = 0$.

However, the higher order derivatives of the regularization, while they improve the large momentum behaviour of the fermion propagator, generate new, more singular, gauge interactions, and it is no longer clear whether the theory can be rendered finite.

Correlation functions in the gauge background then are generated by

$$\begin{aligned} \mathcal{Z}(\bar{\eta}, \eta; \mathbf{A}) \\ = \int [d\psi(x) d\bar{\psi}(x)] \exp \left[-\mathcal{S}_F(\bar{\psi}, \psi, \mathbf{A}) + \int d^d x (\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)) \right], \end{aligned} \quad (21.74)$$

where $\bar{\eta}, \eta$ are Grassmann sources. Integrating over fermions explicitly, one obtains

$$\begin{aligned} \mathcal{Z}(\bar{\eta}, \eta; \mathbf{A}) &= \mathcal{Z}_0(\mathbf{A}) \exp \left[- \int d^d x d^d y \bar{\eta}(y) \Delta_F(\mathbf{A}; y, x) \eta(x) \right], \\ \mathcal{Z}_0(\mathbf{A}) &= \mathcal{N} \det \left[(M + \not{D}) \prod_i (1 - \not{D}^2/M_i^2) \right], \end{aligned} \quad (21.75)$$

where \mathcal{N} is a gauge field-independent normalization, and $\Delta_F(\mathbf{A}; y, x)$ the fermion propagator in an external gauge field.

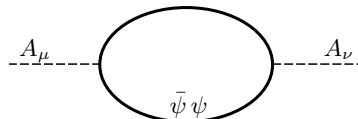


Fig. 21.1 One-loop contribution to the gauge field two-point function

The fermion determinant. The first problem involves the determinant that generates closed fermion loops in a gauge background (like in Fig. 21.1). Using $\ln \det = \text{tr} \ln$, one finds

$$\ln \mathcal{Z}_0(\mathbf{A}) = \text{tr} \ln (M + \not{D}) + \sum_r \text{tr} \ln (1 - \not{D}^2/M_r^2) - (\mathbf{A} = 0).$$

In the most relevant dimension, $d = 4$, we can then use the anticommutation of γ_5 with \not{D} ,

$$\det(\not{D} + M) = \det \gamma_5 (\not{D} + M) \gamma_5 = \det(M - \not{D}),$$

$$\ln \mathcal{Z}_0(\mathbf{A}) = \frac{1}{2} \text{tr} \ln (M^2 - \not{D}^2) + \sum_i \text{tr} \ln (1 - \not{D}^2/M_i^2) - (\mathbf{A} = 0).$$

We note that the regularization does not change the power counting of the determinant and, therefore, of one-loop diagrams of the form of fermion closed loops with external gauge fields, a problem that requires an additional regularization. This analysis signals a difficulty in constructing in general a regularized gauge-invariant expression for the determinant of operators of the form $\not{D} + M$ in the continuum and at fixed dimension (see Section 23.6).

The fermion determinant can be further regularized by adding boson regulator fields with fermion spin to the action and, therefore, a propagator similar to Δ_F but with different masses

$$\mathcal{S}_B(\bar{\phi}, \phi; \mathbf{A}) = \int d^d x \bar{\phi}(x) (M_0^B + \not{D}) \prod_{i=1} (1 - \not{D}^2/(M_i^B)^2) \phi(x), \quad M_i^B = O(\Lambda). \quad (21.76)$$

Section 30.3 provides an explicit example.

The integration over the boson fields $\bar{\phi}, \phi$ adds to $\ln \mathcal{Z}_0$ the quantity

$$\delta \ln \mathcal{Z}_0(\mathbf{A}) = -\frac{1}{2} \text{tr} \ln ((M_0^B)^2 - \not{D}^2) - \sum_{i=1} \text{tr} \ln (1 - \not{D}^2/(M_i^B)^2) - (\mathbf{A} = 0).$$

Expanding in inverse powers of \not{D} , one adjusts the masses to cancel as many powers as possible. However, the unpaired initial fermion mass M is the source of a problem. The corresponding determinant can only be regularized with an unpaired boson $M_0^B = O(\Lambda)$. In the chiral limit $M = 0$, we have two options: either we give a chiral charge to the boson field and the mass M_0^B breaks chiral symmetry, or we leave it invariant in a chiral transformation. Then, after a chiral gauge transformation, one obtains the determinant of the operator

$$e^{i\theta(x)\gamma_5} \not{D} e^{i\theta(x)\gamma_5} (\not{D} + M_0^B)^{-1}.$$

For $\theta(x)$ constant $e^{i\theta\gamma_5}$ anticommutes with \not{D} and cancels. Otherwise, a non-trivial contribution remains. The analysis thus indicates possible difficulties with space-dependent chiral transformations.

Since the problem reduces to the study of a determinant in an external background, we study it directly, starting with Section 23.6. We examine whether it is possible to define some regularized form in a way consistent with chiral symmetry. The one-loop renormalized diagrams can then be inserted in a general diagram already partially regularized by the preceding cut-off methods.

The gauge field propagator. After integration over the gauge field, diagrams constructed from $\Delta_F(\mathbf{A}; y, x)$ belong to loops with gauge field propagators (Fig. 21.2), and, therefore, can be rendered finite if the gauge field propagator can be improved.

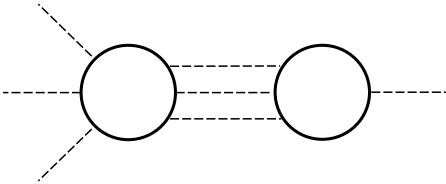


Fig. 21.2 Example of a multi-loop diagram

For the free gauge action in a covariant gauge, simple derivatives can be used, because in an Abelian theory the gauge field is neutral. The tensor $F_{\mu\nu}$, which is gauge invariant, and $\nabla \cdot \mathbf{A}$, which is generated by the arbitrary action for the scalar $\chi(x)$ of equation (21.16), can be multiplied by polynomials of ∇^2 . Therefore, the large momentum behaviour of the gauge field propagator can be arbitrarily improved.

Scalar matter. In the case of scalar matter, a similar analysis holds. For multi-loop diagrams, scalar self-interaction vertices can be added, but then the number of matter propagators exceeds the number of gauge field vertices and again the diagrams can be rendered superficially convergent.

Finally, the determinant $\det D_\mu^2$ generated by integrating over charged scalar fields in a gauge background can be regularized by Schwinger's proper time method (see Section A21.3). In this way, the determinant is expressed in terms of the evolution operator corresponding to a non-relativistic Hamiltonian in a magnetic field.

21.8.2 Lattice regularization

The construction of a lattice regularized form of a gauge theory is directly based on the notion of parallel transport and the geometric interpretation of the gauge field as a connection [221]. Since, on the lattice, points are split, the gauge field has to be replaced by *link variables* corresponding to parallel transport along links of the lattice (see Section 21.7 and Chapter 25 for a detailed discussion).

A link variable $U(x, y)$ is an element of the $U(1)$ group, lattice analogue of the parallel transporter $U(C)$ defined by equation (21.69), the curve C being the link joining the site x to the neighbouring site y on the lattice. It can be parametrized in terms of an angle θ_{xy} , and is such that

$$U(x, y) \equiv U_{xy} = e^{i\theta_{xy}} = (U_{yx})^{-1}. \quad (21.77)$$

The measure of integration over the gauge variables is the group invariant measure, that is, the flat measure $d\theta_{xy}$.

Lattice-dependent $U(1)$ group transformations are then the lattice equivalents of the gauge transformations of the continuum theory.

Gauge invariant fermion interactions on the lattice have the form (but this leads to subtle issues for the fermion lattice propagator, see Section 12.9)

$$\sum_\mu \bar{\psi}(x + an_\mu) \gamma_\mu U(x + an_\mu, x) \psi(x), \quad (21.78)$$

where we have denoted by n_μ the unit vector in μ direction, and by a the lattice spacing.

Plaquette action. We have shown in Section 21.7 that the curvature tensor is associated with parallel transport around a closed curve. This indicates that a regularized form of the gauge action $\int d^d x \sum_{\mu,\nu} F_{\mu\nu}^2$ is the product of link variables around a closed curve on the lattice, the simplest one on an hypercubic lattice being a square, forming a *plaquette*.

Such a product is clearly gauge invariant. This leads to the famous *plaquette action* (see also Section 25.2):

$$\sum_{\text{all plaquettes}} U_{xy}U_{yz}U_{zt}U_{tx}. \quad (21.79)$$

We have denoted symbolically by x , y , z , and t four sites forming a square on the lattice.

A typical gauge-invariant lattice action corresponding to the continuum action of a gauge field coupled to fermions then has the form

$$\mathcal{S}(U, \bar{\psi}, \psi) = \beta \sum_{\text{plaquettes}} U_{xy}U_{yz}U_{zt}U_{tx} - \kappa \sum_{\text{links}} \bar{\psi}_y \gamma_{yx} U_{yx} \psi_x - \sum_{\text{sites}} M \bar{\psi}_x \psi_x. \quad (21.80)$$

We have denoted by x , y , and so on the lattice sites, $\gamma_{xy} \equiv \sum_\mu (x-y)_\mu \gamma_\mu$ and β and κ are the coupling constants.

Finally, it is possible to add a term of the form $\sum \cos \theta_{xy}$ to the gauge-invariant lattice action to give a mass to the vector field.

Since the Chapter 25 is devoted to lattice gauge theories, we refer to this chapter for a thorough discussion of lattice gauge theory.

21.9 WT identities and renormalization

In this section, for convenience, *we assume dimensional regularization*.

In gauge theories, WT identities [222] play an essential role, because it is necessary to maintain the gauge symmetry throughout the renormalization process, in order to ensure that the theory, although it is not explicitly unitary, is equivalent to a unitary theory for the ‘physical’, that is, gauge-invariant observables (this includes S -matrix elements).

We derive them for the action (21.30) to which we add sources $J_\mu(x)$, $\eta(x)$, and $\bar{\eta}(x)$ for the fields $A_\mu(x)$, $\bar{\psi}(x)$, and $\psi(x)$, respectively. The action then becomes

$$\Sigma(\mathbf{A}, \bar{\psi}, \psi,) = \mathcal{S}_\xi(\mathbf{A}, \bar{\psi}, \psi) - \int d^d x [\mathbf{J}(x) \cdot \mathbf{A}(x) + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)].$$

The corresponding generating functional of correlation functions then is

$$\mathcal{Z}(J, \bar{\eta}, \eta) = \int [dA d\psi, d\bar{\psi}] \exp [-\Sigma(\mathbf{A}, \bar{\psi}, \psi)]. \quad (21.81)$$

In the field integral, we make an infinitesimal change of variables of the form of the gauge transformations (21.28) and (21.29). The measure is invariant and the variation of Σ comes from the \mathbf{A} -mass term, the gauge-fixing term and the sources:

$$\begin{aligned} \delta \Sigma(\mathbf{A}, \bar{\psi}, \psi) &= -\frac{1}{e} \int d^d x \Lambda(x) \{ [(\nabla^2/\xi - m^2) \nabla \cdot \mathbf{A}(x) + \nabla \cdot \mathbf{J}(x)] \\ &\quad + i [\bar{\eta}(x)\psi(x) - \bar{\psi}(x)\eta(x)] \}. \end{aligned} \quad (21.82)$$

This leads, following the usual arguments (see *e.g.* Section 7.5.3) to an equation for the generating functionals $\mathcal{Z}(J, \bar{\eta}, \eta)$ and thus, $\mathcal{W}(J, \bar{\eta}, \eta) = \ln \mathcal{Z}(J, \bar{\eta}, \eta)$, which has the form

$$\begin{aligned} &\left\{ (m^2 - \nabla^2/\xi) \sum_\mu \partial_\mu \frac{\delta}{\delta J_\mu(x)} - ie \left[\bar{\eta}(x) \frac{\delta}{\delta \bar{\eta}(x)} - \eta(x) \frac{\delta}{\delta \eta(x)} \right] \right\} \mathcal{W}(J, \bar{\eta}, \eta) \\ &= \nabla \cdot \mathbf{J}(x). \end{aligned} \quad (21.83)$$

The term linear in J_μ leads, after Fourier transformation, to the equation for the gauge field two-point function,

$$\sum_\mu k_\mu \widetilde{W}_{\mu\nu}^{(2)}(k) = \xi \frac{k_\nu}{k^2 + \xi m^2}. \quad (21.84)$$

More generally, expanding in powers of η and $\bar{\eta}$ and setting $J_\mu = 0$, one finds

$$\begin{aligned} (k^2/\xi + m^2) \sum_\mu k_\mu \widetilde{W}_\mu^{(2n+1)}(k; p_1, \dots, p_n; q_1, \dots, q_n) \\ = e \sum_{i=1}^n \left[\widetilde{W}^{(2n)}(p_1, \dots, p_i + k, \dots, p_n; q_1, \dots, q_n) \right. \\ \left. - \widetilde{W}^{(2n)}(p_1, \dots, p_n; q_1, \dots, q_i + k, \dots, q_n) \right], \end{aligned} \quad (21.85)$$

in which k is the gauge-field momentum, and p_i and q_i the momenta of ψ and $\bar{\psi}$ fields, respectively. Additional external gauge fields do not modify the identities.

Equation (21.83) is a linear first-order partial differential equation. The Legendre transformation is simple and yields an equation for the generating functional $\Gamma(\mathbf{A}, \bar{\psi}, \psi)$ of vertex functions,

$$(\nabla^2/\xi - m^2) \nabla \cdot \mathbf{A}(x) + \sum_\mu \partial_\mu \frac{\delta \Gamma}{\delta A_\mu(x)} + ie \left[\psi(x) \frac{\delta \Gamma}{\delta \psi(x)} - \bar{\psi}(x) \frac{\delta \Gamma}{\delta \bar{\psi}(x)} \right] = 0. \quad (21.86)$$

It can be verified that the equations (21.83) and (21.86) have the same content as the quantum equations of motion of the field χ of Section 21.1.

The general solution of equation (21.86) can be written as

$$\Gamma(A_\mu, \bar{\psi}, \psi) = \Gamma_{\text{sym.}}(A_\mu, \bar{\psi}, \psi) + \frac{1}{2} \int \left[m^2 \mathbf{A}^2(x) + (\nabla \cdot \mathbf{A}(x))^2/\xi \right] d^d x, \quad (21.87)$$

where $\Gamma_{\text{sym.}}$ is gauge invariant.

Renormalization. We perform a loop expansion of the functional Γ . Because the equation (21.86) is linear, the tree approximation satisfies the inhomogeneous equation while all higher order terms satisfy the homogeneous equation.

Denoting by Γ_ℓ the ℓ -loop contribution to Γ , for $\ell > 0$, one finds,

$$\sum_\mu \partial_\mu \frac{\delta \Gamma_\ell}{\delta A_\mu(x)} + ie \left[\psi(x) \frac{\delta \Gamma_\ell}{\delta \psi(x)} - \bar{\psi}(x) \frac{\delta \Gamma_\ell}{\delta \bar{\psi}(x)} \right] = 0.$$

Therefore, the generating functional Γ_ℓ of ℓ -loop vertex functions is gauge invariant. The singular part of the Laurent expansion in $\varepsilon = 4 - d$ of Γ_ℓ is also gauge invariant, which means that the divergent part $\Gamma_\ell^{\text{div.}}$ is gauge invariant.

The conclusion is that the action can be completely renormalized by adding gauge-invariant counter-terms. As in the case of the linear symmetry breaking in Section 13.3, one can say that the terms that break the gauge invariance, the gauge-field mass and the gauge-fixing terms, are not renormalized, since they are not modified by counter-terms.

The full renormalized action can be written as

$$\begin{aligned} \mathcal{S}_r(A_\mu, \bar{\psi}, \psi) = \int d^d x \left[\frac{1}{4} Z_A \sum_{\mu, \nu} F_{\mu\nu}^2(x) + \frac{1}{2} m^2 \mathbf{A}^2(x) + \frac{1}{2\xi} (\nabla \cdot \mathbf{A}(x))^2 \right. \\ \left. - Z_\psi \bar{\psi}(x) (\not{D} + M_0) \psi(x) \right], \end{aligned} \quad (21.88)$$

where Z_A is the gauge field, Z_ψ and M_0 are the ψ field and mass renormalization constants.

We now introduce the bare (microscopic) fields $\psi^0, \bar{\psi}^0, A_\mu^0$ and charge e_0 ,

$$\psi^0 = Z_\psi^{1/2} \psi, \quad \bar{\psi}^0 = Z_\psi^{1/2} \bar{\psi}, \quad A_\mu^0 = Z_A^{1/2} A_\mu, \quad e_0 = Z_e^{1/2} e. \quad (21.89)$$

We note that the covariant derivative is not renormalized:

$$\partial_\mu + ie_0 A_\mu^0 = \partial_\mu + ie A_\mu. \quad (21.90)$$

This implies the important relation

$$Z_A Z_e = 1. \quad (21.91)$$

Gauge invariance relates the renormalization of the charge and the gauge field.

Moreover, since the quantities $F_{\mu\nu}(x)$ and $\bar{\psi}(x)\psi(x)$ are gauge invariant, their correlation functions are the same in the unitary and ξ gauges. Therefore, Z_A and M_0 (but not Z_ψ) can be chosen independent of ξ .

21.10 Gauge dependence: The fermion two-point function

In most of the section, we assume dimensional regularization.

Some insight into the physical properties of the theory can be gained from a study of the gauge-dependence of correlation functions. In the covariant gauges of the action (21.30), this amounts to studying the dependence on the parameter ξ .

As an example, we investigate the gauge-dependence of the fermion two-point function, but the method we use has a simple generalization to other correlation functions.

The fermion two-point function is given by

$$\widetilde{W}_\xi^{(2)}(u, v) \equiv \langle \bar{\psi}(u)\psi(v) \rangle = \mathcal{Z}^{-1} \int [d\mathbf{A} d\psi d\bar{\psi}] \bar{\psi}(u)\psi(v) e^{-S_\xi(\mathbf{A}, \bar{\psi}, \psi)}, \quad (21.92)$$

where S_ξ is the action (21.30), \mathcal{Z} the partition function (or vacuum amplitude), which is gauge independent and, thus, ξ independent.

We now invert the algebraic transformations of Section 21.2.1. We introduce an uncoupled scalar field χ and add a mass term to the action:

$$S_\xi(\mathbf{A}, \bar{\psi}, \psi) \mapsto S_\xi(\mathbf{A}, \bar{\psi}, \psi, \chi) = S_\xi(\mathbf{A}, \bar{\psi}, \psi) + \frac{1}{2}\xi m^2 \int d^d x \chi^2(x).$$

A direct integration over χ yields a constant factor that cancels in the ratio (21.92).

Instead, we change variables, shifting $\chi(x)$ as,

$$\chi(x) \mapsto \chi(x) + \frac{i}{\xi m} \nabla \cdot \mathbf{A}(x).$$

After an integration by parts, the action becomes

$$S(\mathbf{A}, \bar{\psi}, \psi, \chi) = S_\xi(\mathbf{A}, \bar{\psi}, \psi) + \int d^d x \left[\frac{1}{2}\xi m^2 \chi^2(x) - im \mathbf{A}(x) \cdot \nabla \chi(x) - \frac{1}{2} (\nabla \cdot \mathbf{A}(x))^2 / \xi \right].$$

The last term cancels the gauge fixing term, and the second term in the right-hand side can be eliminated by the gauge transformation,

$$\mathbf{A}(x) \mapsto \mathbf{A}(x) + i \nabla \chi(x) / m, \quad \psi(x) \mapsto \psi(x) e^{ie\chi(x)/m}, \quad \bar{\psi}(x) \mapsto \bar{\psi}(x) e^{-ie\chi(x)/m}.$$

The field integral representation of the two-point function then factorizes and leads to the relation

$$\widetilde{W}_\xi^{(2)}(u, v) = \widetilde{W}_\infty^{(2)}(u, v) \mathcal{Z}_\chi^{-1} \int [d\chi] e^{-S_\xi(\chi)}, \quad (21.93)$$

in which $\widetilde{W}_\infty^{(2)}(u, v)$ is calculated with the unitary action (21.27) of massive QED, and

$$S_\xi(\chi) = \int d^d x \left[-\frac{1}{2} (\nabla \chi(x))^2 - \frac{1}{2} \xi m^2 \chi^2(x) \right] + ie\chi(v) - ie\chi(u). \quad (21.94)$$

The χ integration can be performed by noting that

$$ie\chi(v) - ie\chi(u) = \int d^d x J(x)\chi(x), \text{ with } J(x) = ie \left[\delta^{(d)}(x-v) - \delta^{(d)}(x-u) \right].$$

Thus, in terms of the χ -propagator Δ_χ ,

$$\begin{aligned} \int [d\chi] e^{-S_\xi(\chi)} &= \exp \left[-\frac{1}{2} \int d^d x d^d y J(x) \Delta_\chi(\xi, x-y) J(y) \right] \\ &= \exp [-e^2 (\Delta_\chi(\xi, 0) - \Delta_\chi(\xi, u-v))]. \end{aligned} \quad (21.95)$$

The χ -propagator has the Fourier representation

$$\Delta_\chi(\xi, x) = \frac{1}{(2\pi)^d} \int \frac{d^d p e^{ipx}}{p^2 + \xi m^2} \quad (21.96)$$

and, thus,

$$\Delta_\chi(\xi, 0) - \Delta_\chi(\xi, x) = \frac{1}{(2\pi)^d} \int \frac{d^d p (1 - e^{ipx})}{p^2 + \xi m^2}.$$

In particular, we infer the ratio of the bare two-point functions for two different values of ξ . For example,

$$\widetilde{W}_\xi^{(2)}(u, v) = \exp \left[\frac{\xi e^2}{(2\pi)^d} \int d^d p \frac{e^{ip(u-v)} - 1}{p^2 (p^2 + \xi m^2)} \right] \widetilde{W}_{(\xi=0)}^{(2)}(u, v). \quad (21.97)$$

Note that, for $d = 4$, this relation has a $m = 0$ limit.

Only the second term in the exponential is divergent, while the first one contains the non-physical χ poles. Introducing the fermion renormalization constant Z_ψ , we obtain the ratio of renormalization constants for two values of ξ (up to additional finite renormalizations, required to avoid IR divergences in the limit $m = 0$),

$$Z_\psi(\xi) = Z_\psi(0) \exp \left[-\frac{\xi e^2}{(2\pi)^d} \int \frac{d^d p}{p^2 (p^2 + \xi m^2)} \right]. \quad (21.98)$$

The relation between renormalized functions then becomes

$$\left[\widetilde{W}_\xi^{(2)}(u, v) \right]_r = \exp \left[\frac{\xi e^2}{(2\pi)^d} \int d^d p \frac{e^{ip(u-v)}}{p^2 (p^2 + \xi m^2)} \right] \left[\widetilde{W}_{(\xi=0)}^{(2)}(u, v) \right]_r. \quad (21.99)$$

Equation (21.97) has additional consequences: no other ξ -dependent renormalization is needed. One recovers the property that the renormalization constant Z_e and thus, from equation (21.91), also Z_A are gauge independent. In the renormalized action (21.88), only Z_ψ is gauge dependent.

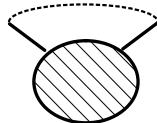


Fig. 21.3 Order e^2 contribution. The full lines correspond to the fermion two-point function, the dotted line to the non-physical scalar field propagator

General correlation functions. All ψ -field correlation functions have a similar structure. The ξ -dependence, which comes only from the χ integration, factorizes, and can be evaluated (see Section A21.2).

This also applies to correlation functions involving, in addition, the gauge-invariant field $F_{\mu\nu}$, which does not introduce any additional gauge dependence.

Finally, since the fermion character plays little role in the derivation, a similar analysis is valid for charged scalar bosons.

Unitarity. An expansion of $\widetilde{W}_\xi^{(2)}$ in powers of ξe^2 of the factor in equation (21.99) immediately shows that only the first term, which is ξ independent, has the physical fermion pole (see Fig. 21.3 for the order e^2 contribution).

This argument generalizes to $\bar{\psi}, \psi, F_{\mu\nu}$ correlation functions. As a consequence of factorization, after renormalization only the first term in an expansion of the factor in powers of ξe^2 , which is ξ independent, contributes to the mass-shell limit. In this limit, the ξ -dependence of the singular part of correlation functions reduces to a multiplicative renormalization. Therefore, the properly normalized S -matrix elements are gauge independent. Being gauge independent, they cannot have the ξ -dependent singularities generated by the field χ . The full S -matrix, in the subspace of physical states, is thus unitary. Note that, by using WT identities, one can also directly prove that the contribution of χ -field cancels in the intermediate states in the extended unitarity relations.

Gauge-invariant operators. We have examined the gauge-dependence of S -matrix elements. From the point of view of correlation functions, the only gauge-independent quantities are expectation values of products of gauge-invariant operators, that is, local polynomials in the fields, invariant under the transformations (21.28) and (21.29).

The simplest such operators are $F_{\mu\nu}$, which selects the transverse part of the gauge field, $\bar{\psi}(x)\psi(x)$ or, more generally, $\bar{\psi}(x)\Gamma_A\psi(x)$ in which the matrix Γ_A is any element of the algebra of γ matrices. Equation (21.97) explicitly shows the mechanism which makes the correlation functions of ψ gauge dependent while $\bar{\psi}(x)\psi(x)$ is gauge independent. When in the product $\bar{\psi}(x)\psi(y)$, y approaches x , the additional gauge-dependent renormalization needed to make the product $\bar{\psi}(x)\psi(x)$ finite cancels the gauge-dependent part of the fermion field renormalization.

To study the renormalization properties of gauge-invariant operators, one can add to the action sources for them. The form of the WT identities is not modified. The arguments of Section 21.9 are still valid: the counter-terms are gauge invariant and can be chosen gauge independent. This also proves that gauge-invariant operators mix under renormalization only with gauge-invariant operators of lower or equal canonical dimensions.

Non-gauge-invariant correlation functions in the unitary gauge. With the original action (21.27), all correlation functions are ‘physical’, but the theory is not renormalizable. However, one is able to define some correlation functions, the renormalized gauge-invariant correlation functions, which have a large cut-off limit. The relation (21.93), in presence of a momentum cut-off Λ ,

$$\widetilde{W}_\infty^{(2)}(u, v) = \exp \left[-\frac{e^2}{(2\pi)^d} \int d^d k \frac{1 - e^{ik(u-v)}}{m^2 k^2} \right] \widetilde{W}_0^{(2)}(u, v),$$

yields an explanation for this surprising property. For $|u - v| \neq 0$, the dominant term in the large cut-off limit in the exponential is (N_d is the loop factor, equation (21.112))

$$\frac{e^2}{(2\pi)^d} \int \frac{d^d k}{m^2 k^2} \propto N_d \frac{\Lambda^2}{m^2} \Lambda^{d-4} e^2. \quad (21.100)$$

Therefore, in the physical representation, all non-gauge-invariant correlation functions vanish. The technical explanation is the following: although the mass term breaks gauge invariance, this breaking is not sufficient to prevent fluctuations coming from the gauge degrees of freedom to suppress these correlation functions.

21.11 Renormalization and RG equations

In this section, we derive RG equations in the example of the action (21.30), which corresponds to massive QED with fermions, displaying the dependence of RG functions on the gauge fixing parameter ξ . We denote by $\Gamma^{(l,n)}$ the vertex (or one-particle irreducible) functions corresponding to l gauge fields, and n fermion pairs ψ and $\bar{\psi}$. In the Fourier representation, the relation between bare and renormalized correlation functions is

$$\tilde{\Gamma}_B^{(l,n)}(p_i, q_j; \alpha_0, \xi_0, m_0, M_0) = Z_A^{-l/2} Z_\psi^{-n} \tilde{\Gamma}^{(l,n)}(p_i, q_j; \mu, \alpha, \xi, m, M), \quad (21.101)$$

in which μ is the renormalization scale, and we have used the standard notation (we have set $\hbar = c = 1$)

$$\alpha = e^2 / 4\pi, \quad (21.102)$$

for the fine structure constant, which is also the loop expansion parameter.

Differentiating equation (21.101) with respect to μ at bare parameters fixed, one obtains the RG equations,

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(\alpha) \frac{\partial}{\partial \alpha} + \delta(\alpha) \xi \frac{\partial}{\partial \xi} + \eta_m(\alpha) m \frac{\partial}{\partial m} + \eta_M(\alpha) M \frac{\partial}{\partial M} - \frac{l}{2} \eta_A(\alpha) - n \eta_\psi(\alpha) \right] \tilde{\Gamma}^{(l,n)}(p_i, q_j; \mu, \alpha, \xi, m, M) = 0. \quad (21.103)$$

The gauge field and coupling constant renormalization constants Z_A and Z_e are related by (equation (21.91)),

$$Z_A Z_e = 1.$$

Therefore,

$$\alpha = Z_A \alpha_0. \quad (21.104)$$

Moreover, we have shown in Section 21.9 that the parameters m and ξ are not renormalized. It follows that,

$$m_0^2 = m^2 Z_A^{-1}, \quad \xi_0 = \xi Z_A. \quad (21.105)$$

Finally, in Section 21.10, we have shown that the renormalization constant Z_A can be chosen to be independent of ξ (the minimal subtraction (MS) scheme satisfies this requirement). Equations (21.104) and (21.105) then imply the three relations between RG functions:

$$\beta(\alpha) = \eta_A(\alpha), \quad \delta(\alpha) = -\beta(\alpha), \quad \eta_m(\alpha) = \beta(\alpha)/2. \quad (21.106)$$

In addition, $\beta(\alpha)$ is independent of ξ . The function η_M can also be chosen independent of ξ , only the fermion field renormalization is necessarily gauge dependent. Actually, from equation (21.98), it is possible to determine the gauge-dependence of η_ψ . A short calculation leads, in the MS scheme, to

$$\eta_\psi(\alpha, \xi) = \eta_\psi(\alpha, 0) - \alpha \xi / 2\pi. \quad (21.107)$$

21.12 One-loop β function and the triviality issue

We calculate now the β -function in four dimensions, using dimensional regularization and minimal subtraction, at one-loop order in the case of the action (21.27), which describes the interaction of a gauge field with charged fermions, and then add charged scalar bosons interacting with the gauge field. It follows from the discussion of Section 21.11 that only the gauge two-point function is required.

21.12.1 Charged Dirac fermions

Using the action (21.27), we evaluate the gauge field renormalization constant, and infer the coupling constant renormalization from relation (21.91).

The one-loop contribution to the generating functional of vertex functions coming from the fermion integration is

$$\Gamma_{\text{1 loop}}(\mathbf{A}) = -\text{tr} \ln (\not{\partial} + ie\not{\mathcal{A}} + M). \quad (21.108)$$

Differentiating twice with respect to A_μ , we obtain the one-loop contribution to the two-point vertex function in the Fourier representation:

$$\tilde{\Gamma}_{\mu\nu}^{(2)}(p) = Z_A (\delta_{\mu\nu} p^2 - p_\mu p_\nu) + p_\mu p_\nu / \xi + \Sigma_{\mu\nu}(p), \quad (21.109)$$

with

$$\Sigma_{\mu\nu}(p) = e^2 \int \frac{d^d k}{(2\pi)^d} \frac{\text{tr} [\gamma_\mu (\not{k} + iM) \gamma_\nu (\not{k} - \not{p} + iM)]}{(k^2 + M^2) [(p - k)^2 + M^2]}.$$

One verifies immediately that, as expected, the one-loop contribution is transverse and, therefore, ξ is not renormalized. We calculate the one-loop integral by introducing Feynman parameters. After some algebra, and with the help of the identity

$$m \frac{d}{dm} \int \frac{d^d k}{k^2 + m^2} = (d-2) \int \frac{d^d k}{k^2 + m^2} = -2m^2 \int \frac{d^d k}{(k^2 + m^2)^2}, \quad (21.110)$$

one obtains,

$$\Sigma_{\mu\nu}(p) = 2e^2 \text{tr} \mathbf{1} (p^2 \delta_{\mu\nu} - p_\mu p_\nu) \int_0^1 ds s(1-s) \int \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 + s(1-s)p^2 + M^2]^2}.$$

In particular, the divergent part of $\Sigma_{\mu\nu}(p)$ is

$$\Sigma_{\mu\nu}(p) = e^2 \text{tr} \mathbf{1} \frac{N_d}{3\varepsilon} (p^2 \delta_{\mu\nu} - p_\mu p_\nu) + O(1), \quad (21.111)$$

in which $\varepsilon = 4 - d$ and N_d is the loop factor:

$$N_d = 2/(4\pi)^{d/2} \Gamma(d/2). \quad (21.112)$$

This determines the \mathbf{A} field renormalization Z_A and thus, from equation (21.91) also Z_e ,

$$Z_A = 1 - N_d \text{tr} \mathbf{1} \frac{e^2}{3\varepsilon} + O(e^4), \quad Z_e = 1 + N_d \text{tr} \mathbf{1} \frac{e^2}{3\varepsilon} + O(e^4). \quad (21.113)$$

Replacing N_d and $\text{tr} \mathbf{1}$ by their values for $d = 4$ ($N_4 = 1/8\pi^2$, $\text{tr} \mathbf{1} = 4$), one obtains the corresponding β -function, for $d = 4$, at one-loop order:

$$\beta(\alpha) = -\varepsilon \left[\frac{d \ln (\alpha Z_e)}{d\alpha} \right]^{-1} = \frac{4}{3} \frac{\alpha^2}{2\pi} + O(\alpha^3). \quad (21.114)$$

21.12.2 The sign of the β -function

Note that in the domain of validity of the expansion (21.113) (α small), Z_A satisfies

$$Z_A \leq 1,$$

a result consistent with the Källen–Lehmann representation for the two-point function (see Section 6.6), which assumes Hermiticity. In the case of the Abelian gauge field, the property remains true, because Z_A is related to the transverse part of the two-point function, to which non-physical states do not contribute.

Since $Z_A Z_e = 1$, we see that the *sign* of the β -function in four dimensions is *determined* by Hermiticity.

The triviality or inconsistency issues. Since, for $d = 4$, and at least for α small enough, the β -function is positive, the physical charge is much smaller than the bare charge (the charge at the cut-off scale), a property that, intuitively, one can understand as a *screening* of the classical charge by relativistic quantum effects. Denoting by Λ the cut-off and μ the physical scale, for $\Lambda \gg \mu$, the effective charge $\alpha(\mu)$ at scale μ has the asymptotic form

$$\alpha(\mu) \sim \frac{3\pi}{2Q^2 \ln(\Lambda/\mu)}, \quad \text{with } Q^2 = \sum_i q_i^2,$$

where the q_i s are the charges of all other charged fermions expressed in unit of the electron charge. If one insists in taking the infinite cut-off limit, one faces the *triviality issue*: the physical charge vanishes. Otherwise, the field theory becomes inconsistent at the cut-off scale. In the SM of particle physics (Chapter 23), $Q^2 = \frac{14}{3}$ if we consider the running of α starting from the top mass (173 GeV), omitting all weak and strong interactions. Then,

$$\ln(\Lambda/\mu) \approx 9\pi/28\alpha(173),$$

and thus, Λ is a non-physical large scale, much larger than Planck's mass.



Fig. 21.4 Charged bosons: the two one-loop contributions to the A_μ two-point function

21.12.3 Charged scalar boson fields

We now consider the action, renormalizable in four dimensions, (see equation (21.33)),

$$\mathcal{S}(\mathbf{A}, \phi) = \mathcal{S}_\xi(\mathbf{A}) + \int d^d x \left(\sum_\mu |\mathbf{D}_\mu \phi(x)|^2 + r |\phi(x)|^2 + \frac{1}{6} g |\phi(x)|^4 \right), \quad g > 0, \quad (21.115)$$

where $\mathcal{S}_\xi(\mathbf{A})$ is the free action (21.21), ϕ a complex scalar boson, and r is positive, at leading order. Moreover,

$$\mathbf{D}_\mu = \partial_\mu + ieA_\mu(x).$$

In a different notation, and without gauge fixing, for $d = 3$ this is also the Landau–Ginzburg macroscopic model of superconductivity [223] (see Section 21.13).

To the two independent couplings e and g correspond two β -functions β_{e^2} and β_g .

We calculate here the divergent one-loop contributions to the gauge field propagator using Feynman diagrams, but an alternative method is described in Section A21.3. The two interaction terms proportional to $\phi^2 \mathbf{A}^2$ and $\phi^2 A_\mu$ generate the two diagrams of Fig. 21.4. Both diagrams can be obtained from the expansion to order A^2 of

$$\Gamma_{\text{1 loop}}(\mathbf{A}) = \text{tr} \ln \left[(\nabla + ie\mathbf{A})^2 - r \right] (\nabla^2 - r)^{-1}. \quad (21.116)$$

One obtains,

$$\tilde{\Gamma}_{\mu\nu, \text{1 loop}}^{(2)}(p) = 2e^2 \delta_{\mu\nu} \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 + r} - e^2 \int \frac{d^d k}{(2\pi)^d} \frac{(2k-p)_\mu (2k-p)_\nu}{(k^2 + r)[(k-p)^2 + r]}. \quad (21.117)$$

One again verifies that the result is transverse. Introducing Feynman parameters, and playing with the same identity (21.110), one finds

$$\tilde{\Gamma}_{\mu\nu, \text{1 loop}}^{(2)}(p) = e^2 (p^2 \delta_{\mu\nu} - p_\mu p_\nu) \int_0^1 ds (1-2s)^2 \int \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 + s(1-s)p^2 + r]^2}.$$

In particular, the divergent part is

$$[\tilde{\Gamma}_{\mu\nu}^{(2)}]_{\text{1 loop}}^{\text{div.}}(p) = \frac{N_d}{3\varepsilon} e^2 (p^2 \delta_{\mu\nu} - p_\mu p_\nu) + O(1). \quad (21.118)$$

The last part of the calculation is the same as in the fermion case.

The difference between fermions and bosons only comes from the trace of the identity in the space of γ matrices, which yields an additional factor 4 in the contribution of fermions to the β -function. For n_B complex scalar fields of charge e , one thus obtains

$$\beta_{e^2} = -\varepsilon e^2 + n_B \frac{e^4}{24\pi^2} + O(\text{two loops}). \quad (21.119)$$

To complete the calculation, one has still to calculate the function β_g . One finds,

$$\beta_g = -\varepsilon g + \frac{1}{24\pi^2} [(n_B + 4)g^2 - 18ge^2 + 54e^4] + O(\text{two loops}). \quad (21.120)$$

For $d = 4$, the origin $e^2 = g = 0$ is a stable IR fixed point only for $n_B \geq 183$, a result relevant, within the $\varepsilon = 4 - d$ expansion, for the field theory in three dimensions (see Section 21.13), as predicting a weak first order transition for $n_B \ll 183$.

Furry's theorem. For more general perturbative calculations, the following observation is useful. Correlation functions without matter field and an odd number of gauge fields vanish. The proof is based on charge conjugation. We consider the contribution to the effective gauge field action, $\det(\mathcal{D} + ie\mathcal{A} + M)$, which is generated by the integration over the fermion fields and use the property of the charge conjugation matrix C introduced in Section A12.1.7 (T means transposed):

$$\begin{aligned} \det(\mathcal{D} + ie\mathcal{A} + M) &= \det[(\mathcal{D} + ie\mathcal{A} + M)^T] = \det C^{-1} [(\mathcal{D} + ie\mathcal{A} + M)^T] C \\ &= \det(\mathcal{D} - ie\mathcal{A} + M). \end{aligned}$$

Therefore, the interaction between gauge fields generated by the fermions is even in A_μ . In particular, this implies that Feynman diagrams with fermion loops with an odd number of external gauge fields can be omitted.

21.13 The Abelian Landau–Ginzburg–Higgs model

We now discuss the Abelian Landau–Ginzburg–Higgs model: a gauge field coupled to a charged scalar field $\phi(x)$ in the spontaneously broken phase.

We consider again the gauge-invariant action (21.115) for a complex field ϕ , (the Ginzburg–Landau Hamiltonian in the terminology of statistical physics [223] and the Abelian Higgs model [224] in the particle-physics context)

$$\mathcal{S}(\mathbf{A}, \phi) = \int d^d x \left(\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) + \sum_{\mu} |\nabla_{\mu} \phi(x)|^2 + r |\phi(x)|^2 + \frac{1}{6} g |\phi(x)|^4 \right), \quad (21.121)$$

with $g > 0$, and $r \propto T - T_c$ for the superconductor. The action has a $U(1)$ symmetry corresponding to multiplying ϕ by a phase.

In three dimensions, this is an effective model for a superconductor in a static magnetic field near T_c and, in four dimensions, a piece of the weak sector of the SM of particle physics (see Section 23.1).

21.13.1 Classical approximation

In the classical approximation, for $r > 0$, the $U(1)$ symmetry is unbroken, the gauge field is massless, and the scalar field ϕ has two real components with equal mass \sqrt{r} .

However, for $r < 0$, the $U(1)$ symmetry is spontaneously broken, and $\phi(x)$ acquires a non-vanishing expectation value $|\langle \phi \rangle| = v/\sqrt{2} = \sqrt{-3r/g}$, which, for convenience, we assume real positive (we comment later on the significance of $\langle \phi \rangle$, which is not gauge invariant). This is a situation we have met in Section 13.4, and we have concluded that the spontaneous symmetry breaking of a continuous symmetry implies the presence of Goldstone modes in the form of a massless scalar bosons. This result can be derived by parametrizing, for $v \neq 0$, the field ϕ as (see Section 13.4)

$$\phi(x) = \rho(x) e^{i\theta(x)} / \sqrt{2}, \quad \text{with } \langle \rho \rangle = v = \sqrt{-6r/g}. \quad (21.122)$$

As a consequence of the symmetry, the resulting action then depends only on θ through $\partial_{\mu} \theta$; the field θ , therefore, is massless.

However, in a gauge theory the transformation $\phi(x) \mapsto \rho(x)$ has the form of a gauge transformation. By performing the corresponding transformation on the gauge field $A_{\mu}(x)$,

$$\mathbf{A}(x) = \mathbf{A}'(x) - \frac{1}{e} \nabla \theta(x), \quad (21.123)$$

one eliminates the field θ from the action completely. After this transformation, the action $\mathcal{S}(\mathbf{A}, \phi)$ becomes

$$\begin{aligned} \mathcal{S}(\mathbf{A}, \rho) = \int d^d x & \left[\frac{1}{4} F_{\mu\nu}^2(x) + \frac{1}{2} (\nabla \rho(x))^2 + \frac{1}{2} e^2 \mathbf{A}^2(x) \rho^2(x) \right. \\ & \left. + \frac{1}{2} r \rho^2(x) + \frac{1}{24} g \rho^4(x) \right]. \end{aligned} \quad (21.124)$$

We then set $\rho = v + \rho'$ and collect the quadratic part of the action. We find that, in the classical limit, the spectrum of the theory contains a massive vector particle and a massive scalar particle, now generally called, in particle physics, the Higgs particle, with masses (inverse penetration length and inverse coherence length, resp.)

$$m(A) = v|e|, \quad m(\rho) = v\sqrt{g/3}. \quad (21.125)$$

As a consequence of gauge invariance, no Goldstone boson has been generated. This is a most remarkable property, which is also at the basis of the Meissner effect in superconductivity. There, it is induced by the long-range forces (in the non-relativistic model the magnetic force) generated by the massless vector field.

Between the symmetric phase and the spontaneously broken phase, the total number of physical degrees of freedom has not changed, since one degree of freedom of the scalar field has been transferred to the longitudinal component of the massive vector field.

As a final remark, for later purpose, we recall that the mechanism of spontaneous symmetry breaking can also be used to give a masses to fermions in a chiral invariant theory (Section 13.6).

21.13.2 Quantization: Unitary gauge

The transformation (21.122) and the elimination of the $\theta(x)$ field corresponds for the action (21.121) to a gauge fixing: the representation (21.124) is the action in the so-called *unitary gauge*. The corresponding theory can be directly quantized, contains only physical fields, but is not renormalizable.

On the other hand, if one starts from the action (21.121) and quantize it in the same way as in the symmetric phase, the theory contains non-physical degrees of freedom, but is expected to be renormalizable. We are reminded of the massive vector field coupled to a conserved current, discussed in the first part of the chapter. Actually, there exists a formal relation between the massive vector field and the Higgs model: if one takes the non-linear σ -model limit of the action (21.121), that is, a limit in which the bare mass of the Higgs field becomes infinite at fixed expectation value v , one recovers the action (21.7) with the identification $m = |e|v$.

21.13.3 Renormalizable gauge

In order to be able to perform perturbative calculations of gauge-invariant observables and S -matrix elements, we return to the action (21.121). We fix the gauge by adding a term proportional to $(\nabla \cdot \mathbf{A})^2$. This amounts to coupling the phase field $\theta(x)$, which plays the role of the $\chi(x)$ field of Section 21.2.1.

We start from the renormalizable action (21.115), to which we add a mass term for the vector field to provide an IR cut-off,

$$\mathcal{S}(\mathbf{A}, \phi) = \mathcal{S}_\xi(\mathbf{A}, \phi) + \int d^d x \left[\frac{1}{2} m^2 \mathbf{A}^2(x) + \sum_\mu |\mathbf{D}_\mu \phi(x)|^2 + r |\phi(x)|^2 + \frac{g}{6} |\phi(x)|^4 \right]. \quad (21.126)$$

We introduce the real and imaginary parts of ϕ and set

$$\phi(x) = [\sigma(x) + i\pi(x)] / \sqrt{2}. \quad (21.127)$$

Since v has been chosen real and positive, $\langle \sigma(x) \rangle = v$. After the translation $\sigma(x) = v + \zeta(x)$, the quadratic part \mathcal{S}_2 of the action becomes

$$\begin{aligned} \mathcal{S}_2(\mathbf{A}, \phi) = & \int d^d x \left[\frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) + \frac{1}{2} (\nabla \cdot \mathbf{A}(x))^2 / \xi + \frac{1}{2} (e^2 v^2 + m^2) \mathbf{A}^2(x) \right. \\ & \left. - ev\pi(x) \nabla \cdot \mathbf{A}(x) + \frac{1}{2} (\nabla \zeta(x))^2 + \frac{1}{2} (\nabla \pi(x))^2 + \frac{1}{6} gv^2 \zeta^2(x) \right]. \end{aligned} \quad (21.128)$$

We note that $\nabla \cdot \mathbf{A}$ is coupled to the Goldstone field π . The corresponding propagators are

$$\widetilde{W}_{\mu\nu}^{(2)}(k) = \frac{\delta_{\mu\nu} - k_\mu k_\nu / k^2}{k^2 + e^2 v^2 + m^2} + \xi \frac{k_\mu k_\nu}{k^2(k^2 + \xi m^2)}, \quad \widetilde{W}_{\mu\pi}^{(2)}(k) = -\xi \frac{iev k_\mu}{k^2(k^2 + \xi m^2)}, \quad (21.129)$$

$$\widetilde{W}_{\pi\pi}^{(2)}(k) = \frac{1}{k^2} + \frac{\xi e^2 v^2}{k^2(k^2 + \xi m^2)}, \quad \widetilde{W}_{\zeta\zeta}^{(2)}(k) = \frac{1}{k^2 + gv^2/3}. \quad (21.130)$$

The spectrum of the theory contains three physical states and the usual state with negative norm coming from the regulator. We note that, in the absence of a mass term for the vector field in the action (21.126), the theory is potentially IR divergent in four dimensions. On the other hand, with the mass term the gauge symmetry is broken, and the χ -field really corresponds to a Goldstone mode. Even in the physical gauge a massless field is then present and coupled.

21.14 The Landau–Ginzburg–Higgs model: WT identities

From the combined analysis of Chapter 13 and Section 21.9, it is possible to derive WT identities, and to prove the renormalizability of the theory in the broken phase [225]. It follows that, after renormalization, the correlation functions satisfy the equivalent of the WT identities (21.83) and (21.86). As a consequence, the dependence of correlation functions on the parameter ξ can be determined as in Section 21.10. In particular, only correlation functions of gauge-invariant operators, and S -matrix elements, are gauge independent.

The explicit form of the WT identities for correlation functions is now rather complicate. We write here only the identities corresponding to the (\mathbf{A}, π) two-point vertex functions. Denoting by v the expectation value of the renormalized φ -field, we obtain by differentiating equation (21.86) with respect to A_μ ,

$$\sum_\nu k_\nu \tilde{\Gamma}_{\mu\nu}^{(2)}(k) + iev \tilde{\Gamma}_{\mu\pi}^{(2)}(k) = k_\mu (k^2/\xi + m^2). \quad (21.131)$$

Differentiating with respect to π , we find,

$$\sum_\mu k_\mu \tilde{\Gamma}_{\mu\pi}^{(2)}(k) - iev \tilde{\Gamma}_{\pi\pi}^{(2)}(k) = 0. \quad (21.132)$$

We parametrize the different functions as

$$\tilde{\Gamma}_{\mu\nu}^{(2)}(k) = a(k^2) \delta_{\mu\nu} - b(k^2) k_\mu k_\nu, \quad \tilde{\Gamma}_{\mu\pi}^{(2)}(k) = iev c(k^2) k_\mu, \quad \tilde{\Gamma}_{\pi\pi}^{(2)}(k) = d(k^2). \quad (21.133)$$

In the tree approximation, the values of a , b , c and d are

$$\begin{cases} a(k^2) = e^2 v^2 + m^2 + k^2, \\ c(k^2) = 1, \end{cases} \quad \begin{cases} b(k^2) = 1 - 1/\xi, \\ d(k^2) = k^2. \end{cases} \quad (21.134)$$

We now express the identity (21.132),

$$d(k^2) = k^2 c(k^2). \quad (21.135)$$

The identity (21.131) leads to

$$a(k^2) - k^2 b(k^2) - e^2 v^2 c(k^2) = k^2 / \xi + m^2. \quad (21.136)$$

In particular, in the $k = 0$ limit, the equation implies

$$a(0) - e^2 v^2 c(0) = m^2. \quad (21.137)$$

The corresponding connected correlation functions are

$$\begin{aligned} \widetilde{W}_{\mu\nu}^{(2)}(k) &= \frac{1}{a} \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) + \frac{\xi k_\mu k_\nu}{k^2(k^2 + \xi m^2)}, \\ \widetilde{W}_{\mu\pi}^{(2)}(k) &= -\frac{iev\xi k_\mu}{k^2(k^2 + \xi m^2)}, \\ \widetilde{W}_{\pi\pi}^{(2)}(k) &= \frac{1}{ck^2}. \end{aligned} \quad (21.138)$$

21.15 Spontaneous symmetry breaking: Decoupling gauge

The quantization method we have used previously leads to massless fields and thus IR divergences, even though the physical theory contained only massive fields. By the cleverer choice of a gauge, which explicitly breaks the global $U(1)$ symmetry of the action (and, therefore, eliminates the Goldstone modes), it is possible to circumvent this difficulty. In the notation of the action (21.126), one imposes, instead of the gauge condition (21.66), the condition ('t Hooft gauge, [226]) :

$$\nabla \cdot [\mathbf{B}(x) + \nabla\theta(x)] + \lambda ev \operatorname{Im}(\phi(x) e^{-ie\theta(x)}) - \nu(x) = 0, \quad (21.139)$$

in which λ is a constant, which has to be adjusted.

The insertion of the constraint (21.139) into the field integral representation generates a Jacobian of the form of the determinant of the functional derivative \mathbf{M} of equation (21.139) with respect to θ . In the bra–ket notation of quantum mechanics (QM),

$$\langle y | \mathbf{M} | x \rangle = \left[\nabla^2 - \lambda e^2 v \operatorname{Re}(\phi(x) e^{-ie\theta(x)}) \right] \delta^{(d)}(x - y). \quad (21.140)$$

The determinant $\det \mathbf{M}$ depends now on the dynamical fields. This is the source of difficulties of a kind we encounter again in Sections 22.4, 26.1.1, or Chapters 34, 35: we have to introduce spinless fermion fields to write $\det \mathbf{M}$ in local form, and then use the induced Becchi–Rouet–Stora–Tyutin (BRST) symmetry [236, 237] to show that renormalization preserves the form of the action. We postpone this analysis until Chapter 26 and discuss here only the tree approximation.

As before, we integrate over $\nu(x)$ with the distribution (21.59) and use the gauge invariance of the initial action:

$$\mathcal{S}_{\text{sym.}}(\mathbf{B}, \phi) = \mathcal{S}_{\text{sym.}}(\mathbf{A}, \phi e^{-ie\theta}). \quad (21.141)$$

Changing then variables in the field integral, $\phi e^{-ie\theta} \mapsto \phi$, we obtain the quantized action,

$$\begin{aligned} \mathcal{S}_{\text{qu}}(\mathbf{A}, \phi, \bar{C}, C) &= \int d^d x \left\{ \frac{1}{4} \sum_{\mu, \nu} F_{\mu\nu}^2(x) + \frac{1}{2\xi} (\nabla \cdot \mathbf{A}(x) + \lambda ev \operatorname{Im} \phi(x))^2 \right. \\ &\quad \left. + \sum_{\mu} |\mathbf{D}_{\mu} \phi(x)|^2 + r |\phi(x)|^2 + \frac{1}{6} g |\phi^4(x)| - C(x) (\nabla^2 - \lambda e^2 v \operatorname{Re} \phi(x)) \bar{C}(x) \right\}, \end{aligned} \quad (21.142)$$

in which $C(x)$ and $\bar{C}(x)$ are two scalar fermion fields, which are used to express the determinant $\det \mathbf{M}$ in local form.

As we have explained with an example in Section 12.6, scalar fermions cannot be interpreted as physical particles, but are of a nature similar to Pauli–Villars regulator fields. We now use the parametrization (21.127) and choose, at leading order,

$$\lambda = \xi\sqrt{2} . \quad (21.143)$$

The propagators are then

$$\begin{aligned}\widetilde{W}_{\mu\nu}^{(2)} &= \frac{\delta_{\mu\nu}}{k^2 + e^2 v^2} + \frac{(\xi - 1)k_{\mu k_\nu}}{(k^2 + e^2 v^2)(k^2 + \xi e^2 v^2)}, \\ \widetilde{W}_{\pi\pi}^{(2)} &= \frac{1}{k^2 + \xi e^2 v^2}, \\ \widetilde{W}_{CC}^{(2)} &= \frac{1}{k^2 + \xi e^2 v^2}.\end{aligned}\quad (21.144)$$

The advantages of this gauge are that, by construction, there is no $A_\mu \pi$ propagator and that all non-physical fields are massive and have the same mass $\xi e^2 v^2$. It suffices to prove gauge independence of physical observables to show that the pole at $k^2 = -\xi e^2 v^2$ cancels. The price to pay here is the more complicated form of WT identities, which now are mixed with BRST symmetry. We examine this question in Chapter 26 in detail.

21.16 Physical observables. Unitarity of the S -matrix

The non-physical pole at $k^2 = -\xi m^2$ can be shown to cancel in physical observables (gauge-invariant operators, S -matrix), either through a gauge-dependence analysis as we have done in Section 21.10, or directly by using the whole set of WT identities and showing explicitly that the pole coming from $\widetilde{W}_{\mu\nu}^{(2)}$ cancels the contribution coming from $\widetilde{W}_{\mu\pi}^{(2)}$ in the intermediate state in unitarity relations. As the expressions (21.138) show, the residues of the pole are related and, therefore, one understands that a cancellation is possible. The proof is not very difficult, but tedious and we refer to the literature [225].

In the limit $m = 0$, we expect also the pole at $k^2 = 0$ to cancel in physical observables. According to relation (21.137), for $k^2 \rightarrow 0$, the different propagators behave like

$$\begin{aligned}\widetilde{W}_{\mu\nu}^{(2)}(k) &\sim \frac{k_\mu k_\nu}{k^2} \left(\frac{1}{m^2} - \frac{1}{m^2 + e^2 v^2 c(0)} \right), \quad \widetilde{W}_{\mu\pi}^{(2)}(k) \sim -iev \frac{k_\mu}{k^2 m^2}, \\ \widetilde{W}_{\pi\pi}^{(2)}(k) &\sim \frac{1}{c(0)k^2}.\end{aligned}\quad (21.145)$$

Again, a direct argument based on WT identities for connected correlation functions and unitarity relations makes it possible to prove that, in the $m^2 = 0$ limit, the π -field decouples from physical observables. Here we do not have an alternative proof based upon gauge dependence. However, in a more general context, a proof can be constructed by using a different gauge.

Gauge-invariant operators. When the action (21.126) is used, only expectation values of gauge-invariant operators are physical correlation functions. The simplest examples are $F_{\mu\nu}(x)$ and $\varsigma(x) = \rho^2(x) = \frac{1}{2}\phi(x)\phi^*(x)$. In particular, the parameter v is really the square root of the expectation value $\langle \varsigma(x) \rangle$. Note that the action (21.124) can be rewritten in terms of the field $\zeta(x)$. One attractive feature of this representation is that the measure of integration in the field integral is just the flat measure $[d\zeta(x)]$.

Furthermore, all correlation functions of the transverse part of the vector field and the scalar field correspond directly, in this physical representation, to gauge-invariant correlation functions of the renormalizable representation. However, an inspection of the action written in terms of the $A_\mu(x)$ and $\varsigma(x)$ fields does not provide a direct explanation for the finiteness, after the introduction of a finite number of counter-terms, of the correlation functions of $F_{\mu\nu}$ and ς .

21.17 Stochastic quantization: The example of gauge theories

It has been suggested by Parisi and Wu [227] that stochastic dynamic equations (see Chapters 35, 36) could be used to quantize field theories when non-trivial quantization problems arise. The time variable in the Langevin equation is then a *fictitious* additional variable, since only the equilibrium distribution is physical (it can be interpreted as the computer time of numerical simulations, where stochastic methods are used to generate field configurations).

We briefly review the application of the idea to gauge theories. The problem in gauge theories is that one field degree of freedom, which corresponds to gauge transformations, is redundant and the conventional quantization method has to be adapted to this peculiar situation.

The idea of stochastic quantization, in the simplest example of the Abelian gauge field without matter, is the following. We consider gauge fields A_μ solutions of a gauge-invariant Langevin equation where the drift force is the functional derivative of the classical gauge-invariant action (see Chapter 35):

$$\dot{A}_\mu(t, x) = -\frac{1}{2}\Omega \sum_\nu \partial_\nu F_{\mu\nu}(t, x) + N_\mu(t, x), \quad (21.146)$$

where $(\mu, \nu = 1, \dots, d)$

$$F_{\mu\nu}(t, x) = \partial_\mu A_\nu(t, x) - \partial_\nu A_\mu(t, x),$$

and $N_\mu(x, t)$ is a Gaussian noise with measure

$$[d\rho(N)] = [dN_\mu] \exp \left[- \int d^d x dt \sum_\mu N_\mu^2(t, x)/2\Omega \right]. \quad (21.147)$$

The equation is clearly invariant under time-independent gauge transformations

$$\mathbf{A}(t, x) \mapsto \mathbf{A}(t, x) + \nabla\theta(x),$$

but not under time-dependent transformations. It provides a dynamics to all components of the gauge field.

In a non-gauge situation, the Langevin equation would generate an equilibrium distribution $e^{-\mathcal{A}}$, where \mathcal{A} is the action, and the Langevin equation, therefore, could be considered as an alternative method of quantization. Here the situation is clearly different. To understand the problem, we solve the equation, after Fourier transformation,

$$\begin{aligned} A_\mu(t, k) &= \sum_\nu (\delta_{\mu\nu} - k_\mu k_\nu/k^2) e^{-\Omega k^2 t/2} \left[A_\nu(0, k) + \int_0^t e^{\Omega k^2 t'/2} N_\nu(t', k) dt' \right] \\ &\quad + \sum_\nu (k_\mu k_\nu/k^2) \left[A_\nu(0, k) + \int_0^t N_\nu(t', k) dt' \right]. \end{aligned}$$

We immediately note that the component of A_μ along k_μ follows a Brownian motion and thus does not equilibrate. This can be verified explicitly by calculating the equal time two-point function averaged over the noise. Denoting by A_L and A_T the components of A_μ , respectively, along and perpendicular to k_μ , one finds,

$$\begin{aligned}\langle A_T(t, k) A_T(t, -k) \rangle &= \frac{1}{k^2} + O\left(e^{-\Omega k^2 t}\right), \\ \langle A_L(t, k) A_L(t, -k) \rangle &= A_L(0, k) A_L(0, -k) + \Omega t.\end{aligned}$$

Due to gauge invariance, the Langevin equation (21.146) does not generate an equilibrium distribution, although the gauge-invariant functions have a large time limit. Since only the latter functions have a physical meaning, the quantization problem has been solved.

It is clear that the same conclusion is reached if the gauge field interacts with matter, and in the non-Abelian case (see Chapter 22).

The next problems are the relation with the standard quantization procedure by gauge fixing, and the problems of power counting and renormalization.

The gauge field propagator obtained from the effective dynamic action in the Fourier representation reads

$$\tilde{\Delta}_{\mu\nu}(\omega, k) = \frac{\delta_{\mu\nu} - k_\mu k_\nu / k^2}{\omega^2 + \Omega k^4 / 4} + \frac{k_\mu k_\nu}{k^2 \omega^2}.$$

The $1/\omega^2$ singularity reflects the absence of equilibrium distribution, which now takes the form of IR divergences. It is necessary to work in a finite time interval. Also, the longitudinal propagator does not decrease at large momentum for fixed ω . A similar problem has also been encountered with non-covariant gauges: the theory is not renormalizable by power counting. One solution to the problem is to add a non-conservative drift force of the form $D_\mu V(A)$ to the Langevin equation, where D_μ is the covariant derivative, and $V(A)$ a linear function of A , for example $\nabla \cdot \mathbf{A}$. It is easy to verify that such a term does not contribute to the evolution equations (equation (34.31)) for equal-time gauge-invariant correlation functions, and thus, these functions are not modified. With this term, the Langevin equation is no longer gauge invariant, an equilibrium distribution is generated and, with a suitable covariant choice of $D_\mu V(A)$, the theory is renormalizable by power counting. The drawback is that much of the aesthetic appeal of the original formulation has been lost, and the proof of renormalizability becomes as complicated as in canonical quantization [228].

A general review of the possible applications of the method can be found in Ref. [229].

A21 Additional remarks

A21.1 Vacuum energy and Casimir effect

We discuss here the vacuum energy of a free vector field. The vacuum energy is a gauge-independent quantity but, because the formalism depends on the gauge, this is not always obvious. We then apply the result to the Casimir effect [230].

Note that while, in a non-gravitational theory, the *vacuum energy is not a physical observable*, a variation of the vacuum energy may be.

A21.1.1 The free electromagnetic field: Vacuum energy

The free massive vector field. We first consider the massive case. One way to calculate the ground state energy is to evaluate the free energy in a large Euclidean volume. The result is proportional to a space-time volume factor. Dividing by the volume, one obtains the vacuum energy density, the energy per unit space volume:

$$\mathcal{E}/\text{space volume} = -\ln \mathcal{Z}/\text{space-time volume}.$$

In the initial unitary gauge, one immediately finds in the Fourier representation,

$$-\ln \mathcal{Z} = \frac{1}{2} \text{tr} \ln [(k^2 + m^2) \delta_{\mu\nu} - k_\mu k_\nu] = \frac{1}{2}(d-1) \text{tr} \ln(k^2 + m^2) + \frac{1}{2} \text{tr} \ln m^2. \quad (\text{A21.1})$$

Up to an irrelevant constant, one finds $(d-1)$ times the vacuum energy of a free scalar field of mass m , not a surprising result, since the massive vector field has $(d-1)$ degrees of freedom. The result can be verified directly by diagonalizing the Hamiltonian (21.6).

By contrast, if one repeats the calculation with the action \mathcal{S}_ξ of equation (21.21), one obtains

$$-\ln \mathcal{Z} = \frac{1}{2}(d-1) \text{tr} \ln(k^2 + m^2) + \frac{1}{2} \text{tr} \ln(k^2 + \xi m^2)/\xi.$$

This gauge-dependent result for the vacuum energy is clearly incorrect. The reason can be simply understood. In the various algebraic transformations that have led to the action (21.21), we have omitted field-independent normalization factors. This is justified for correlation functions but not for the vacuum energy. The additional term originates from the χ field we have added to the theory in equation (21.16). A correct normalization of the χ integral cancels the additional unwanted contribution.

The massless gauge field. We have seen that correlation functions have a smooth massless limit. This is not the case for the vacuum energy. Indeed, in the massless limit, one degree of freedom of the vector field, the longitudinal mode, decouples from the theory but still contributes to the vacuum energy. We thus evaluate the energy directly, and then discuss the gauge dependence.

In the temporal gauge, we can use the Hamiltonian (21.46) (in a vanishing source). The Hamiltonian is partially diagonalized by a Fourier transformation (see Section 6.1.1)

$$\mathcal{H} = \frac{1}{2} \left[(2\pi)^{1-d} \tilde{E}_i(\hat{p}) \tilde{E}_i(-\hat{p}) + (2\pi)^{d-1} \tilde{A}_i(\hat{p}) (\hat{p}^2 \delta_{ij} - \hat{p}_i \hat{p}_j) \tilde{A}_j(-\hat{p}) \right], \quad (\text{A21.2})$$

where \hat{p} is a space momentum. We immediately note that we have $(d-2)$ harmonic oscillators with ground state energy $|\hat{p}|$ in the field direction perpendicular to the vector \hat{p} , and a free QM Hamiltonian in the \hat{p} direction. However, *Gauss's law* implies that $\hat{p} \cdot E(\hat{p})$ annihilates all physical states and, therefore, does not contribute to the energy. We conclude that the vacuum energy is only $(d-2)$ times the vacuum energy of a free massless scalar particle.

Covariant gauge. By contrast, evaluating the energy in a covariant gauge, using the corresponding partition function, naively one finds

$$-\ln \mathcal{Z} = \frac{1}{2} \text{tr} \ln [k^2 \delta_{\mu\nu} - k_\mu k^\nu (1 - 1/\xi)] = \frac{1}{2} d \text{tr} \ln k^2 - \frac{1}{2} \text{tr} \ln \xi.$$

Up to constant, this is the energy of d massless states instead of $(d-2)$. The discussion of the massive example gives an idea of the origin of the difficulty. We have omitted field independent factors in algebraic transformations. For instance, in the Faddeev–Popov quantization of Section 21.6.1, if we want the multiplication by the integral over $\theta(x)$ to be an identity, we have to multiply the integral (21.67) by $\det(-\nabla^2)$. Such a factor cancels a complex, or two real massless scalar bosons. It reduces d to $(d-2)$. In non-Abelian gauge theories, ‘ghost fields’ automatically produce the right book-keeping.

A21.1.2 Casimir effect

In Chapter 32, we discuss quantum field theory in a finite volume or, more generally, in restricted geometries. However, the simplest example of finite size effects concerns the change of vacuum energy produced by a change in boundary conditions. In particular, this leads to a QED effect (the Casimir effect), which takes the form of an attractive force between two parallel perfectly conducting plates in the vacuum.

At leading order, all charged particles can be omitted, because only massless fields have a significant contribution at large plate separation. Hence, the problem is reduced to a calculation of the change of vacuum energy of the free electromagnetic field due to boundary conditions. The conducting plates impose to the electric field to be perpendicular to the plates. It is easy to verify that this condition is satisfied if the vector field A_μ itself vanishes on the plates. Calling L the distance between the plates, $z = 0$ and $z = L$ the plate positions, one has to integrate over fields that have the Fourier representation,

$$A_\mu(\mathbf{x}_\perp, z) = \int d^{d-1}p_\perp \sum_{n \geq 1} e^{i\mathbf{p}_\perp \cdot \mathbf{x}_\perp} \sin(\pi z/L) \tilde{A}(\mathbf{p}_\perp, n), \quad (A21.3)$$

where \mathbf{x}_\perp are the space–time coordinates in the remaining directions. Since all components of A_μ satisfy the same boundary conditions and, since the field vacuum energy is $(d-2)$ (the number of field degrees of freedom) times the free scalar vacuum energy, we now solve the problem for a free massless scalar field.

The free massless scalar theory. We consider the free action for a massless scalar field $\phi(x)$,

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x (\nabla \phi(x))^2,$$

where the situation of physical interest is $d = 4$. We assume that in one direction the field satisfy the boundary conditions

$$\phi(\mathbf{x}_\perp, z = 0) = \phi(\mathbf{x}_\perp, z = L) = 0.$$

For the other directions, we first assume periodic boundary conditions with finite size L_\perp , with $L_\perp \gg L$. Integrating over the field, we find that the vacuum energy \mathcal{E} is given by

$$L_\perp \mathcal{E} = \sum_{n \geq 1} \sum_{\mathbf{n}_\perp} \frac{1}{2} \ln \left(\frac{4\pi^2 \mathbf{n}_\perp^2}{L_\perp^2} + \frac{\pi^2 n^2}{L^2} \right). \quad (A21.4)$$

This expression is only meaningful in the presence of a UV cut-off. Since

$$\int_{\varepsilon}^{\infty} \frac{dt}{t} e^{-st} \underset{s \rightarrow 0}{=} -\ln(\varepsilon s) + \psi(1) + O(\varepsilon),$$

we replace it by the regularized form,

$$L_{\perp} \mathcal{E} = - \sum_{n \geq 1} \sum_{\mathbf{n}_{\perp}} \frac{1}{2} \int_{a^2}^{\infty} \frac{dt}{t} \exp \left[-t \left(\frac{4\pi^2 \mathbf{n}_{\perp}^2}{L_{\perp}^2} + \frac{\pi^2 n^2}{L^2} \right) \right],$$

where $\Lambda = 1/a$ plays the role of a UV cut-off. We now take the large- L_{\perp} limit. We replace sums by integrals,

$$L_{\perp} \mathcal{E} = -\frac{1}{2} \sum_{n \geq 1} \int_{a^2}^{\infty} \frac{dt}{t} \int d^{d-1} \mathbf{n}_{\perp} \exp \left[-t \left(\frac{4\pi^2 \mathbf{n}_{\perp}^2}{L_{\perp}^2} + \frac{\pi^2 n^2}{L^2} \right) \right].$$

The \mathbf{n}_{\perp} integral is Gaussian. We then obtain the energy per unit area

$$\mathcal{E}/L_{\perp}^{d-2} = -\frac{1}{2}(4\pi)^{(1-d)/2} \int_{a^2}^{\infty} \frac{dt}{t^{(d+1)/2}} \sum_{n \geq 1} \exp(-\pi^2 n^2 t/L^2). \quad (A21.5)$$

We evaluate the sum, using identity (A32.3):

$$\sum_{n \geq 1} \exp(-\pi^2 n^2 t/L^2) = \frac{L}{2\sqrt{\pi t}} - \frac{1}{2} + \frac{L}{\sqrt{\pi t}} \sum_{n \geq 1} \exp(-L^2 n^2/t).$$

The first term yields a contribution proportional to $L\Lambda^d$, which is the vacuum energy \mathcal{E}_0 in the absence of boundaries. The second term yields an L -independent surface energy of order Λ^{d-1} due to the boundaries. Finally, the remaining terms which are cut-off independent, but depend on L , give the interesting contribution. After integration over t , the contribution takes the form

$$(\mathcal{E} - \mathcal{E}_0)/L_{\perp}^{d-2} = \text{const.} - A(d)L^{1-d}, \quad (A21.6)$$

with

$$A(d) = \frac{\Gamma(d/2)}{(4\pi)^{d/2}} \zeta(d). \quad (A21.7)$$

For $d = 4$, one finds

$$A(4) = \frac{\pi^2}{1440}.$$

Casimir effect. The resulting force between plates is, thus, attractive (a feature that, however, depends on the geometry). To pass from the scalar result to the electromagnetic result, we have to take into account the $d - 2 = 2$ degrees of freedom of the gauge field for $d = 4$. For the energy and force F per unit area one finds, restoring the physical units,

$$(\mathcal{E} - \mathcal{E}_0)/L_{\perp}^2 = \text{const.} - \frac{\pi^2}{720} \frac{\hbar c}{L^3} \Rightarrow F = -\frac{1}{L_{\perp}^2} \frac{d\mathcal{E}}{dL} = -\frac{\pi^2}{240} \frac{\hbar c}{L^4}. \quad (A21.8)$$

This quantum relativistic effect is very small but has been measured in the $0.1\mu\text{m}-1\mu\text{m}$ range [230]. The effect is remarkable because, although electromagnetic in nature, it is independent of the value of the electric charge.

A21.2 Gauge dependence

To characterize the gauge-dependence of general correlation functions with only matter fields (or with only additional transverse gauge fields), we generalize the derivation of Section 21.10. Again the ξ dependence factorizes into the product of unitary correlation functions and χ integrals. Setting

$$\widetilde{W}_\xi^{(2n)}(x_1, \dots, x_n; y_1, \dots, y_n) = \widetilde{W}_\infty^{(2n)}(x_1, \dots, x_n; y_1, \dots, y_n) \\ \times U_\xi^{(2n)}(x_1, \dots, x_n; y_1, \dots, y_n),$$

one finds,

$$U_\xi^{(2n)}(x_1, \dots, x_n; y_1, \dots, y_n) = \left\langle \exp \left[\frac{ie}{m} \sum_{i=1,n} \chi(x_i) - \chi(y_i) \right] \right\rangle,$$

where $\langle \bullet \rangle$ means Gaussian expectation value with the χ -field measure corresponding to the action (21.94),

$$\mathcal{S}_\xi(\chi) = \int d^d x \left[-\frac{1}{2} (\nabla \chi(x))^2 - \frac{1}{2} \xi m^2 \chi^2(x) \right].$$

The integral can be calculated (for details see, for example, Section 30.1), and yields

$$U_\xi^{(2n)} = \exp \left\{ \frac{e^2}{m^2} \sum_{i,j} \left[\frac{1}{2} \Delta_\xi(x_i - x_j) + \frac{1}{2} \Delta_\xi(y_i - y_j) - \Delta_\xi(x_i - y_j) \right] \right\}, \quad (A21.9)$$

where Δ_ξ is the χ -field propagator (equation (21.96)),

$$\Delta_\xi(x) = \frac{1}{(2\pi)^d} \int \frac{d^d k e^{ikx}}{k^2 + \xi m^2}. \quad (A21.10)$$

As indicated in Section 21.10, it is convenient to take the ratio of correlation functions corresponding to two finite values of ξ . For example, the ratio between the generic ξ and Landau's gauge ($\xi = 0$) is then still given by an expression of the form (A21.9), but with the propagator (A21.10) replaced by

$$m^2 K_\xi(x) = \Delta_\xi(x) - \Delta_0(x) = -\frac{\xi m^2}{(2\pi)^d} \int \frac{d^d k e^{ikx}}{k^2 (k^2 + \xi m^2)}.$$

In expression (A21.9), only the diagonal terms $i = j$ with vanishing arguments yield UV divergences. One recognizes immediately the divergent factor $(Z_\psi(\xi)/Z_\psi(0))^n$ as given by equation (21.98):

$$\frac{U_\xi^{(2n)}}{U_0^{(2n)}} = \left(\frac{Z_\psi(\xi)}{Z_\psi(0)} \right)^n \\ \times \exp \left\{ e^2 \left[\sum_{i < j} (K_\xi(x_i - x_j) + K_\xi(y_i - y_j)) - \sum_{i,j} K_\xi(x_i - y_j) \right] \right\}. \quad (A21.11)$$

After renormalization, if one expands the ratio between correlation functions in powers of ξe^2 , one notes that all terms but the first one correspond to a matter field correlation function in which at least two external lines have been joined by a χ -field propagator (as in Fig. 21.3). Therefore, in the mass-shell limit, these terms do not have the corresponding poles and do not contribute to the S -matrix, as indicated in Section 21.10.

A21.3 Divergences at one loop from Schwinger's representation

We consider the action (21.115). The one-loop gauge action, obtained after integration over the charged scalar fields of the vector field one-loop contribution, is (equation (21.116))

$$\Gamma_{\text{1 loop}}(A) = \text{tr} \ln \left[(D^2 - r) (\nabla^2 - r)^{-1} \right], \text{ with } D^2 = \sum_{\mu} D_{\mu}^2.$$

To determine the divergent part, we use Schwinger's representation of $\text{tr} \ln H$ for a positive operator H ,

$$\text{tr} \ln H = \text{tr} \int_0^{\infty} \frac{dt}{t} e^{-tH}.$$

The divergent part is related to the integration near $t = 0$. Therefore, we need the small t expansion of the matrix elements of the operator $U = e^{tD^2}$, solution of the equation

$$\begin{aligned} (\nabla_x + ie\mathbf{A}(x))^2 \langle x | U(t) | x' \rangle &= \partial_t \langle x | U(t) | x' \rangle, \\ \text{with for } t \rightarrow 0_+, \quad \langle x | U(t) | x' \rangle &\sim \frac{e^{-(x-x')^2/2t}}{(4\pi t)^{d/2}}. \end{aligned} \quad (A21.12)$$

For $t > 0$, we set (x' is implicit and defines the boundary condition)

$$\langle x | U(t) | x' \rangle = e^{-\sigma(t,x)}. \quad (A21.13)$$

Equation (A21.12) then takes the form ($\dot{\sigma} \equiv \partial_t \sigma$)

$$\dot{\sigma}(t,x) = \nabla_x^2 \sigma(t,x) - ie\nabla \cdot \mathbf{A}(x) - (\nabla_x \sigma(t,x) - ie\mathbf{A}(x))^2. \quad (A21.14)$$

For $t \rightarrow 0$, the function σ has an expansion of the form,

$$\sigma(t,x) = \frac{1}{4t}(x-x')^2 + \frac{d}{2} \ln 4\pi t + \sigma_0(x) + \sigma_1(x)t + \sigma_2(x)t^2 + O(t^3). \quad (A21.15)$$

At this order,

$$\partial_{\mu} \sigma = \frac{1}{2t}(x-x')_{\mu} + \partial_{\mu} \sigma_0(x) + \partial_{\mu} \sigma_1(x)t + \partial_{\mu} \sigma_2(x)t^2, \quad \dot{\sigma} = -\frac{1}{4t^2}(x-x')^2 + \frac{d}{2t} + \sigma_1 + 2t\sigma_2.$$

It follows that

$$\begin{aligned} (\nabla_x \sigma(t,x))^2 &= \frac{1}{4t^2}(x-x')^2 + \frac{1}{t}(x-x') \cdot \nabla_x \sigma_0(x) + (\nabla_x \sigma_0(x))^2 + (x-x') \cdot \nabla_x \sigma_1(x) \\ &\quad + 2t \nabla_x \sigma_0(x) \cdot \nabla_x \sigma_1(x) + t(x-x') \cdot \nabla_x \sigma_2(x), \\ \nabla_x^2(t,x)\sigma &= \frac{d}{2t} + \nabla_x^2 \sigma_0(x) + t \nabla_x^2 \sigma_1(x) + t^2 \nabla_x^2 \sigma_2(x). \end{aligned}$$

The order t^{-1} yields the first non-trivial equation:

$$(x-x') \cdot (\nabla_x \sigma_0(x) - ie\mathbf{A}(x)) = 0 \Rightarrow \sigma_0(x) = ie \int_0^1 ds (x-x') \cdot \mathbf{A}(x' + s(x-x')).$$

The solution has the expected gauge transformations,

$$\delta A_{\mu}(x) = \partial_{\mu} \theta(x) \Rightarrow \delta \sigma_0(x) = ie(\theta(x) - \theta(x')).$$

The term of order t^0 yields

$$\sigma_1(x) + (x-x') \cdot \nabla \sigma_1(x) = \nabla \cdot (\nabla \sigma_0(x) - ie\mathbf{A}(x)) - (\nabla \sigma_0(x) - ie\mathbf{A}(x))^2. \quad (A21.16)$$

Because

$$\partial_\mu \sigma_0(x) - ie A_\mu(x) = ie \int_0^1 ds s \sum_\nu (x - x')_\nu F_{\mu\nu}(x' + s(x - x')),$$

is gauge invariant, σ_1 is also gauge invariant. Then,

$$\nabla \cdot (\nabla \sigma_0(x) - ie \mathbf{A}(x)) = ie \int_0^1 ds s^2 \sum_{\mu,\nu} (x - x')_\nu \partial_\mu F_{\mu\nu}(x' + s(x - x')).$$

We solve the general equation

$$\sigma_1(x) + (x - x') \cdot \nabla \sigma_1(x) = X(x) \Rightarrow \sigma_1(x, x') = \int_0^1 ds X(x' + s(x - x')).$$

The first and second terms in the right-hand side of equation (A21.16) yield

$$\begin{aligned} & ie \int_0^1 ds s(1-s) \sum_{\mu,\nu} (x - x')_\nu \partial_\mu F_{\mu\nu}(x' + s(x - x')), \\ & 2e^2 \int u du v dv \sum_{\mu,\nu,\rho} (x - x')_\nu (x - x')_\rho F_{\mu\nu}(x' + u(x - x')) F_{\mu\rho}(x + v(x' - x)), \end{aligned}$$

with $u \geq 0$, $v \geq 0$ and $v + u \leq 1$.

We need the last equation coming from the coefficient of t only for $x = x'$, because $x - x' = O(\sqrt{t})$. Thus,

$$2\sigma_2(x) = \nabla^2 \sigma_1(x).$$

The contribution to σ_1 linear in e vanishes in this limit. The solution is

$$\sigma_2(x) = \frac{1}{12} e^2 \sum_{\mu,\nu} F_{\mu\nu}^2(x).$$

In addition $\sigma_0(x') = \sigma_1(x') = 0$. The divergent contribution, at one-loop, to the gauge field two-point function generated by charged scalars can thus be inferred from

$$\begin{aligned} \text{tr} \ln[(-D^2 + r)(-\nabla^2 + r)^{-1}] &= \int \frac{dt}{t} e^{-rt} \text{tr}(e^{t\nabla^2} - U(t)) \\ &\sim \frac{e^2}{12} \frac{1}{(4\pi)^{d/2}} \int \frac{dt}{t^{1+d/2}} t^2 \int d^d x \sum_{\mu,\nu} F_{\mu\nu}^2(x). \end{aligned}$$

In dimensional regularization, the divergent term is

$$\text{tr} \ln(-D^2 + M^2) - \text{tr} \ln(-\nabla^2 + M^2) \sim \frac{1}{8\pi^2} \frac{e^2}{3\varepsilon} \frac{1}{4} \int d^4 x \sum_{\mu,\nu} F_{\mu\nu}^2(x),$$

in agreement with the Feynman diagram calculation.

The fermion determinant can also be related to the scalar determinant by

$$\begin{aligned} \text{tr} \ln(\not{D} + M) &= \frac{1}{2} \text{tr} \ln [M^2 - (\not{D})^2] = \frac{1}{2} \text{tr} \ln \left(M^2 - D^2 + \frac{1}{2} e \sum_{\mu,\nu} F_{\mu\nu} \sigma_{\mu\nu} \right), \\ &= \frac{1}{2} \text{tr } \mathbf{1} \text{tr}' \ln(M^2 - D^2) - \frac{1}{8} \text{tr } \mathbf{1} e^2 \text{tr}' \sum_{\mu,\nu} F_{\mu\nu} \nabla^{-2} F_{\mu\nu} \nabla^{-2} + O(A^4), \\ &\sim \frac{1}{2} \text{tr } \mathbf{1} e^2 \frac{1}{8\pi^2} \left(\frac{1}{3\varepsilon} - \frac{1}{\varepsilon} \right) \frac{1}{4} \int d^4 x \sum_{\mu,\nu} F_{\mu\nu}^2(x), \end{aligned}$$

where tr' means trace over space variables only. Taking into account the minus sign in front of the fermion determinant, one finds $\text{tr } \mathbf{1}$ times the boson result, in agreement with Feynman diagram calculations.

22 Non-Abelian gauge theories: Introduction

In Chapter 21, we have described the structure and the formal properties of Abelian gauge theories, which provide a framework for the construction of quantum electrodynamics (QED). However, to be able to describe the other fundamental interactions, weak and strong, at the microscopic scale, it is necessary to generalize the concept of gauge symmetry to non-Abelian groups. Therefore, in this chapter, we construct a quantum field theory (QFT) invariant under *local*, that is, space-time-dependent, transformations of matrix representations of a general compact Lie group G [231]. Inspired by the Abelian example of Chapter 21, we immediately introduce the geometric concept of *parallel transport*, a concept discussed more extensively in Chapter 28 in the framework of Riemannian manifolds. All the required mathematical quantities then appear naturally.

Early articles on the quantization problem include Refs. [26, 231–233]. We first quantize gauge theories in the temporal gauge [218]. We show the equivalence with covariant gauges [219, 226, 234, 235]. We exhibit some formal properties of the quantum theory, like the Becchi–Rouet–Stora–Tyutin (BRST) symmetry [236, 237]. We establish the Feynman rules of perturbation theory [238, 41] and discuss how perturbation theory can be regularized, a somewhat non-trivial problem [239–221]. Finally, we discuss general properties of the non-Abelian Higgs mechanism [242, 226, 243]. A selection of early articles can be found in Ref. [245], and a few reviews and books in Refs. [5, 246, 215].

In the appendix, we quantize massive non-Abelian gauge fields, and explain why they lead to non-renormalizable QFTs [244].

We work with a Euclidean formalism, except when explicitly stated otherwise.

22.1 Geometric construction: Parallel transport

We consider a scalar field $\phi(x)$, $x \in \mathbb{R}^d$, transforming under a linear, unitary (but the generalization to orthogonal is simple) representation $\mathcal{R}(G)$ of a compact Lie group G . We want to construct a QFT that has a *local* G symmetry, that is, a theory where the action is invariant under space-dependent group transformations, also called *gauge transformations*. Denoting a matrix belonging to a representation $\mathcal{R}(G)$ by \mathbf{g} , we write the ϕ -field transformation as

$$\phi^{\mathbf{g}}(x) = \mathbf{g}(x)\phi(x). \quad (22.1)$$

If we consider only products of fields taken at the same point, global invariance (\mathbf{g} constant) implies local (or gauge) invariance. However, if we consider invariant functions of fields and their derivatives, or more generally products of fields taken at different points, this is no longer true. An analogous problem arises in the study of Riemannian manifolds, and is solved by the introduction of parallel transport. We refer to Section 28.2 for additional geometric and algebraic considerations.

Parallel transport. Let C_{xy} be a continuous, piecewise differentiable oriented curve, y being the origin and x the end-point. We denote by C^{-1} the curve C with opposite orientation. To the curve C , we associate a parallel transporter $\mathbf{U}(C)$, element of the representation $\mathcal{R}(G)$. Parallel transporters $\mathbf{U}(C)$ satisfy the rules (28.15) and (28.17),

$$\mathbf{U}(C_{xy})\mathbf{U}(C_{yz}) = \mathbf{U}(C_{xy} \cup C_{yz}) \Rightarrow \mathbf{U}(C_{xx} \equiv 1 \text{ point}) = \mathbf{1}, \quad \mathbf{U}(C^{-1}) = \mathbf{U}^{-1}(C).$$

Associated with a space-dependent group element $\mathbf{g}(x)$, we define the corresponding gauge transformation of $\mathbf{U}(C_{xy})$ by

$$\mathbf{U}^{\mathbf{g}}(C_{xy}) = \mathbf{g}(x)\mathbf{U}(C_{xy})\mathbf{g}^{-1}(y). \quad (22.2)$$

It then follows that the vector ϕ_U defined by

$$\phi_U(x) = \mathbf{U}(C_{xy})\phi(y), \quad (22.3)$$

transforms by $\mathbf{g}(x)$ instead of $\mathbf{g}(y)$, and the product $\phi^\dagger(x)\mathbf{U}(C_{xy})\phi(y)$ is *gauge invariant*.

Connection. We assume now that $\mathbf{U}(C)$ is differentiable in the curve C for small enough differentiable curves. Considering an infinitesimal straight line in d dimensions,

$$y_\mu = x_\mu + dx_\mu, \quad (22.4)$$

we expand \mathbf{U} as

$$\mathbf{U}(C_{xy}) = \mathbf{1} + \mathbf{A}(x) \cdot d\mathbf{x} + o(\|d\mathbf{x}\|), \quad (22.5)$$

where the *connection* $\mathbf{A}_\mu(x)$ is a vector from the point of view of space transformations, and an anti-Hermitian matrix belonging to the adjoint representation of the Lie algebra of G . Parallel transporters can then be defined by line-ordered integrals in terms of the connection (see equation (21.69) for the Abelian case).

Expanding equation (22.2) at first order in dx_μ , one finds how $\mathbf{A}_\mu(x)$ transforms under gauge transformations ($\partial_\mu \equiv \partial/\partial x_\mu$):

$$\mathbf{A}_\mu^{\mathbf{g}}(x) = \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x). \quad (22.6)$$

From the point of view of global transformations ($\mathbf{g}(x)$ constant), the field $\mathbf{A}_\mu(x)$ transforms by the adjoint representation of the group G . However, $\mathbf{A}_\mu(x)$, which is usually called the *gauge field* or *Yang–Mills field*, is not a tensor for gauge transformations, the transformation (22.6) being affine.

Covariant derivative. To the connection $\mathbf{A}_\mu(x)$ is associated a covariant derivative \mathbf{D}_μ , whose explicit form depends on the tensor on which it is acting. To derive its expression when acting on $\phi(x)$, we consider in equation (22.3) the limit of an infinitesimal curve, as in equation (22.4). Equation (22.3) becomes

$$\begin{aligned} \phi_U &= (\mathbf{1} + \mathbf{A}(x) \cdot d\mathbf{x})(\phi(x) + \nabla\phi(x) \cdot d\mathbf{x}) + o(\|d\mathbf{x}\|) \\ &= (\mathbf{1} + d\mathbf{x} \cdot \mathbf{D})\phi(x) + o(\|d\mathbf{x}\|), \end{aligned} \quad (22.7)$$

where we have defined

$$\mathbf{D}_\mu \equiv \mathbf{D}_\mu(\mathbf{A}) = \mathbf{1}\partial_\mu + \mathbf{A}_\mu. \quad (22.8)$$

\mathbf{D}_μ is both a differential operator acting on space variables and a matrix. The identity

$$\mathbf{g}(x)\mathbf{D}_\mu(\mathbf{A})\mathbf{g}^{-1}(x) = \mathbf{1}\partial_\mu + \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x) = \mathbf{D}_\mu(\mathbf{A}^{\mathbf{g}}) \quad (22.9)$$

shows that \mathbf{D}_μ is a tensor, since $\mathbf{D}_\mu^{\mathbf{g}}$, the transform of \mathbf{D}_μ under the gauge transformation (22.6), is

$$\mathbf{D}_\mu^{\mathbf{g}} \equiv \mathbf{D}_\mu(\mathbf{A}^{\mathbf{g}}) = \mathbf{g}(x)\mathbf{D}_\mu\mathbf{g}^{-1}(x) \equiv \mathbf{g}(x)\mathbf{D}_\mu(\mathbf{A})\mathbf{g}^{-1}(x). \quad (22.10)$$

In equations (22.9) and (22.10), the products must be understood as products of differential and multiplicative operators.

Parallel transport: Integrated form. In Section 21.7, we have expressed the parallel transporter in terms of gauge field (equation (21.69)). In the non-Abelian case (Section 22.1), the explicit relation is more complicated, because the gauge field $\mathbf{A}_\mu(x)$ is an element of the Lie algebra of G , and the matrices representing the field at different points do not commute. The parallel transporter, which can be formally written as

$$\mathbf{U}[C(x, y)] = \text{P} \left\{ \exp \left[\oint_C \mathbf{A}(s) \cdot ds \right] \right\}, \quad (22.11)$$

in which the symbol P means path-ordered integral, transforms as in equation (22.2). The path-ordered integral is the continuum limit of an ordered product of parallel transporters corresponding to small s intervals.

Infinitesimal gauge transformations. Setting,

$$\mathbf{g}(x) = \mathbf{1} + \boldsymbol{\omega}(x) + o(\|\boldsymbol{\omega}\|), \quad (22.12)$$

in which $\boldsymbol{\omega}(x)$ belongs to the Lie algebra of $\mathcal{R}(G)$, we derive from equation (22.6) the form of the infinitesimal gauge transformation of the field \mathbf{A}_μ ,

$$-\delta_\omega \mathbf{A}_\mu(x) = \partial_\mu \boldsymbol{\omega}(x) + [\mathbf{A}_\mu(x), \boldsymbol{\omega}(x)] \equiv \mathbf{D}_\mu^{\text{adj}} \cdot \boldsymbol{\omega}(x). \quad (22.13)$$

In equation (22.8), we have given the form of the covariant derivative corresponding to the representation $\mathcal{R}(G)$. The equation (22.13) yields the form of the covariant derivative when acting on a vector transforming by the adjoint representation. Indeed, one verifies the identity

$$\partial_\mu \boldsymbol{\omega}^g(x) + [\mathbf{A}_\mu^g(x), \boldsymbol{\omega}^g(x)] = \mathbf{g}(x) \{ \partial_\mu \boldsymbol{\omega}(x) + [\mathbf{A}_\mu(x), \boldsymbol{\omega}(x)] \} \mathbf{g}^{-1}(x), \quad (22.14)$$

in which \mathbf{A}_μ^g is given by equation (22.6) and $\boldsymbol{\omega}^g$ by

$$\boldsymbol{\omega}^g(x) = \mathbf{g}(x) \boldsymbol{\omega}(x) \mathbf{g}^{-1}(x). \quad (22.15)$$

Curvature tensor. The commutator of two covariant derivatives,

$$\mathbf{F}_{\mu\nu}(x) = [\mathbf{D}_\mu, \mathbf{D}_\nu] = \partial_\mu \mathbf{A}_\nu(x) - \partial_\nu \mathbf{A}_\mu(x) + [\mathbf{A}_\mu(x), \mathbf{A}_\nu(x)], \quad (22.16)$$

is no longer a differential operator. It is again an element of the Lie algebra of $\mathcal{R}(G)$ and transforms, as a consequence of equation (22.10), as

$$\mathbf{F}'_{\mu\nu}(x) = \mathbf{g}(x) \mathbf{F}_{\mu\nu}(x) \mathbf{g}^{-1}(x). \quad (22.17)$$

$\mathbf{F}_{\mu\nu}$ is a tensor, the curvature tensor, generalization of the electromagnetic field of QED. As we show in Chapter 28, in a different context, the curvature tensor is associated with parallel transport along infinitesimal closed curves.

Differential forms and gauge fields. It is sometimes convenient to use the language of differential forms and to associate a one-form with the gauge field, and a two-form with the curvature tensor (see Sections 1.5 and 28.1.3):

$$\mathfrak{A}(x) = \mathbf{A}(x) \cdot d\mathbf{x}, \quad \mathfrak{F}(x) = \sum_{\mu,\nu} \mathbf{F}_{\mu\nu}(x) dx_\mu \wedge dx_\nu = 2(d\mathfrak{A}(x) + \mathfrak{A}^2(x)), \quad (22.18)$$

where the differential operator d acting on forms is defined by equation (1.37): $d \equiv d\mathbf{x} \cdot \nabla$.

22.2 Gauge-invariant actions

Gauge field. The simplest gauge-invariant action, function of the gauge field \mathbf{A}_μ , and generalization of the Abelian action (21.7) for $m = 0$, has the form

$$\mathcal{S}(\mathbf{A}) = -\frac{1}{4e^2} \int d^d x \sum_{\mu,\nu} \text{tr} \mathbf{F}_{\mu\nu}^2(x), \quad (22.19)$$

where e is the gauge coupling constant. Note that we have not added a mass term to the gauge action, by contrast with the Abelian action (21.1). Indeed, we show in Section A22 that, in the non-Abelian case, the zero mass limit is singular. We have also chosen the normalization of the gauge field, in such a way that all geometric quantities become independent of the gauge coupling constant. The sign in front of the action takes into account that, with our definition, the matrix $\mathbf{F}_{\mu\nu}$ is anti-Hermitian.

Two remarks are immediately in order:

(i) In contrast with the Abelian case, because the gauge field transforms non-trivially under the group, as equation (22.6) shows (the gauge field is ‘charged’), the curvature tensor $\mathbf{F}_{\mu\nu}$ is not gauge-invariant, and thus not directly a physical observable. The action (22.19) is no longer a free-field action; the gauge field has self-interactions and even the spectrum of the pure gauge action is not perturbative (some analytic results can be obtained in dimension 2).

We indicate, in Chapter 25, how lattice gauge theory provides a framework for non-perturbative investigations. The difference between Abelian and non-Abelian theory is reminiscent of the non-linear σ model (Chapter 19).

(ii) Like in the Abelian case, the action, because it is gauge invariant, does not provide a dynamics to the degrees of freedom of the gauge field that correspond to gauge transformations and, therefore, some *gauge fixing* is required.

Matter fields. For boson fields transforming as shown in equation (22.1), and taking into account the transformation (22.10) of the covariant derivatives, one verifies that the action

$$\mathcal{S}_B(\phi) = \int d^d x \left[\sum_\mu (\mathbf{D}_\mu \phi(x))^\dagger \mathbf{D}_\mu \phi(x) + V(\phi(x)) \right],$$

is gauge invariant if $V(\phi)$ is a group-invariant function of the scalar field ϕ .

Similarly, the simplest gauge-invariant action for fermions, transforming by $\mathcal{R}(G)$ has the form,

$$\mathcal{S}_F(\bar{\psi}, \psi) = - \int d^d x \bar{\psi}(x) (\mathbf{D} + M) \psi(x).$$

22.3 Hamiltonian formalism. Quantization in the temporal gauge

Notation. In this section, we work with a *real-time formalism*, denoting by $t \equiv x_0 = ix_d$ time and the corresponding field component by $\mathbf{A}_0 = -i\mathbf{A}_d$. We use the notation \dot{Q} for the time derivative of Q . Space components will carry Roman indices (\mathbf{A}_i, x_i) .

We consider a gauge theory without matter, corresponding to the action (22.19), because the extension of all arguments to a general gauge-invariant theory is simple. We show that non-Abelian gauge theories can be quantized, using a simple Hamiltonian formalism, in the temporal or Weyl’s gauge by the method explained in the Abelian case in Section 21.5.2 [218]. This leads to a QFT that, at least at the formal level, is unitary, because it corresponds to a Hermitian Hamiltonian.

22.3.1 Classical field equations

From the continuation to real time of the action (22.19), one derives the classical field equation,

$$\sum_{\mu} \mathbf{D}_{\mu} \mathbf{F}^{\mu\nu}(t, x) = 0, \quad (22.20)$$

with $x \in \mathbb{R}^{d-1}$, in which the explicit form of \mathbf{D}_{μ} is given by equation (22.13).

The equation (22.20) does not lead to a standard quantization, because, as we have already discussed in the Abelian case, the action does not depend on $\dot{\mathbf{A}}_0$, the time derivative of \mathbf{A}_0 . Thus, again, \mathbf{A}_0 is not a dynamical variable; the \mathbf{A}_0 field equation is a constraint equation that can be used to eliminate \mathbf{A}_0 from the action. However, in the absence of a mass term, as in the Abelian case, the reduced action does not depend on all space components of the gauge field. Only the combination $\sum_j [\delta_{ij} - D_i(D_{\perp}^2)^{-1}D_j] \dot{\mathbf{A}}_j$ appears (D_{\perp}^2 is the covariant space Laplacian). But, in contrast with the Abelian case, the projector acting on \mathbf{A}_i depends on the field itself and, therefore, the procedure which led to Coulomb's gauge does not work here, at least in its simplest form. Therefore, we discuss the quantization directly in the temporal (or Weyl's) gauge.

22.3.2 Temporal or Weyl's gauge

As in the Abelian case, we note that if $\mathbf{A}_{\mu}(t, x)$ is a solution of equation (22.20), any gauge transform of $\mathbf{A}_{\mu}(t, x)$ is also a solution. To describe all solutions, one can thus make a gauge section, that is, consider a section in the space of all gauge fields that intersects all gauge orbits once. We then represent all solutions by an element of the section and a gauge transformation. A gauge condition is especially well-suited to the construction of a Hamiltonian formalism [217],

$$\mathbf{A}_0(t, x) = 0, \quad (22.21)$$

which defines the *temporal gauge*.

The field equation (22.20), after separation of the space and time components, is equivalent to

$$\mathbf{D}_0 \mathbf{F}_{0k} - \sum_l \mathbf{D}_l \mathbf{F}_{lk} = 0, \text{ and } \sum_l \mathbf{D}_l \mathbf{F}_{l0} = 0. \quad (22.22)$$

The indices k, l vary only from 1 to $(d-1)$ and correspond to space components. In the gauge (22.21), the field equations simplify and become

$$\dot{\mathbf{E}}_k = \sum_l \mathbf{D}_l \mathbf{F}_{lk}, \quad \text{with} \quad \mathbf{E}_k = -\dot{\mathbf{A}}_k/e^2, \quad \text{and} \quad (22.23)$$

$$\sum_l \mathbf{D}_l \mathbf{E}_l = 0. \quad (22.24)$$

Equation (22.23) is a dynamical equation that can be directly derived from the initial Lagrangian in which the condition (22.21) has been used:

$$\mathcal{L}(\mathbf{A}_k) = -\text{tr} \int d^{d-1}x \left[\frac{1}{2e^2} \sum_k \dot{\mathbf{A}}_k^2(t, x) - \frac{1}{4e^2} \sum_{k,l} \mathbf{F}_{kl}^2(t, x) \right]. \quad (22.25)$$

The expression has the form of a usual Lagrangian for the space components of the gauge field: \mathbf{E}_k is the conjugated momentum of \mathbf{A}_k and the corresponding Hamiltonian then is

$$\mathcal{H}(\mathbf{E}, \mathbf{A}) = -\text{tr} \int d^{d-1}x \left[\frac{e^2}{2} \sum_k \mathbf{E}_k^2(x) + \frac{1}{4e^2} \sum_{k,l} \mathbf{F}_{kl}^2(x) \right]. \quad (22.26)$$

By contrast, equation (22.24) is a constraint equation, a non-Abelian generalization of Gauss's law. The only relevant solutions of the field equations are those that satisfy the constraint. The Poisson brackets of the equation with the Hamiltonian vanish. Therefore, the equation (22.24) is compatible with the classical motion. This is a consequence of the property that the gauge condition (22.21) is left invariant by time-independent gauge transformations. Therefore, time-independent gauge transformations form a symmetry group of the Lagrangian (22.25) and thus of the Hamiltonian (22.26). The quantities $\sum_l \mathbf{D}_l \mathbf{E}_l$ are the generators, in the sense of Poisson brackets, of the symmetry group.

These arguments generalize to the quantum theory; the quantum operators $\sum_l \mathbf{D}_l \mathbf{E}_l$ generators of a symmetry group, commute with the Hamiltonian. The space of admissible physical states $\Psi(\mathbf{A})$ is restricted by the quantum form of Gauss's law:

$$\sum_l \mathbf{D}_l \mathbf{E}_l \Psi(\mathbf{A}) \equiv \sum_l \mathbf{D}_l \frac{1}{i} \frac{\delta}{\delta \mathbf{A}_l(x)} \Psi(\mathbf{A}) = 0. \quad (22.27)$$

The equation implies that physical states are gauge invariant: they belong to the invariant sector of the symmetry group, a subspace left invariant by quantum evolution.

Euclidean field integral. Quantization in the temporal gauge, as in the Abelian case, then follows conventional lines. Returning to the *Euclidean formalism*, we conclude that the partition function can be written as

$$\mathcal{Z} = \int [d\mathbf{A}_\mu] \prod_x \delta(\mathbf{A}_d(x)) \exp \left[\frac{1}{4e^2} \int d^d x \sum_{\mu,\nu} \text{tr} \mathbf{F}_{\mu\nu}^2(x) \right]. \quad (22.28)$$

Note that, at zero temperature, the perturbative vacuum is automatically gauge invariant and Gauss's law plays no role. This is no longer true at finite temperature.

Non-covariance. The theory we have constructed is not explicitly space-time covariant and this leads to serious difficulties, as we have already emphasized in the Abelian case (see Section 21.5). In particular, in the temporal gauge, the theory is not renormalizable in the sense of power counting. Indeed, the (Euclidean) propagator in this gauge, in the Fourier representation, is proportional to (equation (21.63))

$$\tilde{W}_{ij}^{(2)}(\mathbf{k}_\perp, k_d) = \frac{1}{k^2} \left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}_\perp^2} \right) + \frac{1}{k_d^2} \frac{k_i k_j}{\mathbf{k}_\perp^2}, \quad (22.29)$$

in which \mathbf{k}_\perp is the 'space' part of \mathbf{k} , does not decrease, at k_d fixed, for large spatial momenta $|\mathbf{k}_\perp|$.

Therefore, in the next section, we show that gauge-invariant observables can equivalently be calculated using another quantum action, which leads to a theory that is explicitly covariant and renormalizable by power counting.

22.4 Covariant gauges

Covariant gauges are much better suited than the temporal gauge (22.28) for perturbative calculations. For this purpose, we now use a method that is explained more thoroughly in Sections 26.1.1 and 26.2. To implement the covariant gauge condition

$$\nabla \cdot \mathbf{A}(x) - \boldsymbol{\nu}(x) = 0, \quad (22.30)$$

where $\boldsymbol{\nu}(x)$ is an arbitrary field, and an anti-Hermitian matrix belonging to the Lie algebra of $\mathcal{R}(G)$, we start from the equation

$$E(\mathbf{g}, x) \equiv \nabla \cdot \mathbf{A}^g(x) - \boldsymbol{\nu}(x) = 0, \quad (22.31)$$

where \mathbf{A}^g is the gauge transform of \mathbf{A} (equation (22.6)) by \mathbf{g} . Equation (22.31) determines the space-dependent group element $\mathbf{g}(x)$ as a function of $\mathbf{A}(x)$ and $\boldsymbol{\nu}(x)$. If $\boldsymbol{\nu}(x)$ is a stochastic field, then $\mathbf{g}(x)$ has a stochastic distribution.

We assume that the equation (22.31) has a unique solution, which is equivalent to assert that, in the space of gauge fields, the surface defined by equation (22.31) intersects once and only once all gauge orbits. This is certainly possible for small fields and, thus, in perturbation theory (see the remark at the end of the section concerning this problem beyond perturbation theory).

The Jacobian J of the transformation $E(\mathbf{g}) \mapsto \mathbf{g}$ is related to the variation of the equation with respect to \mathbf{g} . For $\delta\mathbf{g}(x) = \boldsymbol{\omega}(x)\mathbf{g}(x)$, $\boldsymbol{\omega}(x)$ belonging to the Lie algebra,

$$\delta E(\mathbf{g}, x) = -\nabla \cdot \mathbf{D}(\mathbf{A}^g)\boldsymbol{\omega}(x) \Rightarrow J = \det(-\nabla \cdot \mathbf{D}(\mathbf{A}^g)), \quad (22.32)$$

where $\mathbf{D}_\mu(\mathbf{A}^g) = \partial_\mu + \mathbf{A}_\mu^g$. We also introduce a boson field $\boldsymbol{\lambda}$, to implement equation (22.31) and spinless fermions $\bar{\mathbf{C}}$ and \mathbf{C} , the Faddeev–Popov ‘ghosts’ [219], to express the Jacobian J , all transforming under the adjoint representation.

These relations can be combined into the identity (a version of equation (26.15))

$$1 = \int [d\mathbf{g} d\bar{\mathbf{C}} d\mathbf{C} d\boldsymbol{\lambda}] \exp[-\mathcal{S}_{\text{gauge}}(\mathbf{A}^g, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})], \quad (22.33)$$

with

$$\begin{aligned} \mathcal{S}_{\text{gauge}}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \\ = \int d^d x \text{tr} \{ \boldsymbol{\lambda}(x) [\nabla \cdot \mathbf{A}(x) - \boldsymbol{\nu}(x)] + \mathbf{C}(x) \nabla \cdot \mathbf{D}(\mathbf{A}) \bar{\mathbf{C}}(x) \}. \end{aligned} \quad (22.34)$$

We introduce the identity (22.33) into the representation (22.28) of the partition function and obtain

$$\mathcal{Z} = \int [d\mathbf{g} d\bar{\mathbf{C}} d\mathbf{C} d\boldsymbol{\lambda} d\mathbf{A}] \delta(\mathbf{A}_d) \exp[-\mathcal{S}(\mathbf{A}) - \mathcal{S}_{\text{gauge}}(\mathbf{A}^g, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})]. \quad (22.35)$$

We then change variables $\mathbf{A}_\mu^g \mapsto \mathbf{A}_\mu$. The gauge action $\mathcal{S}(\mathbf{A})$ (equation (22.19)) and the integration measure over fields are gauge invariant. Only the gauge condition $\delta(\mathbf{A}_d)$ is affected. Changing \mathbf{g} into \mathbf{g}^{-1} , we find

$$\mathcal{Z} = \int [d\bar{\mathbf{C}} d\mathbf{C} d\boldsymbol{\lambda} d\mathbf{A}] \exp[-\mathcal{S}(\mathbf{A}) - \mathcal{S}_{\text{gauge}}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})] \int [d\mathbf{g}] \delta(\mathbf{A}_d^g). \quad (22.36)$$

We first integrate over the group field $\mathbf{g}(x)$. The result of the integral $\int [d\mathbf{g}] \delta(\mathbf{A}_d^g)$ is gauge invariant. One can thus calculate it only for fields satisfying the gauge condition $\mathbf{A}_d = 0$. Then, $\mathbf{A}_d^g(x) = \mathbf{g}(x)\partial_d\mathbf{g}^{-1}(x)$, and the integral is a constant, independent of \mathbf{A} .

This leads to the field integral representation

$$\mathcal{Z} = \int [d\mathbf{A} d\bar{\mathbf{C}} d\mathbf{C} d\boldsymbol{\lambda}] \exp [-\mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})], \quad (22.37)$$

where the quantum action $\mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ is given by

$$\mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathcal{S}(\mathbf{A}) + \mathcal{S}_{\text{gauge}}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu}). \quad (22.38)$$

Since the partition function is independent of $\boldsymbol{\nu}$, one can average over the ‘noise’ field $\boldsymbol{\nu}(x)$ with the Gaussian distribution (ξe^2 is a positive constant, width of the distribution),

$$[d\rho(\boldsymbol{\nu})] = [d\boldsymbol{\nu}] \exp \left[\frac{1}{2\xi e^2} \int d^d x \operatorname{tr} \boldsymbol{\nu}^2(x) \right]. \quad (22.39)$$

The $\boldsymbol{\nu}$ -averaged partition function \mathcal{Z} becomes

$$\mathcal{Z} = \int [d\mathbf{A} d\bar{\mathbf{C}} d\mathbf{C} d\boldsymbol{\lambda}] \exp [-\mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda})], \quad (22.40)$$

where \mathcal{S} is the local action,

$$\mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}) = \mathcal{S}(\mathbf{A}) + \mathcal{S}_{\text{gauge}}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}), \quad (22.41)$$

with

$$\mathcal{S}_{\text{gauge}}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}) = \int d^d x \operatorname{tr} \left[\frac{1}{2} \xi e^2 \boldsymbol{\lambda}^2(x) + \boldsymbol{\lambda}(x) \nabla \cdot \mathbf{A}(x) + \mathbf{C}(x) \nabla \cdot \mathbf{D} \bar{\mathbf{C}}(x) \right]. \quad (22.42)$$

Except in the limit in which ξ vanishes, it is also possible to integrate over $\boldsymbol{\lambda}(x)$ to find the new quantum action

$$\begin{aligned} & \mathcal{S}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}) \\ &= \int d^d x \operatorname{tr} \left[-\frac{1}{e^2} \left(\frac{1}{4} \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2\xi} (\nabla \cdot \mathbf{A}(x))^2 \right) + \mathbf{C}(x) \nabla \cdot \mathbf{D} \bar{\mathbf{C}}(x) \right]. \end{aligned} \quad (22.43)$$

This form of the quantum action is better-suited for perturbative calculations, although geometric properties of the action are more apparent in expression (22.41), in particular, the BRST symmetry, as we will show.

Therefore, we have established the formal equivalence between the two expressions (22.40) and (22.28) of the partition function. The formal equivalence with other gauges can be proven by a similar method. The obvious drawback of the covariant gauge, which leads to a covariant, local, and renormalizable theory, is the lack of explicit positivity, and thus unitarity. In particular, Faddeev–Popov fermions, being spinless, do not obey to the spin–statistics connection and are, thus, non-physical.

Gribov’s ambiguity. As pointed out by Gribov [247], in contrast with the Abelian case, depending on the value of the gauge field $\mathbf{A}_\mu(x)$, the gauge condition (22.31) may not always have a unique solution in $\mathbf{g}(x)$, a problem called Gribov’s ambiguity. When two solutions merge, an eigenvalue of the operator $\partial_\mu \mathbf{D}_\mu(\mathbf{A})$ vanishes. This implies that the representation (22.40) is not meaningful beyond perturbation theory. The same ambiguity has been shown to arise for a large class of gauge conditions.

22.4.1 BRST symmetry

One consequence of the quantization procedure is that the action (22.41) is no longer gauge invariant. On the other hand, the general analysis of Section 26.2, applied to the quantization procedure, shows that the action now has a BRST symmetry [236, 237], consequence of the stochastic dynamics given to the degrees of freedom of the gauge group variables. To understand the form of the BRST transformations, it is convenient to separate the gauge group degrees of freedom, which induce the BRST symmetry, from the other degrees of freedom of the gauge field, which play no role, setting

$$\mathbf{A}_\mu(x) = \mathbf{B}_\mu^g(x),$$

where $\mathbf{B}_\mu(x)$ satisfies the gauge condition (22.31): $\nabla \cdot \mathbf{B}(x) = \boldsymbol{\nu}(x)$. The general analysis of Section 26.2 shows that the BRST transformations in the case of group manifolds have the form (ε is an additional element of the Grassmann algebra),

$$\begin{cases} \delta_{\text{BRST}} \mathbf{g}(x) = \varepsilon \bar{\mathbf{C}}(x) \mathbf{g}(x), & \delta_{\text{BRST}} \bar{\mathbf{C}}(x) = \varepsilon \bar{\mathbf{C}}^2(x), \\ \delta_{\text{BRST}} \mathbf{C}(x) = \varepsilon \boldsymbol{\lambda}(x), & \delta_{\text{BRST}} \boldsymbol{\lambda}(x) = 0. \end{cases} \quad (22.44)$$

The field $\mathbf{B}_\mu(x)$ has a dynamics provided by the gauge action and is BRST invariant,

$$\delta_{\text{BRST}} \mathbf{B}_\mu(x) = 0, \quad (22.45)$$

and this implies that the gauge action (22.19) is also BRST invariant.

Calculating the effect of a BRST transformation on the field \mathbf{A}_μ , one then finds

$$\delta_{\text{BRST}} \mathbf{A}_\mu(x) = \delta \mathbf{B}_\mu^g(x) = -\varepsilon \mathbf{D}_\mu \bar{\mathbf{C}}(x), \quad (22.46)$$

which is also an infinitesimal gauge transformation.

This result is not surprising: the transformations (22.44) correspond to an infinitesimal change in the gauge group degrees of freedom, and thus to an infinitesimal gauge transformation for \mathbf{A}_μ . Equation (22.46) is also similar to the transformation of the current associated with group elements, as discussed in Section 35.6 (equations (35.68–35.70)).

Since, for all physical fields, the BRST transformations have the form of infinitesimal gauge transformations, all gauge-invariant quantities are also BRST invariant.

The action of a BRST transformation on a function of $\mathbf{A}, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}$ can be reproduced by the action of the BRST functional differential operator

$$\mathcal{D} = \int d^d x \operatorname{tr} \left[- \sum_\mu \mathbf{D}_\mu \bar{\mathbf{C}}(x) \frac{\delta}{\delta \mathbf{A}_\mu(x)} + \bar{\mathbf{C}}^2(x) \frac{\delta}{\delta \bar{\mathbf{C}}(x)} + \boldsymbol{\lambda}(x) \frac{\delta}{\delta \mathbf{C}(x)} \right], \quad (22.47)$$

The BRST symmetry of the quantized action is expressed by the equation

$$\mathcal{D}\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}) = 0. \quad (22.48)$$

It is simple to verify directly that $\mathcal{D}^2 = 0$, as one also infers from the general discussion of Section 26.2, and thus, \mathcal{D} is a *cohomology operator*. Equation (22.48) then implies that the expression (22.34) is BRST *closed*. Furthermore, it is even BRST *exact*, since it can be written as

$$\mathcal{S}_{\text{gauge}} = \mathcal{D} \int d^d x \operatorname{tr} \mathbf{C}(x) [\nabla \cdot \mathbf{A}(x) - \boldsymbol{\nu}(x)].$$

Similarly, after integration over the ν field, the expression (22.42)) can be written as

$$\mathcal{S}_{\text{gauge}} = \mathcal{D} \int d^d x \operatorname{tr} \mathbf{C}(x) [\nabla \cdot \mathbf{A}(x) + \frac{1}{2}\xi e^2 \boldsymbol{\lambda}(x)]. \quad (22.49)$$

We show in Chapter 26, in a general framework, that the BRST symmetry, expressed by equations (22.44) and (22.46), implies that the action (22.41) satisfies the more general Zinn-Justin (ZJ) equation [40], which ensures its structural stability under renormalization.

22.5 Perturbation theory, regularization

For perturbative calculations, it is useful to express all expressions in component form (see Section 13.1.1 for conventions).

The component form. We introduce a basis $\{\mathbf{t}^a\}$ of anti-Hermitian matrices, with elements t_{ij}^a , generators of the Lie algebra in the representation \mathcal{R} . Then, the gauge field has the expansion

$$\mathbf{A}_\mu(x) = \sum_a A_\mu^a(x) \mathbf{t}^a. \quad (22.50)$$

The matrix elements of the covariant derivative (22.8) read

$$(\mathbf{D}_\mu)_{ij} = \partial_\mu \delta_{ij} + \sum_a A_\mu^a(x) t_{ij}^a. \quad (22.51)$$

Equation (22.13), after expansion of $\boldsymbol{\omega}$ on the basis, involves the structure constants f_{abc} of the Lie algebra (which can be chosen completely antisymmetric for orthogonal or unitary groups). It takes the form

$$-\delta A_\mu^a(x) = \partial_\mu \omega_a(x) + \sum_{b,c} f_{bca} A_\mu^b(x) \omega_c(x). \quad (22.52)$$

The equation also yields the form of the covariant derivative in the adjoint representation.

The curvature tensor has an expansion on the basis of the form

$$\mathbf{F}_{\mu\nu}(x) = \sum_a F_{\mu\nu}^a(x) \mathbf{t}^a, \quad (22.53)$$

and, therefore,

$$F_{\mu\nu}^a(x) = \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + \sum_{b,c} f_{bca} A_\mu^b(x) A_\nu^c(x). \quad (22.54)$$

This last expression is independent of the group representation.

Compared with the Abelian theory, the new features of the action (22.43) are the presence of gauge field self-interactions and ghost terms.

We redefine the coupling constant e in such a way that the gauge field action, in the covariant gauge (22.43), takes the form

$$\mathcal{S}(A_\mu^a) = \frac{1}{e^2} \int d^d x \left[\frac{1}{4} \sum_{a,\mu,\nu} F_{\mu\nu}^a(x) F_{\mu\nu}^a(x) + \frac{1}{2\xi} \sum_a (\nabla \cdot \mathbf{A}^a(x))^2 \right], \quad (22.55)$$

The ghost-field action becomes

$$\mathcal{S}_{\text{ghost}} = \int d^d x \sum_{a,c,\mu} C^a(x) \partial_\mu \cdot \left(\delta_{ac} \partial_\mu + \sum_b f_{bca} A_\mu^b(x) \right) \bar{C}^c(x). \quad (22.56)$$

22.5.1 Feynman rules in the Fourier representation

In the covariant gauge of Section 22.4, the gauge field propagator is (equation (21.35))

$$[\tilde{\Delta}_\xi]_{\mu\nu}^{ab}(k) = e^2 \delta_{ab} \left(\frac{\delta_{\mu\nu}}{k^2} + (\xi - 1) \frac{k_\mu k_\nu}{(k^2)^2} \right). \quad (22.57)$$

In four dimensions, the gauge field has dimension 1. The ghost-field propagator is

$$\tilde{\Delta}^{ab}(k) = \delta_{ab}/k^2. \quad (22.58)$$

One can assign to the ghost fields a canonical dimension 1 in four dimensions. The interaction terms have all dimension 4 and, therefore, the theory is renormalizable by power counting in four dimensions. The power counting for matter fields is the same as in the Abelian theory.

Compared with the Abelian theory, the non-Abelian theory has three new vertices generated by the gauge field self-interactions and the interaction with the ghost fields. The A^3 three-point vertex is

$$[\tilde{\Gamma}^{(3)}]_{\mu\nu\rho}^{abc}(p, q, r) = \frac{i}{e^2} f_{abc} [(r - q)_\mu \delta_{\nu\rho} + (p - r)_\nu \delta_{\rho\mu} + (q - p)_\rho \delta_{\mu\nu}], \quad (22.59)$$

and the A^4 four-point vertex is given by

$$\begin{aligned} [\tilde{\Gamma}^{(4)}]_{\mu\nu\rho\sigma}^{abcd} &= \frac{1}{e^2} [f_{eab} f_{ecd} (\delta_{\mu\rho} \delta_{\nu\sigma} - \delta_{\mu\sigma} \delta_{\nu\rho}) + f_{eac} f_{ebd} (\delta_{\mu\nu} \delta_{\rho\sigma} - \delta_{\mu\sigma} \delta_{\nu\rho}) \\ &\quad + f_{ead} f_{ecb} (\delta_{\mu\rho} \delta_{\nu\sigma} - \delta_{\mu\nu} \delta_{\sigma\rho})]. \end{aligned} \quad (22.60)$$

All terms are obtained from the first one by exchanging the indices to make the correlation function totally symmetric. Finally, the $C\bar{C}A$ ghost gauge field vertex is

$$[\tilde{\Gamma}^{(3)}]_\mu^{abc}(p, q, r) = -i f_{abc} p_\mu. \quad (22.61)$$

Note that, in a generic gauge, the two ghost fields play a different role. In a graphic representation of Feynman diagrams, ghost propagator lines are oriented. However, in Landau's gauge ($\xi = 0$), because $\nabla \cdot \mathbf{A}$ vanishes, the vertex can be antisymmetrized, and a symmetry between ghost fields is established.

Matter fields. The coupling to matter fields differs from the Abelian case only by geometric factors corresponding to group indices. For example, the coupling to fermions generated by the covariant derivative (22.51) is simply $\gamma_\mu t_{ij}^a$.

Infrared divergences. In the covariant gauge, and in the absence of a Higgs mechanism which provides a mass to gauge fields, only the gauge $\xi = 1$, Feynman's gauge, leads to a theory that is obviously infrared (IR) finite. In contrast to the Abelian case, it is impossible to give an explicit mass to the gauge field and to then construct a theory that is both unitary and renormalizable (for more details see Section A22). However, eventually, to prove the gauge independence of physical observables, we must be able to define a theory for more than one gauge. One way to introduce an IR regulator is to consider the theory in a finite volume. This necessitates a discussion of finite volume effects (see Chapter 32).

22.5.2 Regularization

The problem of regularization in non-Abelian gauge theories has many features in common with the Abelian case, as well as with the non-linear σ -model (Chapter 19). We review the three regularization methods that we have always considered in this work. Dimensional regularization is the most convenient for practical calculations and works in the absence of chiral fermions (γ_5 has no dimensional continuation).

Momentum or Pauli–Villars’s type regularizations [239] work only partially in geometric models, because *they regularize all diagrams except one-loop diagrams*. The regularized gauge action takes the form

$$\mathcal{S}(\mathbf{A}_\mu) = \int d^d x \sum_{\mu,\nu} \text{tr} \mathbf{F}_{\mu\nu}(x) P(\mathbf{D}^2/\Lambda^2) \mathbf{F}_{\mu\nu}(x), \quad (22.62)$$

in which P is a polynomial.

Similarly, the gauge function $\nabla \cdot \mathbf{A}$ is changed into

$$\nabla \cdot \mathbf{A}(x) \mapsto Q(\nabla^2/\Lambda^2) \nabla \cdot \mathbf{A}(x), \quad (22.63)$$

in which P and Q are polynomials of arbitrary high degrees. Therefore, both the gauge-field propagator and the ghost-field propagator can be made arbitrarily convergent. However, as in the Abelian case, the covariant derivatives generate new interactions which are more singular. One verifies that the power counting of one-loop diagrams remains unchanged while all higher-order diagrams can be made superficially convergent. For matter fields, the situation is the same as in the Abelian theory (Section 21.8.1).

The one-loop diagrams are generated by determinants. Boson determinants can be regularized by Schwinger’s representation. Like in the Abelian example, a potential problem is generated by the fermion matter determinant. The addition of auxiliary, non-physical fields cures the problem, including for fermions with global chiral symmetries, but problems arise with local chiral transformations, where *anomalies* are indeed possible (see Section 23.6).

Lattice regularization [221], which is also relevant for non-perturbative calculations, can be used generally (see Chapter 25 for details), since a (non-trivial) method for handling chiral fermions has been found (related to Ginsparg–Wilson’s relation [98]), but has also problems with chiral gauge theories in the integration measure.

We emphasize again that the preservation of gauge symmetry is necessary for the cancellation of non-physical states in physical amplitudes, and thus essential for the physical consistency of the gauge theory.

22.6 The non-Abelian Higgs mechanism

We have already discussed the Abelian Higgs mechanism in Section 21.13. The basic idea is the same in non-Abelian theories: the spontaneous breaking of a global symmetry associated with a gauge invariance leads to masses for gauge fields [242], without generating massless Goldstone particles [248, 226, 243, 249]. Simply, because the group structure is richer, a number of different situations may arise.

We consider a classical gauge-invariant action, expressed here in terms of real fields, for a gauge field coupled to a scalar boson ϕ transforming under an orthogonal representation of the symmetry group (the generators \mathbf{t}^α are antisymmetric matrices):

$$\mathcal{S}(\mathbf{A}, \phi) = \int d^d x \left[-\frac{1}{4e^2} \sum_{\mu u, \nu} \text{tr} \mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2} \sum_\mu \mathbf{D}_\mu \phi(x) \cdot \mathbf{D}_\mu \phi(x) + V(\phi(x)) \right]. \quad (22.64)$$

We assume that the symmetric potential $V(\phi)$ has non-symmetric minima. In the absence of gauge symmetry, this is the situation which we have already analysed in Section 13.4. Since the spectrum in the classical limit depends on the group structure and the representation content of the field ϕ , we consider here only two families of examples.

22.6.1 Simple compact Lie symmetry groups

First, we assume that G , the symmetry group of the action, is simple and is thus also the gauge group. Moreover, for simplicity, we assume that the field ϕ belongs to an irreducible representation. We denote by \mathbf{v} one constant field that minimizes the potential and H the stabilizer of \mathbf{v} , the subgroup of G that leaves \mathbf{v} invariant. We separate the generators \mathbf{t}^α of G in the matrix representation into two subsets $\alpha \leq p$, corresponding to the Lie algebra $\mathcal{L}(H)$ of the subgroup H , and the complementary set $\mathcal{L}(G/H)$. We parametrize the scalar field $\phi(x)$ as

$$\phi(x) = \exp\left(\sum_{\alpha > p} \theta^\alpha(x) \mathbf{t}^\alpha\right) (\mathbf{v} + \boldsymbol{\rho}(x)), \quad (22.65)$$

in which the vectors $\boldsymbol{\rho}$ and $\{t_{ij}^\alpha v_j\}$ span two orthogonal subspaces. The transformation

$$\phi(x) \longmapsto \{\theta^\alpha(x), \boldsymbol{\rho}(x)\}, \quad (22.66)$$

is such that the new fields $\boldsymbol{\rho}(x)$ and $\theta^\alpha(x)$ can be expanded in powers of $\phi(x) - \mathbf{v}$.

In the absence of gauge symmetry, we have used the representation (22.65) to show that the fields $\theta^\alpha(x)$ correspond to massless Goldstone modes induced by the spontaneous breaking of the G -symmetry (see Section 13.4).

Here, equation (22.65) can also be viewed as a local group transformation relating the two fields ϕ and $\boldsymbol{\rho} + \mathbf{v}$. If we perform on the field \mathbf{A}_μ a gauge transformation of the form (22.6) with

$$\mathbf{g}(x) = \exp\left(\sum_{\alpha > p} \mathbf{t}^\alpha \theta^\alpha(x)\right), \quad (22.67)$$

we eliminate the fields θ^α from the action completely. In fact, we have fixed (at least partially) the gauge. If we now examine the scalar field contribution to the action, we see that, for $\boldsymbol{\rho} = 0$, it reduces to a mass term for the gauge field:

$$\frac{1}{2} \sum_\mu \mathbf{D}_\mu \phi \cdot \mathbf{D}_\mu \phi|_{\boldsymbol{\rho}, \theta^\alpha=0} = \frac{1}{2} \sum_{\alpha, \beta, \mu} \sigma_{\alpha\beta} A_\mu^\alpha A_\mu^\beta, \quad (22.68)$$

with the mass matrix

$$e^2 \sigma_{\alpha\beta} = e^2 \sum_{i,j,k} t_{ij}^\alpha v_j t_{ik}^\beta v_k. \quad (22.69)$$

The matrix $\sigma_{\alpha\beta}$ is positive and has a rank equal to the number of generators of $\mathcal{L}(G/H)$, which is also the number of fields θ^α , that is, the number of would-be Goldstone bosons. We conclude that the spontaneous breaking of the G -symmetry generates no Goldstone bosons but, by contrast, gives masses to all gauge fields except those which are associated with the unbroken subgroup H . In particular, when the symmetry is completely broken, all components of the gauge field acquire a mass.

If one considers directly the classical action obtained after the gauge transformation associated with group element (22.67), the set of massive vector fields can be quantized in a completely standard way [249]. However, as in the Abelian example, the quantized theory is not renormalizable.

22.6.2 The $G \times G$ symmetry with simple compact Lie groups

Another possibility is that the symmetry group of the action is the direct product of the gauge group G by another group G' . We consider here only the simplest example where the symmetry group is $G \times G$, and G is a simple unitary group. We assume that the scalar boson field ϕ is a matrix transforming under $G \times G$ by

$$\phi' = \mathbf{g}_1 \phi \mathbf{g}_2^{-1}, \quad (22.70)$$

in which \mathbf{g}_1 and \mathbf{g}_2 are two elements of G in a matrix representation, and only the group acting on the left is gauged:

$$\mathbf{D}_\mu \phi(x) = (\partial_\mu + \mathbf{A}_\mu(x)) \phi(x).$$

We further assume that one minimum of the potential is proportional to the unit matrix $\phi = v \mathbf{1}$, in such a way that the subgroup H is isomorphic to G with elements of the form (\mathbf{g}, \mathbf{g}) . We recall that the coset space $G \times G/G$ is then a symmetric space (see Section A29.4). As in the previous example, gauge transformations make it possible to eliminate would-be Goldstone bosons from the action. In this specific example, all components of the gauge field acquire the same mass m_A , because the symmetry corresponding to the gauge field is completely broken, and a remaining diagonal G symmetry survives:

$$\sum_\mu \text{tr} (\mathbf{D}_\mu \phi)^\dagger (\mathbf{D}_\mu \phi) |_{\phi=v\mathbf{1}} = -v^2 \sum_\mu \text{tr} \mathbf{A}_\mu^2 \Rightarrow m_A = |ev|. \quad (22.71)$$

After the gauge transformation specified by equation (22.67), the action only contains physical degrees of freedom, and the quantization of all vector fields is straightforward, hence the name of *unitary gauge*. From the point of view of the initial theory, the gauge has been completely fixed. We have constructed an action for massive vector fields transforming under the adjoint representation of a symmetry group G . The corresponding QFT is not renormalizable by power counting. However, by contrast with the massive vector QFT we examine in Section A22, the suitable addition of some scalar fields makes this theory equivalent, at least for physical observables, with a renormalizable theory with additional non-physical degrees of freedom.

The non-linear σ model limit. If we formally take the non-linear model limit, that is, send the masses of all remaining scalar fields towards infinity at v fixed, we obtain an action for a self-interacting massive vector field (see Section A22).

22.6.3 The $SU(2) \times SU(2)$ example

We discuss more specifically the important example of the $SU(2)$ group, because it can be considered as a simplified version of the weak-electromagnetic sector of the Standard Model (SM) of particle physics, which is described in Section 23.1. We choose for scalar field ϕ a 2×2 complex matrix transforming under the $(1/2, 1/2)$ representation of $SU(2) \times SU(2)$. We know that the representation can be reduced, and we restrict the matrix to the form (see Section 13.6)

$$\phi(x) = \frac{1}{2} (\sigma(x) + i \boldsymbol{\tau} \cdot \boldsymbol{\pi}(x)), \quad (22.72)$$

in which σ and $\boldsymbol{\pi}$ are real fields and the $\boldsymbol{\tau}$ matrices are identical to the $\boldsymbol{\sigma}$ Pauli matrices defined in Section A12.1.4.

We represent the gauge field as a three-component real vector \mathbf{A}_μ , and the covariant derivative acts as

$$\mathbf{D}_\mu \phi = (\partial_\mu + \frac{1}{2} i \mathbf{A}_\mu \cdot \boldsymbol{\tau}) \phi. \quad (22.73)$$

We define the curvature tensor $\mathbf{F}_{\mu\nu}$, a three vector, by

$$[\mathbf{D}_\mu, \mathbf{D}_\nu] = \frac{i}{2} \mathbf{F}_{\mu\nu} \cdot \boldsymbol{\tau} \Rightarrow \mathbf{F}_{\mu\nu}(x) = \partial_\mu \mathbf{A}_\nu(x) - \partial_\nu \mathbf{A}_\mu(x) - \mathbf{A}_\mu(x) \times \mathbf{A}_\nu(x). \quad (22.74)$$

The simplest action can then be written as

$$\begin{aligned} \mathcal{S}(\mathbf{A}, \phi) = \int d^d x & \left[\frac{1}{4e^2} \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}^2(x) + \sum_\mu \text{tr}(\mathbf{D}_\mu \phi(x))^\dagger \mathbf{D}_\mu \phi(x) + r \text{tr} \phi^\dagger(x) \phi(x) \right. \\ & \left. + \frac{\lambda}{6} (\text{tr} \phi^\dagger(x) \phi(x))^2 \right]. \end{aligned} \quad (22.75)$$

In the $(\sigma, \boldsymbol{\pi})$ parametrization, infinitesimal gauge transformations take the form

$$\begin{cases} \delta \mathbf{A}_\mu(x) = \partial_\mu \boldsymbol{\omega}(x) - \mathbf{A}_\mu(x) \times \boldsymbol{\omega}(x), \\ \delta \sigma(x) = \frac{1}{2} \boldsymbol{\omega}(x) \cdot \boldsymbol{\pi}(x), \\ \delta \boldsymbol{\pi}(x) = -\frac{1}{2} \sigma(x) \boldsymbol{\omega}(x) + \frac{1}{2} \boldsymbol{\omega}(x) \times \boldsymbol{\pi}(x). \end{cases} \quad (22.76)$$

The covariant derivative becomes,

$$\mathbf{D}_\mu \phi(x) \mapsto \begin{cases} \partial_\mu \sigma(x) - \frac{1}{2} \mathbf{A}_\mu(x) \cdot \boldsymbol{\pi}(x) \\ \partial_\mu \boldsymbol{\pi}(x) + \frac{1}{2} \sigma(x) \mathbf{A}_\mu(x) - \frac{1}{2} \mathbf{A}_\mu(x) \times \boldsymbol{\pi}(x). \end{cases}$$

The scalar field action in these variables then reads

$$\begin{aligned} \mathcal{S}_{\text{scalar}}(\mathbf{A}, \phi) = \int d^d x & \left\{ \frac{1}{2} \sum_\mu \left[(\partial_\mu \sigma(x) - \frac{1}{2} \boldsymbol{\pi}(x) \cdot \mathbf{A}_\mu(x))^2 \right. \right. \\ & \left. \left. + \frac{1}{2} (\partial_\mu \boldsymbol{\pi}(x) + \frac{1}{2} \sigma(x) \mathbf{A}_\mu(x) - \frac{1}{2} \mathbf{A}_\mu(x) \times \boldsymbol{\pi}(x))^2 \right] + \tilde{V}(\sigma^2(x) + \boldsymbol{\pi}^2(x)) \right\}, \end{aligned} \quad (22.77)$$

with

$$\tilde{V}(s) = \frac{1}{2} rs + \frac{1}{24} \lambda s^2.$$

Note that, for the potential \tilde{V} , $SU(2) \times SU(2)$ symmetry implies $O(4)$ symmetry. The two groups are locally isomorphic.

As we have already discussed in Section 13.6, if the potential \tilde{V} has degenerate classical minima, the field ϕ has a non-zero expectation value. Without loss of generality, we choose the expectation value of ϕ to be proportional to the unit matrix and, thus, the component σ to have a non-zero expectation value:

$$\langle \sigma \rangle = v > 0.$$

Then the symmetry $SU(2) \times SU(2)$ is broken down to the diagonal $SU(2)$ subgroup. In the absence of gauge fields, the $\boldsymbol{\pi}$ -field becomes a massless Goldstone boson.

Here, the $\boldsymbol{\pi}$ -field can be eliminated by a gauge transformation, in such a way that the total action written in the unitary gauge becomes

$$\begin{aligned} \mathcal{S}(\mathbf{A}, \sigma) = & \int d^d x \left[\frac{1}{4e^2} \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}^2(x) \right. \\ & \left. + \frac{1}{2} (\nabla \sigma(x))^2 + \frac{1}{8} \sigma^2(x) \mathbf{A}^2(x) + \frac{r}{2} \sigma^2(x) + \frac{\lambda}{24} \sigma^4(x) \right]. \end{aligned} \quad (22.78)$$

The action has an $O(3)$ symmetry. From the point of view of the $O(3)$ group, the gauge field \mathbf{A} is a three-vector and the field σ a scalar. In the classical approximation, the gauge field mass m_A is given in terms of the σ -field expectation value v by

$$m_A = |ev|/2, \quad m_\sigma = \sqrt{\lambda/3} v, \quad \text{and thus, } m_\sigma/m_A = \sqrt{4\lambda/3}/|e|. \quad (22.79)$$

22.6.4 Gauge fixing of the Higgs model in a covariant gauge

In the contribution (22.77) to the action, the σ -field expectation value generates a term of the form $\sum_\mu \partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$, which introduces a mixing between the would-be Goldstone boson $\boldsymbol{\pi}$ and the longitudinal part of the vector field. This is a feature already encountered in the Abelian theory (Section 21.13). As suggested by 't Hooft [226], it is possible to use the gauge-fixing function to eliminate such a term. In the $SU(2)$ example, instead of the gauge condition (22.30), one can choose (s is an adjustable constant)

$$F(\mathbf{A}, \boldsymbol{\pi}, x) = \nabla \cdot \mathbf{A}(x) + \frac{1}{2} s \xi \boldsymbol{\pi}(x) - \boldsymbol{\nu}(x) = 0. \quad (22.80)$$

After a Gaussian integration, the contribution to the action, $\mathcal{S}_F = \mathcal{S}_{\text{gauge}} + \mathcal{S}_{\text{ghost}}$, is

$$\mathcal{S}_{\text{gauge}} = \frac{1}{2\xi e^2} \int d^d x (\nabla \mathbf{A}(x) + \frac{1}{2} s \xi \boldsymbol{\pi}(x))^2, \quad (22.81)$$

$$\begin{aligned} \mathcal{S}_{\text{ghost}} = & \int d^d x \left[\sum_\mu \partial_\mu \mathbf{C}(x) \cdot (\partial_\mu \bar{\mathbf{C}}(x) - \mathbf{A}_\mu(x) \times \bar{\mathbf{C}}(x)) \right. \\ & \left. + \frac{s\xi}{4} \mathbf{C}(x) (\sigma(x) \bar{\mathbf{C}}(x) + \boldsymbol{\pi}(x) \times \bar{\mathbf{C}}(x)) \right]. \end{aligned} \quad (22.82)$$

At leading order, the term $\sum_\mu \partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$ is eliminated by the choice $s = e^2 v$. This gauge has two advantages: it decouples the gauge field from the would-be Goldstone field and, therefore, simplifies the propagators; by explicitly breaking the global $SU(2) \times SU(2)$ -symmetry, it generates a mass for the $\boldsymbol{\pi}$ -field, which is no longer a Goldstone boson. In this gauge, denoting by m_A the mass of \mathbf{A}_μ in the tree approximation (equation (22.79)), we find the propagators (equations (21.144)),

$$\begin{aligned} \widetilde{W}_{\mu\nu}^{(2)}(k) = & \frac{e^2 \delta_{\mu\nu}}{k^2 + m_A^2} + \frac{e^2 (\xi - 1) k_{\mu k_\nu}}{(k^2 + m_A^2)(k^2 + \xi m_A^2)}, \\ \widetilde{W}_{\boldsymbol{\pi}\boldsymbol{\pi}}^{(2)}(k) = & \frac{1}{k^2 + \xi m_A^2}, \quad \widetilde{W}_{\mathbf{C}\bar{\mathbf{C}}}^{(2)}(k) = \frac{1}{k^2 + \xi m_A^2}. \end{aligned} \quad (22.83)$$

They have no poles at zero momentum and no IR problem is encountered. Furthermore, all non-physical states have a mass that explicitly depends on the gauge parameter ξ .

Unitarity. This property can be used to prove unitarity of the physical S -matrix: the S -matrix satisfies a generalized unitarity relation in which, in the intermediate states, all particles both physical and non physical must be included. By showing that the S -matrix does not depend on the gauge, one proves simultaneously that the contribution of non-physical states cancels in the intermediate states, and thus the S -matrix is unitary in the physical subspace. A general proof of this kind is given in Chapter 26.

22.6.5 Renormalization

The proof of renormalizability of non-Abelian gauge theories is the same in the unbroken and broken phase [238, 41]. It is based on an invariant regularization (see Section 22.5.2), and a set of Ward–Takahashi (WT) identities satisfied by vertex functions, which express the consequences of the gauge symmetry, under the form of BRST transformations (equations (22.44–22.46)). The form of the early proofs [238, 41], which are not based on the BRST symmetry, is somewhat complicated, and we postpone the discussion to Sections 26.7–26.10, where we derive the form of the renormalized action for a general gauge theory, using the more powerful formalism of the ZJ equation [40].

In the example of the pure gauge action for a simple group G in the covariant ξ -gauge, the renormalized form of the action (22.43) is given by the substitution

$$e^2 \mapsto Z_e e^2, \quad \mathbf{A}_\mu \mapsto Z_A^{1/2} \mathbf{A}_\mu, \quad \xi \mapsto Z_A Z_e^{-1} \xi, \quad \mathbf{C}\bar{\mathbf{C}} \mapsto Z_C \mathbf{C}\bar{\mathbf{C}}. \quad (22.84)$$

This result has a simple interpretation: the gauge structure (22.43) is preserved, and the coefficient of $\nabla \cdot \mathbf{A})^2$ is not renormalized, as in the Abelian theory. However, in contrast with the Abelian theory, the gauge transformation of the gauge field and, more generally, the form of the covariant derivative, are modified by the gauge field renormalization.

A22 Massive Yang–Mills fields

For completeness, we briefly explain why, in contrast with the Abelian example, it is not possible to construct a renormalizable QFT in which a mass is given to the gauge field by directly adding a mass term to the action.

We consider the *real-time* Lagrangian density ($x \in \mathbb{R}^{d-1}$), [244]

$$\mathcal{L}(\mathbf{A}) = -\frac{1}{e^2} \text{tr} \left[-\frac{1}{4} \mathbf{F}_{\mu\nu}(t, x) \mathbf{F}^{\mu\nu}(t, x) + \frac{1}{2} m^2 \mathbf{A}_\mu(t, x) \mathbf{A}^\mu(t, x) \right], \quad (A22.1)$$

where the first term corresponds to a G -gauge symmetry.

Quantization. The first problem one meets is quantization. We know that \mathbf{A}_0 , the time component of the gauge field, has no conjugated momentum and thus is not a dynamical variable. It can be eliminated, using the corresponding field equation. Actually, it is algebraically convenient to first determine the Hamiltonian, and use the equation of motion afterwards. The Hamiltonian here is obtained by performing a Legendre transformation only on the space components \mathbf{A}_i (Roman indices mean space components). The conjugated momenta \mathbf{E}_i are

$$\mathbf{E}_i = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{A}}_i} = -\frac{1}{e^2} \mathbf{F}_{0i}.$$

After an integration by parts, the Hamiltonian density \mathcal{H} can be written as

$$\begin{aligned} \mathcal{H} &= \sum_i \text{tr} \mathbf{E}_i \dot{\mathbf{A}}_i - \mathcal{L} \\ &= \text{tr} \left[- \sum_i \left(\frac{1}{2} e^2 \mathbf{E}_i^2 + \partial_i \mathbf{E}_i \mathbf{A}_0 - \frac{m^2}{2e^2} \mathbf{A}_i^2 \right) - \frac{m^2}{2e^2} \mathbf{A}_0^2 + \frac{1}{4e^2} \sum_{i,j} \mathbf{F}_{ij}^2 \right]. \end{aligned} \quad (A22.2)$$

Using a familiar property (Section 1.8) of the Legendre transformation, one infers the \mathbf{A}_0 equation of motion,

$$0 = -\frac{\partial \mathcal{L}}{\partial \mathbf{A}_0} = \frac{\partial \mathcal{H}}{\partial \mathbf{A}_0} = \frac{m^2}{e^2} \mathbf{A}_0 + \sum_i \partial_i \mathbf{E}_i. \quad (A22.3)$$

The equation determines \mathbf{A}_0 , which can be eliminated from the Hamiltonian. We then express the field integral representation of the evolution operator in terms of the reduced Hamiltonian, integrating over $\{\mathbf{E}_i, \mathbf{A}_i\}$. However, in this expression, the reduced Hamiltonian can be replaced by the initial Hamiltonian (A22.2) provided one integrates also over \mathbf{A}_0 . Indeed, the integral over \mathbf{A}_0 is Gaussian: the integration thus sets \mathbf{A}_0 to the solution of the field equation (A22.3) and yields a constant determinant, because the coefficient of \mathbf{A}_0^2 in expression (A22.2) is field independent. Finally, integrating with the initial Hamiltonian over \mathbf{E}_i at \mathbf{A}_μ fixed, one recovers the initial action. Thus, the partition function can be expressed in terms of the *Euclidean action*,

$$\mathcal{S}(\mathbf{A}_\mu) = -\frac{1}{e^2} \int d^d x \text{tr} \left(\frac{1}{4} \sum_{\mu,\nu} \mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2} m^2 \sum_\mu \mathbf{A}_\mu^2(x) \right), \quad (A22.4)$$

with a flat field integration measure. In contrast with the massless gauge-invariant situation, this field integral is defined.

Massless (gauge-invariant) and massive theory. In the massless case, one can follow the same strategy. However, the \mathbf{A}_0 field equation (A22.3), no longer determines \mathbf{A}_0 but, instead, yields Gauss's law. One may then wonder why one does not impose Gauss's law by integrating over \mathbf{A}_0 as illustrated previously. The reason is that, due to space gauge invariance, Gauss's law commutes with the Hamiltonian. Therefore, if it is satisfied at initial time it is satisfied at all later times. An integration over $\mathbf{A}_0(t, x)$ is thus infinitely redundant, as one immediately verifies, since this procedure leads to the initial action and thus to an undefined field integral.

The massless limit. Provided one considers only gauge-invariant correlation functions, one can introduce a gauge condition and follow all the algebraic steps of Section 22.4. The only modification, induced by the non-gauge invariance of the mass term, is that the field \mathbf{g} , associated with the gauge transformations, remains coupled. It is easy to verify that the resulting action is the sum of terms, due to the gauge fixing procedure and a gauge-invariant part, obtained from the action (A22.4) by the substitution $\mathbf{A}_\mu \mapsto \mathbf{A}_\mu^{\mathbf{g}}$ (equation (22.6)):

$$\mathbf{A}_\mu^{\mathbf{g}}(x) = \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x) = \mathbf{g}(x)\mathbf{D}_\mu\mathbf{g}^{-1}(x), \quad (A22.5)$$

with

$$\mathbf{D}_\mu = \partial_\mu + \mathbf{A}_\mu(x). \quad (A22.6)$$

One finds the action,

$$\mathcal{S}(\mathbf{A}, \mathbf{g}) = \frac{1}{e^2} \int d^d x \text{tr} \left[-\frac{1}{4} \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2} m^2 \sum_\mu \mathbf{D}_\mu \mathbf{g}^{-1}(x) \mathbf{D}_\mu \mathbf{g}(x) \right]. \quad (A22.7)$$

One recognizes a $G \times G/G$ non-linear σ -model (see Section 29.4), in which one of the G components of the symmetry group has been gauged. In contrast with the Abelian example, the scalar field $\mathbf{g}(x)$ is self-coupled and coupled to the gauge field.

The non-linear σ -model action is not renormalizable for dimensions $d > 2$ (see Chapter 19). If we assume that the theory has a cut-off Λ , the \mathbf{g} -field fluctuations will be only weakly damped, because $m \ll \Lambda$; perturbation theory is not particularly reliable. Moreover, the zero gauge field mass limit appears as a strong coupling limit and, therefore, the scalar field $\mathbf{g}(x)$ is not expected to decouple (in perturbation theory IR divergences are generated).

We also note that the complete action can be considered as a limit of a Higgs model action in a limit in which the potential of the Higgs field is sent to infinity. For example, denoting by ϕ a matrix scalar field and assuming that $G \equiv U(N)$, we can view the action (A22.7) as the formal limit, when λ goes to zero, of

$$\begin{aligned} \mathcal{S}(\mathbf{A}, \phi) = & \int d^d x \text{tr} \left[\frac{1}{2} \sum_\mu \mathbf{D}_\mu \phi^\dagger(x) \mathbf{D}_\mu \phi(x) + \frac{1}{\lambda} \left(\phi^\dagger(x) \phi(x) - \frac{m^2}{e^2} \right)^2 \right. \\ & \left. - \frac{1}{4e^2} \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}^2(x) \right]. \end{aligned} \quad (A22.8)$$

Recalling the equivalence between the non-linear σ model and the $(\phi^2)^2$ QFT (see Chapter 19), one may speculate that, beyond perturbation theory, the actions (A22.7) and (A22.8) lead to the same renormalized correlation functions.

23 The Standard Model (SM) of fundamental interactions

In Chapter 22, we have discussed the structure and the formal properties of non-Abelian gauge theories. We have now the necessary technical tools to describe the (2020 updated) *SM* of weak, electromagnetic, and strong interactions [250], which are all known fundamental interactions at the microscopic scale, except gravity and, perhaps, interactions with dark matter.

The weak and electromagnetic theory with a spontaneously broken $SU(2) \times U(1)$ gauge symmetry and three quark and lepton generations [251], thanks to the smallness of the coupling constant at low energy, has now been tested quite systematically, in particular, in e^+e^- colliders, and when radiative corrections are taken into account, provides a precise description of all collider experiments. With the discovery in 2012 of the Higgs particle at the Large Hadron Collider (LHC) at the European Council for Nuclear Research (CERN) with a mass of 125 GeV, all particles of the SM have now been identified, and most parameters have been measured. Still, *the Higgs particle remains the most mysterious particle of the SM*, since it is involved in all the parameters of the SM except gauge couplings, and since it leads to the important fine-tuning problem. The discovery of its origin, and the precise study of its properties should remain, in the future, one of the most important fields of research in particle physics.

Observations have demonstrated the existence of oscillations between the three neutrino species ν_e , ν_μ , and ν_τ . This implies neutrino masses and has led to a SM modification. We assume here the simplest extension, *Dirac neutrinos* (although, lepton number violating, Majorana-type mass terms are not excluded). With such a minimal modification, consistent so far (2020) with experimental data, the lepton and quark sectors have analogous structures: the lepton sector involves a mixing matrix, like the quark sector (three angles have been measured, the fourth CP violating angle is still unknown).

The strong interaction sector of the SM model is described by quantum chromodynamics (QCD), a theory of quarks interacting through *gluons* associated to the gauge fields corresponding to a $SU(3)$ colour gauge group. The QCD renormalization group (RG) β -function, calculated at leading order in perturbation theory, exhibits the property of large momentum *asymptotic freedom* (AF) [252]. Thus, a number of high-energy properties, like the so-called inclusive cross-sections, can be predicted (see also Chapter 24). However, low-energy properties, such as $SU(3)$ colour confinement, cannot be derived from perturbation theory, the effective interaction being too large. Theoretical evidence for the validity of the quark confinement scheme mainly relies on increasingly precise, non-perturbative, numerical investigations of lattice gauge theories (Chapter 25).

Although the SM depends on many parameters, its structure is tightly constrained and satisfies a number of non-trivial consistency conditions. For example, in some cases, when gauge fields are coupled to axial currents, the Ward–Takahashi (WT) identities that are necessary to prove the consistency of gauge theories are not satisfied beyond the tree approximation. They are spoiled by *anomalies*. Therefore, the second part of this chapter is devoted to the discussion of this physically important problem. General results are illustrated by some physical consequences, like the cancellation of anomalies in the SM, the π_0 decay, and the solution to the $U(1)$ problem.

A number of original articles devoted to non-Abelian gauge theories is collected in Ref. [245]. An elementary textbook is, for example, Ref. [253].

23.1 Weak and electromagnetic interactions: Gauge and scalar fields

We first briefly describe the SM sector of weak and electromagnetic interactions (suitably modified to accommodate assumed *Dirac neutrinos*), which provides a physics application of the non-Abelian Higgs mechanism. We mainly discuss the model with one *generation* and two *flavours*, eventually indicating how it generalizes to three generations (with essential new features like the Cabibbo–Kobayashi–Maskawa (CKM) and Pontecorvo–Maki–Nakagawa–Sakata (PMNS) mixing matrices, and *CP violating angles*). For phenomenological applications, the reader is referred to the literature.

The gauge group of the weak–electromagnetic sector is $SU(2) \times U(1)$. The form of the gauge action with one scalar field can be obtained from the action considered in Section 22.6.3, by gauging a $U(1)$ subgroup of the remaining $SU(2)$ non-gauge symmetry. The pattern of symmetry breaking is similar, but here an unbroken $U(1)$ gauge symmetry, associated with electromagnetic interactions, remains, whose generator is a linear combination of the original $U(1)$ generator and one of the $SU(2)$ generators. Since the gauge group is a product of two groups, the model depends on two independent gauge couplings and, therefore, weak and electromagnetic interactions are combined rather than completely unified.

Note that, in the SM, since the left-handed and right-handed components of the fermion fields are treated differently, the breaking of parity symmetry is explicit.

Notation. In this chapter, the convention of implicit summation over repeated indices is everywhere assumed.

We use the conventions of Section 22.6.3, in particular, for the $SU(2)$ transformations and the τ Pauli matrices.

23.1.1 Gauge and scalar fields: The action

In Section 22.6.3, we have represented the field ϕ as a 2×2 complex matrix of the form (22.72). For what follows, it is more convenient to parametrize the scalar field ϕ in terms of a complex $SU(2)$ doublet φ . The gauge-invariant scalar field action then reads

$$\mathcal{S}_{\text{scal.}} = \int d^d x [\overline{\mathbf{D}_\mu \varphi}(x) \cdot \mathbf{D}_\mu \varphi(x) + \mathcal{U}(|\varphi(x)|^2)], \quad (23.1)$$

with (g and g' are the two independent gauge coupling constants)

$$\mathbf{D}_\mu = \mathbf{1} \left(\partial_\mu + \frac{1}{2} i Y_\varphi g' B_\mu \right) + \frac{1}{2} i g \mathbf{A}_\mu \cdot \boldsymbol{\tau},$$

where the real vector B_μ is associated with the $U(1)$ group and \mathbf{A}_μ , a three-component real vector, is associated with the $SU(2)$ group. The constant Y_φ is the hypercharge of φ with respect to the $U(1)$ group, and we normalize g' in such a way that $Y_\varphi = 1$.

We parametrize the φ potential in terms of two constants r and λ as

$$\mathcal{U}(\rho) = r\rho + \frac{1}{6}\lambda\rho^2, \quad \lambda > 0. \quad (23.2)$$

We normalize the curvature tensors by

$$[\mathbf{D}_\mu, \mathbf{D}_\nu] = \frac{1}{2}i [g \mathbf{F}_{\mu\nu} \cdot \boldsymbol{\tau} + g' B_{\mu\nu} \mathbf{1}],$$

in which $\mathbf{F}_{\mu\nu}$ is the curvature corresponding to the $SU(2)$ subgroup, and $B_{\mu\nu}$ the curvature corresponding to the $U(1)$ subgroup.

In the $SO(3)$ notation,

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu - g \mathbf{A}_\mu \times \mathbf{A}_\nu, \quad \text{and} \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (23.3)$$

With this choice of normalizations, the gauge field action is the simple sum

$$\mathcal{S}(\mathbf{A}, B) = \frac{1}{4} \int d^4x [\mathbf{F}_{\mu\nu}^2(x) + B_{\mu\nu}^2(x)]. \quad (23.4)$$

23.1.2 The Higgs mechanism: Classical approximation

We now assume that, in the potential (23.2), the coefficient r is negative. In the classical approximation, the field φ then has a non-zero expectation value that can be chosen of the form

$$\langle \varphi(x) \rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v > 0. \quad (23.5)$$

With this form, the expectation value is invariant under the $U(1)$ transformations,

$$\varphi(x) \mapsto e^{i\theta(1+\tau_3)/2} \varphi(x). \quad (23.6)$$

This $U(1)$ group is a subgroup of $SU(2) \times U(1)$, whose generator is the sum of the generator of the initial $U(1)$ group and the τ_3 generator of $SU(2)$. As anticipated, the group $SU(2) \times U(1)$ is broken down to a $U(1)$ subgroup.

Replacing φ by its expectation value in action (23.1), one reads off the mass terms for the gauge fields in the classical approximation. One finds

$$\frac{1}{8}v^2 \int d^4x \left[(g'B_\mu - gA_\mu^{(3)})^2 + g^2 |A_\mu^{(1)} + iA_\mu^{(2)}|^2 \right]. \quad (23.7)$$

First, it follows that the linear combination $g'B_\mu - gA_\mu^{(3)}$ is massive while the orthogonal combination remains massless, and thus represents the electromagnetic field.

One defines the weak angle θ_W by

$$g'/g = \tan \theta_W. \quad (23.8)$$

The electromagnetic field A_μ (associated to the photon) and the massive neutral field Z_μ then correspond to

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W A_\mu^{(3)}, \quad Z_\mu = -\sin \theta_W B_\mu + \cos \theta_W A_\mu^{(3)}. \quad (23.9)$$

Conversely,

$$B_\mu = \cos \theta_W A_\mu - \sin \theta_W Z_\mu, \quad A_\mu^{(3)} = \sin \theta_W A_\mu + \cos \theta_W Z_\mu. \quad (23.10)$$

The components $A_\mu^{(1,2)}$ are coupled to A_μ and correspond to charged vector fields, which are usually written in complex notation as

$$W_\mu^\pm = (A_\mu^{(1)} \pm iA_\mu^{(2)})/\sqrt{2}. \quad (23.11)$$

From the coupling of the charged vector bosons with the photon, one derives the relation between electric charge e and coupling constants g and g' ,

$$e = gg' / \sqrt{g^2 + g'^2} = g \sin \theta_W = g' \cos \theta_W. \quad (23.12)$$

The Z mass is then

$$m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} = \frac{gv}{2\cos \theta_W} = \frac{ev}{\sin 2\theta_W}. \quad (23.13)$$

The two charged vector bosons W_μ^\pm have the common mass,

$$m_W = \frac{1}{2}gv = \frac{ev}{2\sin \theta_W} = m_Z \cos \theta_W. \quad (23.14)$$

The mass of the remaining massive scalar particle, the Higgs boson, is (equation (22.79))

$$m_H = v\sqrt{\lambda/3}. \quad (23.15)$$

The experimental values (2019) are

$$m_W = 80.379 \pm 0.012 \text{ GeV}, \quad m_Z = 91.1876 \pm 0.0021 \text{ GeV}, \quad m_H = 125.10 \pm 0.14 \text{ GeV}.$$

The direct measurements of the W and Z masses, for example, determine the parameter θ_W . At leading order, the weak angle is given by $\sin^2 \theta_W = 0.223$, instead of the experimental value $\sin^2 \theta_W = 0.231$. When radiative corrections are taken into account, the agreement between theory and different experimental determinations, like the one coming from neutral currents, improves. One also infers $v = 246.2$ GeV (equation (23.28)).

Inserting the mass of the Higgs particle, one finds $\lambda \approx 0.75$, which is rather small.

23.2 Leptons: Minimal SM extension with Dirac neutrinos

We consider here only the electron e^- and the corresponding neutrino ν_e , since the couplings of the two other generations (μ , ν_μ and τ , ν_τ) have an analogous structure. In the SM, the neutrino and the left-handed part of the electron are combined into a left-handed doublet of $SU(2)$:

$$L = \begin{pmatrix} (\nu_e)_L \\ e_L \end{pmatrix} \equiv \frac{1}{2}(1 - \gamma_5) \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}. \quad (23.16)$$

The right-handed parts of the neutrino $R_1 = \frac{1}{2}(1 + \gamma_5)\nu_e$ (an addition to the SM) and of the electron $R_2 = \frac{1}{2}(1 + \gamma_5)e^-$ are $SU(2)$ singlets.

23.2.1 Lepton gauge action

The action for the two leptons coupled to gauge fields can be written as

$$\mathcal{S}_{\nu,e} = - \int d^4x \left[\sum_i \bar{R}_i (\partial^\mu + \frac{1}{2}iY_{i,R}g'\mathcal{B}) R_i + \bar{L} (\partial^\mu + \frac{1}{2}iY_Lg'\mathcal{B} + \frac{1}{2}ig\mathbf{A} \cdot \boldsymbol{\tau}) L \right]. \quad (23.17)$$

Using relations (23.10) and (23.12) to replace B_μ and $A_\mu^{(3)}$ by their components on A_μ , one obtains the coupling terms to the electromagnetic field,

$$\frac{1}{2}ie \sum_i Y_{i,R} \bar{R}_i \mathcal{A} R_i + \frac{1}{2}ie \bar{L} (Y_L + \tau_3) \mathcal{A} L.$$

Denoting by T_3 the eigenvalue of τ_3 , one of the generators of $SU(2)$, Y the $U(1)$ hypercharge and Q the electric charge, one derives from the relations (23.9) and (23.12) for each fermion,

$$Q = \frac{1}{2}(T_3 + Y) \quad (23.18)$$

or, explicitly,

$$Q_1 = \frac{1}{2}Y_{1R}, \quad Q_2 = \frac{1}{2}Y_{2R}, \quad Q_1 = \frac{1}{2}(Y_L + 1), \quad Q_2 = \frac{1}{2}(Y_L - 1). \quad (23.19)$$

First, one notes the consistency condition

$$Q_1 - Q_2 = 1, \quad (23.20)$$

a direct consequence of the $SU(2)$ doublet assignment. Then, all $U(1)$ charges are related to Q_1 by

$$Y_{1R} = 2Q_1, \quad Y_{2R} = 2(Q_1 - 1), \quad Y_L = 2Q_1 - 1. \quad (23.21)$$

Identifying Q_1 with the neutrino charge, $Q_1 = 0$, one finds $Q_2 = -1$ (consistent with the electron charge) and the $U(1)$ hypercharges,

$$Y_{1R} = 0, \quad Y_{2R} = -2, \quad Y_L = -1. \quad (23.22)$$

The value $Y_{1R} = 0$ is consistent with the observation that ν_R does not participate in the weak and electromagnetic interactions.

23.2.2 Lepton masses

The $SU(2)$ symmetry forbids fermion mass terms. On the other hand, the coupling terms between the scalar field and the leptons,

$$\mathcal{S}_1 = iG_\nu \int d^4x [\bar{R}_1(x)(\varphi(x)\tau_2 L(x)) + (\bar{L}(x)\tau_2 \varphi^*(x))R_1(x)], \quad (23.23a)$$

$$\mathcal{S}_2 = G_e \int d^4x [\bar{R}_2(x)(\varphi^*(x) \cdot L(x)) + (\bar{L}(x) \cdot \varphi(x))R_2(x)], \quad (23.23b)$$

are allowed by the $SU(2) \times U(1)$ symmetry (if \mathbf{U} belongs to $SU(2)$, $\mathbf{U}^* = \tau_2 \mathbf{U} \tau_2$), independently of the value of Q_1 provided $Y_\varphi = 1$. Replacing the field φ by its expectation value, we note that the spontaneous breaking of the $SU(2) \times U(1)$ symmetry generates the masses,

$$m_\nu = v|G_\nu| \quad m_e = v|G_e|.$$

Therefore, the fermion masses are calculable, but, in the absence of a new dynamic principle, in terms of two arbitrary parameters, G_e and G_ν .

In the SM, the Yukawa couplings of leptons are proportional to their mass. In particular, the perturbative approximation becomes worse for heavier leptons.

23.2.3 The Fermi constant

The coupling constant G_F (the Fermi constant), characteristic of the strength of weak interactions, is defined in terms of the effective low-energy current-current, and thus four-fermion interaction:

$$\frac{G_F}{\sqrt{2}} \int d^4x J_\mu(x) J_\mu^\dagger(x). \quad (23.24)$$

The contribution of the electron and the neutrino to the charged current J_μ has the form:

$$J_\mu(x) = \bar{e}(x)(1 - \gamma_5)\gamma_\mu \nu_L(x) = 2\bar{e}_L(x)\gamma_\mu \nu_L(x). \quad (23.25)$$

The relation between G_F and the coupling constants g and g' is obtained by taking the large W -mass limit of the electron-neutrino scattering amplitude in the tree approximation. The result can be obtained by integrating over the vector fields $A_\mu^{(1,2)}$, taking only into account the mass term, and neglecting the kinetic part.

The corresponding part of the action is

$$\frac{1}{8}g^2 v^2 \left[\left(A_\mu^{(1)} \right)^2 + \left(A_\mu^{(2)} \right)^2 \right] + \frac{ig}{2} \bar{L} \gamma_\mu \left(A_\mu^{(1)} \tau_1 + A_\mu^{(2)} \tau_2 \right) L. \quad (23.26)$$

Completing squares, one immediately obtains the result of the integration,

$$\frac{1}{2v^2} \left[(\bar{L} \gamma_\mu \tau_1 L)^2 + (\bar{L} \gamma_\mu \tau_2 L)^2 \right] = \frac{2}{v^2} \bar{\nu}_L \gamma_\mu e_L \bar{e}_L \gamma_\mu \nu_L. \quad (23.27)$$

Comparing with the definition (23.24), one concludes,

$$G_F/\sqrt{2} = 1/2v^2. \quad (23.28)$$

Inserting the experimental value $G_F = 1.16637\dots \times 10^{-5} \text{ GeV}^{-2}$ one finds the value $v = 246.22 \text{ GeV}$.

The effective low-energy Fermi model of charged weak interactions by itself determines all parameters of the SM (with massless neutrinos) but two, for example, the weak angle θ_W and the Higgs field self-coupling λ , which have to be inferred from additional experimental information. At leading order, the W mass can be rewritten as

$$M_W^2 = \frac{e^2}{4\sqrt{2}G_F \sin^2 \theta_W},$$

that is,

$$M_W = \left(\frac{\pi\alpha}{\sqrt{2}G_F} \right)^{1/2} \frac{1}{\sin \theta_W} = \frac{37.28}{\sin \theta_W} \text{ GeV}, \quad (23.29)$$

which yields the slightly worse value: $\sin^2 \theta_W = 0.215$.

The coupling of the charged vector bosons W^\pm to $e^- \nu_e$ is obtained by introducing the definition (23.11) into equation (23.17):

$$\begin{aligned} \frac{g}{2} \bar{L} \gamma_\mu \left(\tau^1 A_\mu^{(1)} + \tau^2 A_\mu^{(2)} \right) L &= \frac{g}{2\sqrt{2}} \left[(\bar{\nu}_L \gamma_\mu e_L + \bar{e}_L \gamma_\mu \nu_L) (W_\mu^+ + W_\mu^-) \right. \\ &\quad \left. + (\bar{\nu}_L \gamma_\mu e_L - \bar{e}_L \gamma_\mu \nu_L) (W_\mu^+ - W_\mu^-) \right] \\ &= \frac{g}{\sqrt{2}} (\bar{\nu}_L \gamma_\mu e_L W_\mu^+ + \bar{e}_L \gamma_\mu \nu_L W_\mu^-). \end{aligned} \quad (23.30)$$

Using the definitions (23.8) and (23.9), one also derives the couplings of leptons to the neutral vector fields A_μ and Z_μ ,

$$\begin{aligned} \frac{eZ_\mu}{\sin 2\theta_W} &(2 \sin^2 \theta_W \bar{e}_R \gamma_\mu e_R - \cos 2\theta_W \bar{e}_L \gamma_\mu e_L + \bar{\nu}_L \gamma_\mu \nu_L) \\ &- e A_\mu (\bar{e}_R \gamma_\mu e_R + \bar{e}_L \gamma_\mu e_L). \end{aligned} \quad (23.31)$$

Note that, from the study of the Z decays, one infers that the number of generations with ‘light’ (*i.e.* with a mass less than $m_Z/2 \approx 45$ GeV), active (*i.e.* weakly interacting) neutrinos is exactly three.

The SM: Neutrino parameters with three generations. The experimental discovery of neutrino oscillations between the three flavours, ν_e , ν_μ , and ν_τ , implies that neutrinos have masses.

This has necessitated an update of the SM. One infers (2019),

$$m^2(\nu_\mu) - m^2(\nu_e) \approx 7.5 \times 10^{-5} \text{ eV}^2, \quad |m^2(\nu_\tau) - m^2(\nu_\mu)| \approx 2.5 \times 10^{-3} \text{ eV}^2.$$

Including cosmological observations, one finds a bound on the sum of masses of about 0.2 eV.

Like in the quark sector, the mass eigenstates differ from the linear combinations of neutrinos appearing in the weak interactions. If neutrinos are Dirac fermions, as we have assumed, this implies a mixing PMNS matrix, as described in Section 23.3, with three measured CP-conserving angles,

$$\sin^2(2\theta_{12}) = 0.846 \pm 0.021, \quad \sin^2(2\theta_{23}) = 1 + 0 - 0.017, \quad \sin^2(2\theta_{13}) = 9.3 \pm 0.8 \times 10^{-2},$$

and one unknown CP violating angle.

A major issue is the extremely small values of the neutrino masses. This problem could have a solution in the framework of so-called grand unified theories, in a scheme with Majorana mass terms for neutrinos. This would imply some form of lepton number violation (see Section 12.3.3).

23.3 Quarks and weak–electromagnetic interactions

With the addition of the right-handed neutrino, the structures of the lepton and quark sectors become quite similar.

We first consider again only one generation with two flavours, corresponding to the quarks **u** and **d**, the coupling of other generations being analogous. Each quark has a colour quantum number and forms a $SU(3)$ triplet (see next section). The left components of the quarks belong to a $SU(2)$ doublet,

$$\mathbf{Q}_L = \{\mathbf{u}_L, \mathbf{d}_L\}. \quad (23.32)$$

All right-handed components $\mathbf{Q}_{1,2R}$ form $SU(2)$ singlets. A gauge-invariant quark action has the form

$$\begin{aligned} \mathcal{S}_Q = - \int d^4x & \left[\sum_{i=1}^2 \bar{\mathbf{Q}}_{iR}(x) (\not{D} + \frac{1}{2}ig'Y_{iR}\not{B}(x)) \mathbf{Q}_{iR}(x) \right. \\ & \left. + \bar{\mathbf{Q}}_L(x) (\not{D} + \frac{1}{2}ig'Y_L\not{B}(x) + \frac{1}{2}ig\not{A}(x) \cdot \boldsymbol{\tau}) \mathbf{Q}_L(x) \right]. \end{aligned} \quad (23.33)$$

We now use the relations (23.20) and (23.21) for the charges Q_i and hypercharges Y_i ,

$$Q_1 - Q_2 = 1, \quad Y_{1R} = 2Q_1, \quad Y_{2R} = 2(Q_1 - 1), \quad Y_L = 2Q_1 - 1.$$

The proton is a **uud** state and has charge +1. Thus, $2Q_1 + Q_2 = 3Q_1 - 1 = 1$ and the quark **u** has charge 2/3. Then, the quark **d** has charge $Q_2 = Q_1 - 1 = -1/3$. The neutron is a **udd** state, and one verifies that it has indeed charge 0, a result consistent with the $SU(2)$ doublet assignment. Finally,

$$Y_L = 1/3, \quad Y_{1R} = 4/3, \quad Y_{2R} = -2/3. \quad (23.34)$$

In Section 23.7.2, we verify that the $SU(3)$ triplet structure of quarks leads to the cancellation of the possible anomaly due to the chiral coupling of gauge fields to fermions in each generation and, therefore, ensures the consistency of the gauge theory of weak and electromagnetic interactions.

Couplings to Higgs field and quark masses. As for leptons, direct quark mass terms are forbidden by the $SU(2)$ symmetry. The quark masses are produced by the coupling to the Higgs scalar field and the spontaneous symmetry breaking. The $SU(2) \times U(1)$ -invariant Higgs–quark coupling terms can be easily inferred from the expressions (23.23),

$$\begin{aligned} \mathcal{S}_{Q\text{mass}} = \int d^4x & \left\{ iG_{q1} [\bar{\mathbf{Q}}_{1R}(x)(\varphi(x)\tau_2\mathbf{Q}_L(x)) + (\bar{\mathbf{Q}}_L(x)\tau_2\varphi^*(x))\mathbf{Q}_{1R}(x)] \right. \\ & \left. + G_{q2} [\bar{\mathbf{Q}}_{2R}(x)(\varphi^*(x) \cdot \mathbf{Q}_L(x)) + (\bar{\mathbf{Q}}_L(x) \cdot \varphi(x))\mathbf{Q}_{2R}(x)] \right\}, \end{aligned} \quad (23.35)$$

which can provide masses for the two quarks. This is at least the situation for one generation. However, six quarks, belonging to three generations, have been observed (see Table 23.1). Therefore, in the interactions (23.35), the spinors that appear on the right and the left need not be, and are not, the same. When one replaces the scalar field φ by its expectation value, one obtains in general a non-diagonal mass matrix of the form

$$\bar{\mathbf{Q}}_{1R}^\alpha M_{\alpha\beta} \mathbf{Q}_{1L}^\beta + \bar{\mathbf{Q}}_{1L}^\alpha M_{\alpha\beta}^\dagger \mathbf{Q}_{1R}^\beta, \quad (23.36)$$

for the quarks of charge 2/3, and a similar one for the charge $-1/3$ quarks.

Performing independent unitary transformations $\mathbf{U}_{R,L}$ on the right and left quark components, it is possible to replace the matrix \mathbf{M} by a real diagonal matrix \mathcal{M} :

$$\mathbf{U}_R^\dagger \mathbf{M} \mathbf{U}_L = \mathcal{M}. \quad (23.37)$$

In this representation, the quarks are mass eigenstates. However, the weak interactions no longer have the simple form (23.33), because the unitary transformations on the quark components \mathbf{Q}_{1L} and \mathbf{Q}_{2L} are in general different. It is customary to put the blame onto the charge $-1/3$ quarks. The mismatch is expressed in terms of a 3×3 unitary matrix (in the case of three generations), the PMNS matrix, which relates the quark mass eigenstates \mathbf{d} , \mathbf{s} , and \mathbf{b} to the quarks appearing in the weak interactions:

$$[\mathbf{Q}_2^\alpha]_{\text{weak int.}} = U_{\alpha\beta} [\mathbf{Q}_2^\beta]_{\text{mass eigenst.}}. \quad (23.38)$$

With only two generations (\mathbf{d} , and \mathbf{s}), it was possible to cast the CKM matrix into the form

$$\mathbf{U}_C = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix},$$

in which θ_C is the Cabibbo angle, after unobservable changes of the relative phases between the quarks.

In the presence of the third \mathbf{b} quark, the 3×3 PMNS matrix can be parametrized in terms of three rotation angles and one CP violating phase, responsible for the observed direct CP violation in neutral kaon and B_0 meson decay. However, this CP violation seems to be much too small to explain the disappearance of antimatter.

Table 23.1
Quarks and leptons. The three generations (2019).

Charge 2/3 quarks	Charge $-1/3$ quarks	Charge -1 leptons	Neutrinos
\mathbf{u} , $m = 1.9$ to 2.7 MeV	\mathbf{d} , $m = 4.5$ to 5.2 MeV	e , $m = 0.511$ MeV	see Section 23.2
\mathbf{c} , $m = 1.27 \pm 0.02$ GeV	\mathbf{s} , $m = 88$ to 104 MeV	μ , $m = 105.6$ MeV	see Section 23.2
\mathbf{t} , $m = 172.9 \pm 0.4$ GeV	\mathbf{b} , $m = 4.18 \pm 0.03$ GeV	τ , $m = 1.777$ GeV	see Section 23.2

Beyond the SM: Grand unification. Theoretical speculations based on the search for a unifying simple group including $U(1) \times SU(2) \times SU(3)$ as a subgroup, have mainly focused on $SU(5)$ (the larger $SO(10)$ has also been discussed). This group deals nicely with fermions, has 12 additional super-heavy gauge bosons, but necessitates a large collection of additional Higgs fields. Running the three independent couplings of the SM model to higher energies, one observes an apparent unification at energies of order 10^{15} GeV. The non-observation of the proton decay, predicted in the model, has shifted the focus to the minimal supersymmetric extension of $SU(5)$, where the problem with proton decay is less severe and the apparent unification of the running coupling constants more precise. However, no supersymmetric particle has been found yet at the LHC (2019). Moreover, the extrapolation up to 10^{15} GeV, which assumes a form of desert beyond the SM, is quite questionable. In particular, the problem of the fine tuning of the bare Higgs mass arises (Section 8.8.3), as we will now explain.

23.3.1 Elementary scalar fields: Parameter proliferation and fine tuning

The SM has (Higgs) scalar fields as an essential ingredient. This is the source of many difficulties. In the absence of the Higgs field, all gauge bosons and fermions would be massless and this property would be protected by gauge and chiral symmetries. The SM would depend only three gauge couplings and, perhaps, the CP violating θ angle of QCD. The Higgs field is at the origin of a large number of new arbitrary parameters, the masses of all fundamental particles, the CKM and PMNS matrices, and the Higgs self-coupling. This could perhaps be expected from an effective low-energy theory. However, the diversity of these parameters is rather puzzling. For example, if the couplings Higgs–fermions were ‘natural’ (of order unity), all fermion masses would be in the few 100 GeV range, as the Higgs expectation value and the W, and Z masses. In this sense, only the top quark (**t**) mass (173 GeV), is natural. By contrast, in the quark sector the masses span about five orders of magnitude, and in the lepton sector the range is even larger.

Another, perhaps more fundamental problem, is related to the Higgs mass. Generically, bare and physical masses of scalar particles are expected to be of the order of the momentum ‘cut-off’, which reflects the scale of some new physics. The scalar bare mass has to be *tuned* to render the physical mass much smaller than the cut-off. In the statistical physics interpretation of the ϕ^4 QFT, the divergence of the correlation length (the inverse of the particle physical mass) is obtained by tuning the temperature, represented by the bare mass, close to a critical value where a second-order phase transition occurs.

However, in particle physics, all parameters are given, and it is somewhat unnatural for the scalar bare mass to lie accidentally close to such a critical value. This leads to the *fine-tuning problem*. To get a rough idea about the severity of the problem, we expect the coefficient of ϕ^2 to be of order Λ^2 , where Λ is the scale of some new physics. The fine-tuning factor is thus of order $(m_H/\Lambda)^2$. Even, for some unknown reason, the initial contribution is smaller than expected, the next perturbative contribution generated by the ϕ^4 self-coupling λ for large momentum cut-off Λ is

$$\delta m_0^2 = \frac{\lambda}{16\pi^4} \int \frac{d^4 p}{p^2(1+p^2/\Lambda^2)^2} = \frac{\lambda}{16\pi^2} \Lambda^2.$$

The Higgs mass m_H at leading order is given by $m_H^2 = \frac{1}{3}\lambda v^2$. Therefore,

$$f \equiv \delta m_0^2/m_H^2 = \frac{3}{16\pi^2} (\Lambda/v)^2.$$

Another important contribution comes a top quark loop, which is of order $(\Lambda/v)^2/\pi^2$. Considering the unexplained range of fermion masses, it is difficult to decide how much fine tuning is acceptable. If one demands that the correction should be smaller than the leading term, then one obtains $\Lambda < 700$ GeV, which is essentially excluded by the LHC. If no new physics beyond the SM is found at the LHC, depending on assumptions, the fine tuning factor will lie between 10 and a few 100. On the other hand, an absence of any new physics below Planck’s scale (10^{19} GeV), or the proposed unification scale (about 10^{15} GeV), is extremely difficult to believe.

Proposed solutions to the fine-tuning problem. At some scale, the fine-tuning problem must be solved with new physics. Three types of schemes have been proposed so far:

(i) The Higgs boson is a bound state of a new type of fermions. This requires a specific model, hopefully not involving new scalars again. Models in which the forces are again due to gauge interactions have been proposed and fall under the generic name of *technicolour* (see next section). Such models have problems generating fermion masses and no compelling model has been proposed yet.

(ii) The Higgs boson remains associated with a fundamental field, but the mass renormalization problem is solved with the help of *supersymmetry*. Since fermions, due to chiral symmetry, can be naturally massless, the idea is to use supersymmetry (see for example Section 27.3), to relate them to scalars. In such models, the scalar mass renormalization grows only logarithmically with the cut-off (it would be absent in the absence of supersymmetry breaking), and thus the problem is much less severe even if the cut-off is of the order of the Planck mass. The difficulty with this approach is that none of the superpartners of existing particles predicted by supersymmetry have been found yet (2020). Therefore, even if signs of supersymmetry are found at some higher scale, the cancellation can only be partial at best.

(iii) The Higgs field is related to additional dimensions of space, for instance, is a component of a gauge vector field, but then, in the simplest realizations, it must transform under the adjoint representation, and this complicates models.

23.4 QCD. RG equations and β function

We now focus on QCD, a gauge theory of quarks and vector particles called *gluons*, which produces the observed hadrons and their strong interactions, neglecting the weak and electromagnetic interactions, which we have described in the preceding sections.

23.4.1 QCD

QCD describes the interactions between quarks, characterized by a *flavour quantum number* f relevant for weak interactions, which are triplets for transformations of a gauged group, the *SU(3) colour group*, realized in the symmetric phase. Their interactions are mediated by the corresponding gauge fields (gluon fields). The action takes the form

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{Q}}, \mathbf{Q}) = - \int d^4x \left[\frac{1}{4g^2} \text{tr} \mathbf{F}_{\mu\nu}^2(x) + \sum_f \bar{\mathbf{Q}}_f(x) (\mathbf{D} + m_f) \mathbf{Q}_f(x) \right], \quad (23.39)$$

where g is the gauge coupling constant, and m_f denotes the quark masses. The *SU(3)* coupling constant is generally parametrized as $\alpha_s = g^2/4\pi$. The experimental value of the running coupling constant α_s , determined at the Z vector boson mass, is $\alpha_s(m_Z) = 0.1203(3)$ (2020).

In what follows, since no confusion with quantum electrodynamics (QED) is possible, we denote α_s simply as α .

The most important physical arguments in favour of such a model are:

(i) Quarks behave almost like free particles at short distances, as indicated by deep inelastic scattering experiments, or the spectrum of bound states of heavy quarks. We calculate below the RG β -function, and show that a pure non-Abelian gauge theory is asymptotically free at large momentum in four dimensions (like the non-linear σ -model in two dimensions).

In Chapter 24, we prove that this property survives the inclusion of a limited number of fermions and, furthermore, that this property is specific to non-Abelian gauge theories.

(ii) No free quarks and gluons have ever been observed at large distance (but they manifest themselves indirectly in the form of jets). This is consistent with the simplest picture in which the β -function (which, due to AF, is negative at small coupling) remains negative for all couplings in such a way that the effective coupling constant grows without bounds at large distances. Numerical simulations strongly support this conjecture, called the colour *confinement* hypothesis (see Chapter 25).

23.4.2 Semi-classical vacuum and fine-tuning problem

In Section 39.6, we show, in the temporal gauge, that the action has an infinite number of degenerate minima (related to gauge transformations homotopy classes) leading to degenerate semi-classical ground states. However, these minima are connected by barrier penetration effects due to instantons (a situation similar to the cosine potential in one-dimensional quantum mechanics (QM)). The true eigenstates are linear combinations, depending on an angle θ , of the states concentrated around minima, leading to the notion of θ -vacuum [255]. To project on a θ -vacuum, one can add to the action the topological quantity (equation (23.102) and for details, see Section 39.6),

$$\frac{i\theta}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \int d^4x \operatorname{tr} \mathbf{F}_{\mu\nu}(x) \mathbf{F}_{\rho\sigma}(x), \quad (23.40)$$

where $\epsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor $\epsilon_{1234} = 1$. For $\theta \neq 0$, this leads to the strong CP violation problem, since the experimental bound on the neutron dipolar electric moment implies $|\theta| < 10^{-10}$. A ‘natural’ solution to this fine-tuning problem has been proposed, based on introducing in the SM a new pseudoscalar particle, dubbed *axion*, leading to the substitution $\theta \mapsto \theta(x)$ [256], where $\theta(x)$ is the axion field. More recently, it has been suggested that such an axion could also contribute to *dark matter*.

23.4.3 RG equations in a covariant gauge

We first discuss the gauge dependence of RG equations, and of RG functions in *matterless gauge theories*, quantized in the covariant gauge used in equation (22.43), in the form of the dependence on the parameter ξ . A short discussion of the Abelian case can be found in Section 21.11. We denote by Z_A the gauge field renormalization constant, and Z_g the renormalization constant of α .

In terms of the renormalization scale μ , the RG equation for the gauge field n -point renormalized vertex function reads:

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(\alpha, \xi) \frac{\partial}{\partial \alpha} + \delta(\alpha, \xi) \xi \frac{\partial}{\partial \xi} - \frac{n}{2} \eta_A(\alpha, \xi) \right] \Gamma_r^{(n)}(\mu, \alpha, \xi) = 0, \quad (23.41)$$

where α and ξ are the renormalized parameters, α_0 and ξ_0 are the bare parameters, and

$$\delta(\alpha, \xi) \equiv \xi^{-1} \left. \mu \frac{\partial}{\partial \mu} \right|_{\alpha_0, \xi_0 \text{ fixed}} \quad \xi = \beta(\alpha, \xi) - \eta_A(\alpha, \xi). \quad (23.42)$$

It is shown in Section 26.10.5 that for such gauges, as in the Abelian case, the gauge-fixing term is not renormalized. Therefore, ξ_0 , the parameter of the bare quantum action (22.43), is related to the renormalized parameter ξ by

$$\xi_0 = \xi Z_A / Z_g. \quad (23.43)$$

Gauge-invariant operators. In Section 26.11, we prove that the bare correlation functions of gauge-invariant operators are gauge independent. Therefore, they are independent of ξ_0 . This also applies to the renormalization constants needed to render these correlation functions finite. It is thus possible to construct renormalized correlation functions which are also ξ_0 independent.

We consider such a correlation function Γ . It satisfies

$$\left. \frac{\partial}{\partial \xi} \right|_{\alpha_0, \text{cut-off fixed}} \Gamma = \left(\frac{\partial}{\partial \xi} + \rho(\alpha, \xi) \frac{\partial}{\partial \alpha} \right) \Gamma(\mu, \alpha, \xi) = 0, \quad (23.44)$$

with

$$\rho(\alpha, \xi) = \left. \frac{\partial \alpha}{\partial \xi} \right|_{\alpha_0, \text{cut-off fixed}}. \quad (23.45)$$

Γ also satisfies an RG equation, which we assume to be homogeneous:

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(\alpha, \xi) \frac{\partial}{\partial \alpha} + \delta(\alpha, \xi) \xi \frac{\partial}{\partial \xi} - \eta_\Gamma(\alpha, \xi) \right] \Gamma(\mu, \alpha, \xi) = 0. \quad (23.46)$$

Using equation (23.44) to eliminate $\partial/\partial\xi$, we obtain a new RG equation for Γ :

$$\left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(\alpha, \xi) \frac{\partial}{\partial \alpha} - \eta_\Gamma(\alpha, \xi) \right] \Gamma(\mu, \alpha, \xi) = 0, \quad (23.47)$$

with

$$\tilde{\beta} = \beta - \xi \delta \rho. \quad (23.48)$$

Expressing then the compatibility condition between the two linear equations (23.44) and (23.47), we obtain two equations:

$$\left(\frac{\partial}{\partial \xi} + \rho(\alpha, \xi) \frac{\partial}{\partial \alpha} \right) \eta_\Gamma(\alpha, \xi) = 0, \quad (23.49)$$

$$\left(\frac{\partial}{\partial \xi} + \rho(\alpha, \xi) \frac{\partial}{\partial \alpha} \right) \tilde{\beta} = \frac{\partial \rho}{\partial \alpha} \tilde{\beta}. \quad (23.50)$$

The first equation expresses that, as expected, the multiplicative renormalization of Γ is independent of ξ_0 . The second equation shows that the zeros of $\tilde{\beta}$ are gauge independent. Moreover, differentiating the equation with respect to α , one also finds that the slope of β at its zeros is gauge independent. Also, one verifies that, in an MS scheme (see Section 10.4), the function ρ vanishes. Indeed, in dimensional regularization, the relation between α_0 and α takes the form

$$\alpha_0 = \mu^\varepsilon \alpha Z_g = \mu^\varepsilon \alpha \left(1 + \frac{Z_g^1(\alpha, \xi)}{\varepsilon} + \frac{Z_g^2(\alpha, \xi)}{\varepsilon^2} + \dots \right). \quad (23.51)$$

The important point is that the term without pole in ε in the expansion of Z_g is ξ independent. Using the definition (23.45) of ρ , one then finds

$$0 = \rho \left(1 + \frac{\partial Z_g^1}{\partial \alpha} \frac{1}{\varepsilon} + \dots \right) + \alpha \left(\frac{\partial Z_g^1}{\partial \xi} \frac{1}{\varepsilon} + O\left(\frac{1}{\varepsilon^2}\right) \right). \quad (23.52)$$

Therefore, the expansion of ρ for ε small has only singular contributions. Since ρ is finite, all singular contributions must cancel and thus ρ vanishes identically. It follows that, in the MS scheme, the β -function and η_Γ are independent of ξ .

23.5 General RG β -function at one-loop order: Asymptotic freedom

We now calculate the RG β -function at one-loop order, in a gauge theory with gauge fields and fermions, corresponding to a simple group G , in particular, to verify AF, because an understanding of its sign without explicit calculation is not simple.

The calculation can be done by various methods, for example, one can use the background field method, as in the case of models on homogeneous spaces in Section 29.6.2. Here, instead, we calculate directly the β -function from the renormalization of the gauge coupling constant, as defined by the fermion-gauge field vertex.

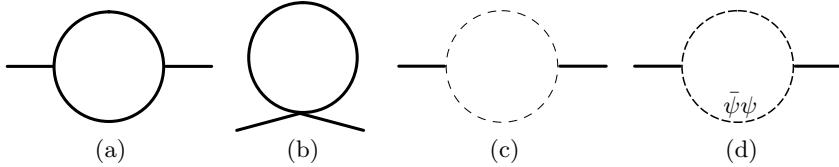


Fig. 23.1 The gauge field two-point function at one-loop (dotted lines represent fermions).

23.5.1 The RG β -function at one-loop order

We need the divergent parts of the gauge field and fermion two-point functions, and the fermion gauge field three-point function. We work in Feynman's gauge and use *dimensional regularization*. The Feynman rules and normalizations of vertices and propagators are given in Section 22.5.

The gauge field two-point vertex function. Four diagrams contribute to the two-point function, (a) and (b) corresponding to the gauge field loops, (c) to the Faddeev–Popov ghost loop and (d) to the matter-fermion loop (see Fig. 23.1). The diagram (b) corresponding to the self-contraction of the gauge four-point vertex vanishes in dimensional regularization. The fermion loop contribution (d) has already been calculated in Section 21.12.1 up to a simple geometric factor. Diagram (a) is given by (f_{abc} are structure constants of the Lie algebra)

$$(a) = \frac{1}{2} f_{acd} f_{bcd} \int \frac{d^d q}{(2\pi)^d} \frac{N_{\mu\nu}(k, q)}{q^2(k+q)^2}, \quad (23.53)$$

with

$$\begin{aligned} N_{\mu\nu}(k, q) &= \delta_{\mu\nu}(5k^2 + 2k \cdot q + 2q^2) + k_\mu k_\nu(d-6) + (q^\mu k_\nu + q^\nu k_\mu)(2d-3) \\ &\quad + 2q_\mu q_\nu(2d-3). \end{aligned} \quad (23.54)$$

To calculate the diagrams, we project the integrand over $\delta_{\mu\nu}$ and $k_\mu k_\nu$ (see equations (10.10) and (10.11)), and use repeatedly the identity

$$2k \cdot q = (k+q)^2 - k^2 - q^2.$$

We set

$$f_{acd} f_{bcd} = C(G) \delta_{ab}. \quad (23.55)$$

A short calculation then yields the divergent part ($\alpha = g^2/4\pi$):

$$(a)_{\text{div.}} = \delta_{ab} \frac{C(G)}{12} (19k^2 \delta_{\mu\nu} - 22k_\mu k_\nu) \frac{\alpha}{2\pi\varepsilon}. \quad (23.56)$$

Diagram (c) is given by

$$(c) = -f_{acd} f_{bcd} \int \frac{d^d q}{(2\pi)^d} \frac{q_\mu(k+q)_\nu}{q^2(k+q)^2}. \quad (23.57)$$

The divergent part is

$$(c)_{\text{div.}} = \delta_{ab} (k^2 \delta_{\mu\nu} + 2k_\mu k_\nu) \frac{1}{12} C(G) \frac{\alpha}{2\pi\varepsilon}. \quad (23.58)$$

Note that both divergent contributions are not separately transverse. By adding them, we find the divergent part of the two-point function in the absence of fermions, which now is transverse as expected:

$$\left[\tilde{\Gamma}_{\mu\nu}^{(2)ab} \right]_{\text{div.}} = \delta_{ab} (k^2 \delta_{\mu\nu} - k_\mu k_\nu) \frac{5}{3} C(G) \frac{\alpha}{2\pi\varepsilon}. \quad (23.59)$$

Therefore, we obtain the relation, at one-loop order, between renormalization constants,

$$\frac{Z_A}{Z_g} = 1 + \frac{5}{3} C(G) \frac{\alpha}{2\pi\varepsilon}. \quad (23.60)$$

Adapting the result (21.113), we infer the additional fermion contribution,

$$\frac{Z_A}{Z_g} = 1 + \left(\frac{5}{3} C(G) - \frac{4}{3} T(R) \right) \frac{\alpha}{2\pi\varepsilon}, \quad (23.61)$$

where the fermions belong to the representation R , and $T(R)$ is the trace of the square of the generators in this representation:

$$\text{tr } \mathbf{t}^a \mathbf{t}^b = -\delta_{ab} T(R). \quad (23.62)$$

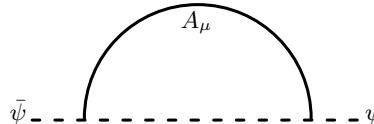


Fig. 23.2 One-loop contribution to the fermion two-point function.

The fermion two-point vertex function. Only the diagram of Fig. 23.2 contributes to the fermion two-point function $\Gamma^{(2)}$. It only differs from its counterpart in QED by a geometric factor:

$$\tilde{\Gamma}_{1\text{loop}}^{(2)}(k) = g^2 \int \frac{d^d q}{(2\pi)^d} \mathbf{t}^a \gamma_\mu \frac{1}{i\cancel{q} + m} \mathbf{t}^a \gamma_\mu \frac{1}{(k - q)^2} \quad (23.63)$$

(for simplicity, we have given the same mass m to all fermions since this does not affect the result). Since we need only the field renormalization, we can project the integrand over \mathbf{k} . The following identity is useful:

$$\gamma_\nu \gamma_\mu \gamma_\nu = (2 - d) \gamma_\mu.$$

Calculating the divergent part of the integral, we obtain the fermion field renormalization Z_F at one-loop order:

$$\mathbf{t}^a \mathbf{t}^a = -C(R) \mathbf{1} \Rightarrow Z_F = 1 - C(R) \frac{\alpha}{2\pi\varepsilon}. \quad (23.64)$$

The gauge field fermion vertex. Two diagrams contribute at one-loop order (Fig. 23.3).



Fig. 23.3 The gauge field fermion vertex at one loop: dotted lines correspond to fermions.

The contribution (a) to $\tilde{\Gamma}_{\mu,a}^{1,2}(k, p_1, p_2)$, which has a QED counterpart, is given by

$$(a) = g^2 \int \frac{d^d q}{(2\pi)^d} t^b \gamma_\nu \frac{1}{iq + m} \mathbf{t}^a \gamma_\mu \frac{1}{iq + ik + m} \mathbf{t}^b \gamma_\nu \frac{1}{(p_1 - q)^2}. \quad (23.65)$$

To calculate the divergent part of the integral, we multiply by γ_μ and take the trace. We also use the identity

$$\mathbf{t}^b \mathbf{t}^a \mathbf{t}^b = (\frac{1}{2}C(G) - C(R)) \mathbf{t}^a.$$

We then find

$$(a)_{\text{div.}} = \left(C(R) - \frac{1}{2}C(G) \right) \mathbf{t}^a \gamma_\mu \frac{\alpha}{2\pi\varepsilon}. \quad (23.66)$$

The diagram (b), which is specific to a non-Abelian theory, is given by

$$(b) = if_{abc}g^2 \int \frac{d^d q}{(2\pi)^d} \mathbf{t}^b \gamma_\nu \frac{1}{ip_1 - iq + m} \mathbf{t}^c \gamma_\rho \frac{V_{\mu\nu\rho}(k, q, -k - q)}{q^2(k + q)^2}, \quad (23.67)$$

with (equation (22.59)):

$$V_{\mu\nu\rho}(k, q, r) = (r - q)_\mu \delta_{\nu\rho} + (k - r)_\nu \delta_{\rho\mu} + (q - k)_\rho \delta_{\mu\nu}. \quad (23.68)$$

The divergent part is

$$(a)_{\text{div.}} = \frac{3}{2}C(G)\mathbf{t}^a \gamma_\mu \frac{\alpha}{2\pi\varepsilon}. \quad (23.69)$$

It follows that

$$Z_F Z_A^{1/2} = 1 - (C(R) + C(G)) \frac{\alpha}{2\pi\varepsilon}, \quad (23.70)$$

and, therefore, (equation (23.64)),

$$Z_A = 1 - 2C(G) \frac{\alpha}{2\pi\varepsilon}. \quad (23.71)$$

Finally, using the result (23.61), one obtains

$$Z_g = 1 - \left(\frac{11}{3}C(G) - \frac{4}{3}T(R) \right) \frac{\alpha}{2\pi\varepsilon}. \quad (23.72)$$

The β -function at one-loop order follows:

$$\beta(\alpha) = -\varepsilon \left[\frac{d \ln(\alpha Z_g)}{d\alpha} \right]^{-1} = \beta_2 \alpha^2 + O(\alpha^3),$$

with

$$\beta_2 = -\frac{1}{2\pi} \left[\frac{11}{3}C(G) - \frac{4}{3}T(R) \right]. \quad (23.73)$$

In the case of the $SU(N)$ group with N_F fermions in the fundamental representation the values of $C(G)$ and $T(R)$ are

$$C(G) = N, \quad T(R) = \frac{1}{2}N_F$$

and, therefore,

$$\beta(\alpha) = - \left(\frac{11N}{3} - \frac{2N_F}{3} \right) \frac{\alpha^2}{2\pi} + O(\alpha^3). \quad (23.74)$$

23.5.2 AF

The β -function is negative for small coupling for $N_F < 11N/2$ [252], which, in the case of $SU(3)$, means at most sixteen flavours. If this condition is met (in the SM, $N_F = 6$), by contrast with QED, $\alpha = 0$ is a UV fixed point (for details see Chapter 24), and the gauge theory is *asymptotically free*. This means that the effective interaction decreases at short distance (and thus the bare coupling constant is smaller than the coupling constant at physical scale), a phenomenon that, intuitively, can be thought as an anti-screening effect due to relativistic quantum fluctuations.

For illustration, we also give here the two-loop order (four-loop order is known [254]) of the β -function. Setting,

$$\beta(\alpha) = \beta_2 \alpha^2 + \beta_3 \alpha^3 + O(\alpha^4), \quad (23.75)$$

then

$$4\pi^2 \beta_3 = -\frac{1}{6} [17C^2(G) - 6C(R)T(R) - 10C(G)T(R)]. \quad (23.76)$$

As shown in Section 9.10.1, these two first coefficients are independent of the renormalization scheme. For the $SU(N)$ group, with N_F flavours in the fundamental representation, one finds

$$4\pi^2 \beta_3 = -\frac{1}{6} [17N^2 - 5NN_F - 3N_F(N^2 - 1)/(2N)]. \quad (23.77)$$

23.6 Axial current, chiral gauge theories, and anomalies

In Sections 21.8 and 22.5, we have pointed out that none of the standard regularization methods can deal with one-loop diagrams in a simple way in the case of chiral gauge theories. We now indeed exhibit gauge theories with massless fermions and chiral symmetry where the axial current is not conserved. The divergence of the axial current, when it does not vanish, is called an *anomaly*. In particular, this may lead to obstructions to the construction of gauge theories where the gauge field couples differently to the two fermion chiral components. Several examples are physically important like the electromagnetic decay of the π_0 meson [257], the consistency conditions in the theory of weak electromagnetic interactions [258], or the $U(1)$ problem [259]. See also Refs. [260].

We first discuss the Abelian axial current, in four dimensions (the generalization to all even dimensions is straightforward), and then the general non-Abelian case. The only possible source of anomalies are one-loop fermion diagrams. This reduces the problem to the study of fermions in the background of gauge fields or, equivalently, of the properties of the determinant of the gauge covariant Dirac operator.

23.6.1 Abelian axial current and Abelian vector gauge field

We consider the QED-like fermion action $\mathcal{S}(\bar{\psi}, \psi)$ for massless Dirac fermions $\psi, \bar{\psi}$ in the background of an Abelian gauge field A_μ :

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^4x \bar{\psi}(x) \not{D} \psi(x), \text{ with } \not{D} \equiv \not{\partial} + ie\not{A}, \quad (23.78)$$

and the corresponding field integral,

$$\mathcal{Z}(A) = \int [d\psi d\bar{\psi}] \exp [-\mathcal{S}(\psi, \bar{\psi})] = \det \not{D}. \quad (23.79)$$

In what follows, we denote by $\langle \bullet \rangle$ expectation values with respect to the corresponding functional measure.

Regularizations can be found that preserve gauge invariance and, since the fermions are massless, chiral symmetry. Therefore, one could naively have expected that the corresponding axial current is conserved. However, the proof of current conservation involves space-dependent chiral transformations and, therefore, steps that cannot be regularized without breaking one of the symmetries.

The coefficient of $\partial_\mu\theta(x)$ in the variation of the action under a space-dependent chiral transformation,

$$\psi_\theta(x) = e^{i\theta(x)\gamma_5} \psi(x), \quad \bar{\psi}_\theta(x) = \bar{\psi}(x) e^{i\theta(x)\gamma_5}, \quad (23.80)$$

yields the axial current $J_\mu^5(x)$. For the action (23.78), one finds,

$$\delta\mathcal{S} = \int d^4x \partial_\mu\theta(x) J_\mu^5(x), \quad \text{with} \quad J_\mu^5(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x). \quad (23.81)$$

After the transformation (23.80), $\mathcal{Z}(A)$ becomes

$$\mathcal{Z}(A, \theta) = \det \left[e^{i\gamma_5\theta(x)} \not{D} e^{i\gamma_5\theta(x)} \right]. \quad (23.82)$$

Since $\det(e^{i\gamma_5\theta}) = 1$, one could, again, naively conclude that $\mathcal{Z}(A, \theta) = \mathcal{Z}(A)$ and, therefore, that the current $J_\mu^5(x)$ is conserved. This is a conclusion we now check by an explicit calculation of the expectation value of $\partial_\mu J_\mu^5(x)$, in the case of the action (23.78).

Regularizations.

(i) For any regularization consistent with the Hermiticity of γ_5 ,

$$|\mathcal{Z}(A, \theta)|^2 = \det(\not{D}\not{D}^\dagger).$$

Therefore, an anomaly can appear only in the imaginary part of $\ln \mathcal{Z}$.

(ii) If the regularization is gauge invariant, $\mathcal{Z}(A, \theta)$ is also gauge invariant. Therefore, a possible anomaly will also be gauge invariant. One regularization scheme, which has the required property, is based on regulator fields. But, as the discussion of Section 8.4.3 shows, at least one regulator field must be an unpaired massive boson with spin, dividing the fermion determinant by a factor $\det(\not{D} + \Lambda)$. If this boson has a chiral charge, global chiral symmetry is broken by the mass Λ ; if it has no chiral charge, global chiral symmetry is preserved, and the determinant is independent of θ for $\theta(x)$ constant, but then the ratio of determinants is not invariant under local chiral transformations.

General form of a possible anomaly. The dimension of the operator $\partial_\mu J_\mu^5(x)$ is 4 and, since a possible anomaly is a large momentum or short-distance effect, $\langle \partial_\mu J_\mu^5(x) \rangle$ can only be a local function of A_μ of dimension 4. The fields A_μ and $\partial_\mu\theta$ have dimension 1, and no mass parameter is available. In addition, parity implies that it is proportional to the completely antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$. This determines $\langle \partial_\mu J_\mu^5(x) \rangle$ up to a multiplicative constant,

$$\langle \partial_\lambda J_\lambda^5(x) \rangle \propto e^2 \epsilon_{\mu\nu\rho\sigma} \partial_\mu A_\nu(x) \partial_\rho A_\sigma(x) \propto e^2 \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x),$$

$F_{\mu\nu}$ being the electromagnetic tensor. The possible anomaly is thus automatically gauge invariant. Moreover, it is a total derivative, since

$$\epsilon_{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x) = 4 \partial_\mu [\epsilon_{\mu\nu\rho\sigma} A_\nu(x) \partial_\rho A_\sigma(x)]. \quad (23.83)$$

Anomaly and three-current function. To determine the multiplicative factor, which is the only regularization dependent feature, it is sufficient to calculate the coefficient of term quadratic in A in the expansion of $\langle \partial_\lambda J_\lambda^5(x) \rangle$ in powers of A . This corresponds to the calculation of the three-point function of one axial current and two vector currents, which, at one-loop order, is given by the sum of two Feynman diagrams.

23.6.2 Regulator fields and explicit anomaly calculation

Instead of calculating Feynman diagrams in the Fourier representation, we evaluate the anomaly more directly, by adding to the action (23.78) the contribution of a charged boson field ϕ with spin 1/2 and a large mass Λ (see Section 12.7.2). We obtain the regularized *gauge-invariant action*,

$$\mathcal{S}(\bar{\psi}, \psi, \bar{\phi}, \phi) = \int d^4x [-\bar{\psi}(x)\not{D}\psi(x) + \bar{\phi}(x)(\not{D} + \Lambda)\phi(x)].$$

The corresponding regularized field integral (23.79) becomes

$$\mathcal{Z}(A) = \int [d\psi d\bar{\psi} d\phi d\bar{\phi}] \exp [-\mathcal{S}(\bar{\psi}, \psi, \bar{\phi}, \phi)]. \quad (23.84)$$

We perform a change of variables of the form of a space-dependent chiral transformation (23.80) acting identically on the fermion and boson fields,

$$\begin{aligned} \psi_\theta(x) &= e^{i\theta(x)\gamma_5} \psi(x), & \bar{\psi}_\theta(x) &= \bar{\psi}(x) e^{i\theta(x)\gamma_5}, \\ \phi_\theta(x) &= e^{i\theta(x)\gamma_5} \phi(x), & \bar{\phi}_\theta(x) &= \bar{\phi}(x) e^{i\theta(x)\gamma_5}. \end{aligned}$$

In the transformation, the integration measure $[d\phi d\bar{\phi} d\psi d\bar{\psi}]$ is invariant.

The variation $\delta\mathcal{S}$ of the action at first order in θ is

$$\delta\mathcal{S} = \int d^4x [\partial_\mu \theta(x) J_\mu^5(x) + 2i\Lambda\theta(x)\bar{\phi}(x)\gamma_5\phi(x)], \quad (23.85)$$

with

$$J_\mu^5(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x) - i\bar{\phi}(x)\gamma_5\gamma_\mu\phi(x).$$

We integrate by parts the first term in expression (23.85) to factorize $\theta(x)$, then expand the integrand in the field integral in powers of θ and integrate. We express that the field integral is not modified by a change of variables. Identifying the coefficient of $\theta(x)$, we obtain the equation

$$\begin{aligned} \langle \nabla \cdot \mathbf{J}^5(x) \rangle &= 2i\Lambda \langle \bar{\phi}(x)\gamma_5\phi(x) \rangle \\ &= i\Lambda \text{tr} \langle x | (\gamma_5(\not{D} + \Lambda)^{-1} + (\not{D} + \Lambda)^{-1}\gamma_5) | x \rangle, \end{aligned} \quad (23.86)$$

where the QM bra–ket notation has been used in the second line to specify matrix elements of operators, and the trace refers only to γ matrices.

The divergence of the axial current comes here from the *boson contribution*. The expression (23.86) can be transformed by using the following identities,

$$\begin{aligned} (\not{D} + \Lambda)^{-1}\gamma_5 &= -\gamma_5(\not{D} - \Lambda)^{-1}, \\ (\not{D} + \Lambda)^{-1} - (\not{D} - \Lambda)^{-1} &= 2\Lambda(\Lambda^2 - \not{D}^2)^{-1}. \end{aligned}$$

Setting $\gamma_\mu\gamma_\nu = \delta_{\mu\nu} + \sigma_{\mu\nu}$ with $\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$, one obtains the relation,

$$\not{D}^2 = D^2 + \frac{1}{2}\sigma_{\mu\nu}F_{\mu\nu}.$$

The contribution of order A^2 is obtained by expanding up to second order in $F_{\mu\nu}$ the expression

$$\begin{aligned} \left(\Lambda^2 - \not{D}^2\right)^{-1} &= (\Lambda^2 - D^2)^{-1} + \frac{1}{2} (\Lambda^2 - D^2)^{-1} F_{\mu\nu} \sigma_{\mu\nu} (\Lambda^2 - D^2)^{-1} \\ &+ \frac{1}{4} (\Lambda^2 - D^2)^{-1} F_{\mu\nu} \sigma_{\mu\nu} (\Lambda^2 - D^2)^{-1} F_{\rho\sigma} \sigma_{\rho\sigma} (\Lambda^2 - D^2)^{-1} + O(A^3). \end{aligned} \quad (23.87)$$

The trace with γ_5 of the two first terms vanishes. In the third term, one uses the identity (A12.49),

$$\text{tr } \gamma_5 \gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma = -4 \epsilon_{\mu\nu\rho\sigma}$$

and, therefore,

$$\text{tr } \gamma_5 \sigma_{\mu\nu} \sigma_{\rho\sigma} = -4 \epsilon_{\mu\nu\rho\sigma}.$$

In the denominators, one can now substitute $D^2 \mapsto \nabla^2$. One obtains,

$$\langle \nabla \cdot \mathbf{J}^5(x) \rangle = -2i\Lambda^2 \epsilon_{\mu\nu\rho\sigma} \langle x | [\Lambda^2 - \nabla^2]^{-1} F_{\mu\nu} [\Lambda^2 - \nabla^2]^{-1} F_{\rho\sigma} [\Lambda^2 - \nabla^2]^{-1} | x \rangle.$$

In the infinite Λ limit, the support of $\langle x | [\Lambda^2 - \nabla^2]^{-1} | y \rangle$ shrinks to $x = y$. The argument of $F_{\mu\nu}$ can be replaced everywhere by x . The expression reduces to

$$\langle \nabla \cdot \mathbf{J}^5(x) \rangle = -2i\Lambda^2 \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x) \langle x | [\Lambda^2 - \nabla^2]^{-3} | x \rangle.$$

Finally,

$$\langle x | [\Lambda^2 - \nabla^2]^{-3} | x \rangle = \frac{1}{16\pi^4} \int \frac{d^4 k}{(k^2 + \Lambda^2)^3} = \frac{1}{32\pi^2 \Lambda^2}.$$

The divergence of the axial current thus is

$$\langle \nabla \cdot \mathbf{J}^5(x) \rangle = -\frac{i}{16\pi^2} F_{\mu\nu}(x) \tilde{F}_{\mu\nu}(x), \quad (23.88)$$

where $F_{\mu\nu}$ is the electromagnetic tensor, and $\tilde{F}_{\mu\nu}$ its dual defined by

$$\tilde{F}_{\mu\nu}(x) = \epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma}(x). \quad (23.89)$$

With the more standard normalization of the electromagnetic tensor, the coefficient becomes $ie^2/16\pi^2 = i\alpha/4\pi$.

Therefore, in a QED-like gauge-invariant field theory with massless fermions, the axial current is not conserved: this is called the *chiral anomaly*.

Since global chiral symmetry is not broken, the integral over the whole space of the anomalous term must vanish at least for a class of ‘small’ gauge fields. This condition is verified since the anomaly can be written as a total derivative,

$$F_{\mu\nu}(x) \tilde{F}_{\mu\nu}(x) = -4e^2 \partial_\mu (\epsilon_{\mu\nu\rho\sigma} A_\nu(x) \partial_\rho A_\sigma(x)). \quad (23.90)$$

The space integral of the anomalous term depends only on the behaviour of the gauge field at the boundaries, and this property establishes a connection between *topology and anomalies*.

Equation (23.88) also implies

$$\ln \det \left[e^{i\gamma_5 \theta(x)} \not{D} e^{i\gamma_5 \theta(x)} \right] = \ln \det \not{D} + \frac{i}{16\pi^2} \int d^4 x \theta(x) F_{\mu\nu}(x) \tilde{F}_{\mu\nu}(x). \quad (23.91)$$

General even dimensions. The generalization to all even dimensions $2n$ is straightforward. The result can be derived by expanding an expression analogous to expression (23.86) up to degree n in A . If again gauge invariance is imposed, the anomaly in the divergence of the axial current $J_\lambda^{2n+1}(x)$ becomes

$$\langle \nabla \cdot \mathbf{J}^{2n+1}(x) \rangle = \frac{2i}{(4i\pi)^n n!} \epsilon_{\mu_1\nu_1\dots\mu_n\nu_n} F_{\mu_1\nu_1}(x) \cdots F_{\mu_n\nu_n}(x), \quad (23.92)$$

where $\epsilon_{\mu_1\nu_1\dots\mu_n\nu_n}$ is the completely antisymmetric tensor with $\epsilon_{1,2,\dots,2n} = 1$.

Chiral gauge theory. A gauge theory is consistent only if the gauge field is coupled to a conserved current. The anomaly thus prevents the construction of a theory that would have both an Abelian gauge vector and axial symmetry, with an action in the fermion sector of the form

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^4x \bar{\psi}(x) [\partial + ie\mathbf{A}(x) + i\gamma_5\mathbf{B}(x)] \psi(x). \quad (23.93)$$

In the gauge theory, current conservation is a WT identity for the AAB correlation function.

In such a theory, the one-loop diagrams contributing to the BBB correlation function are formally identical to those contributing to AAB , because two γ_5 cancel. Therefore, they also yield an anomaly that cannot be removed since the correlation function, by definition, is symmetric in its three arguments. This even prevents the construction of a theory with a purely axial gauge symmetry ($e = 0$).

A solution to both problems can be provided by the introduction of another fermion of opposite chiral coupling in order to cancel the anomaly. With more fermions other coupling combinations are possible. However, in the purely axial case, it is simple to show that a theory with two fermions of opposite chiral charges can be rewritten as a vector theory by combining differently the chiral components of both fermions.

23.6.3 Non-Abelian vector gauge theories and Abelian axial current

We still consider an Abelian axial current, but now in the framework of a non-Abelian gauge theory. The fermion fields transform non-trivially under a unitary or orthogonal gauge group G and \mathbf{A}_μ is the corresponding gauge field. The fermion action is

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^4x \bar{\psi}(x) \mathbf{D} \psi(x), \quad \text{with } \mathbf{D} = \partial + \mathbf{A}. \quad (23.94)$$

The axial current

$$J_\mu^5(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x),$$

is still gauge invariant. Therefore, no new calculation is needed; the anomaly is completely determined by dimensional analysis, gauge invariance and the preceding calculation that yields the term of order \mathbf{A}^2 :

$$\nabla \cdot \mathbf{J}^5(x) = -\frac{i}{16\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{tr} \mathbf{F}_{\mu\nu}(x) \mathbf{F}_{\rho\sigma}(x), \quad (23.95)$$

in which $\mathbf{F}_{\mu\nu}$ is now the corresponding curvature tensor. Again, this expression must be a total derivative. One can verify the relation,

$$\epsilon_{\mu\nu\rho\sigma} \text{tr} \mathbf{F}_{\mu\nu}(x) \mathbf{F}_{\rho\sigma}(x) = 4 \epsilon_{\mu\nu\rho\sigma} \partial_\mu \text{tr} [\mathbf{A}_\nu(x) \partial_\rho \mathbf{A}_\sigma(x) + \frac{2}{3} \mathbf{A}_\nu(x) \mathbf{A}_\rho(x) \mathbf{A}_\sigma(x)]. \quad (23.96)$$

23.6.4 Anomaly and index of the Dirac operator

We assume that the spectrum of \mathbf{D} , the Dirac operator in a non-Abelian gauge field (equation (23.94)), is discrete (temporarily enclosing the fermions in a box, if necessary) and denote by d_n and $\varphi_n(x)$ the corresponding eigenvalues and eigenvectors:

$$\mathbf{D}\varphi_n(x) = d_n\varphi_n(x). \quad (23.97)$$

The eigenvalues are gauge invariant, because in a gauge transformation of element $\mathbf{g}(x)$, the Dirac operator becomes

$$\mathbf{D} \mapsto \mathbf{g}^{-1}(x)\mathbf{D}\mathbf{g}(x) \Rightarrow \varphi_n(x) \mapsto \mathbf{g}(x)\varphi_n(x).$$

For a unitary or orthogonal group, the massless Dirac operator is anti-Hermitian; therefore, the eigenvalues are imaginary and the eigenvectors orthogonal. In addition, we choose them with unit norm.

The anticommutation relation $\mathbf{D}\gamma_5 + \gamma_5\mathbf{D} = 0$ implies

$$\mathbf{D}\gamma_5\varphi_n(x) = -d_n\gamma_5\varphi_n(x). \quad (23.98)$$

Therefore, either d_n is different from 0, and $\gamma_5\varphi_n$ is an eigenvector of \mathbf{D} with eigenvalue $-d_n$, or d_n vanishes. The eigenspace corresponding to the eigenvalue 0 then is invariant under γ_5 , which can be diagonalized: the eigenvectors of \mathbf{D} can be chosen eigenvectors of definite chirality, that is, eigenvectors of γ_5 with eigenvalue ± 1 ,

$$\mathbf{D}\varphi_n(x) = 0, \quad \gamma_5\varphi_n(x) = \pm\varphi_n(x).$$

We denote by n_+ and n_- the dimensions of the eigenspaces of positive and negative chirality, respectively.

We now consider the determinant of the operator $(\mathbf{D} + m)$ regularized by mode truncation (mode regularization):

$$\det_N(\mathbf{D} + m) = \prod_{n \leq N} (d_n + m), \quad (23.99)$$

keeping the N lowest eigenvalues of \mathbf{D} (in modulus), with $N - n_+ - n_-$ even, in such a way that the corresponding subspace is γ_5 invariant.

The regularization is gauge invariant because the eigenvalues of \mathbf{D} are gauge invariant.

In the truncated space, the trace of γ_5 is the index of the Dirac operator:

$$\text{tr } \gamma_5 = n_+ - n_-. \quad (23.100)$$

It does not vanish if $n_+ \neq n_-$, a situation that endangers axial current conservation.

In a chiral transformation (23.80) with θ constant, the regularized determinant of $(\mathbf{D} + m)$ becomes

$$\det_N(\mathbf{D} + m) \mapsto \det_N(e^{i\theta\gamma_5}(\mathbf{D} + m)e^{-i\theta\gamma_5}).$$

We now consider the various eigenspaces.

If $d_n \neq 0$, the matrix γ_5 is represented by the Pauli matrix σ_1 in the sum of eigenspaces corresponding to the two eigenvalues $\pm d_n$ and $(\mathbf{D} + m)$ is represented by $d_n\sigma_3 + m$. The determinant in the subspace is then

$$\det(e^{i\theta\sigma_1}(d_n\sigma_3 + m)e^{-i\theta\sigma_1}) = \det e^{2i\theta\sigma_1} \det(d_n\sigma_3 + m) = m^2 - d_n^2,$$

because σ_1 is traceless.

In the eigenspace of vanishing eigenvalues with positive chirality, of dimension n_+ , γ_5 is diagonal with eigenvalue 1 and thus

$$m^{n_+} \mapsto m^{n_+} e^{2i\theta n_+}.$$

Similarly, in the eigenspace of vanishing eigenvalues of chirality -1 ,

$$m^{n_-} \mapsto m^{n_-} e^{-2i\theta n_-}.$$

Therefore,

$$\det_N(e^{i\theta\gamma_5}(\not{D} + m) e^{i\theta\gamma_5}) = e^{2i\theta(n_+ - n_-)} \det_N(\not{D} + m).$$

The ratio of both determinants,

$$\det_N \left[(e^{i\gamma_5\theta}(\not{D} + m) e^{i\gamma_5\theta}) (\not{D} + m)^{-1} \right] = e^{2i\theta(n_+ - n_-)}, \quad (23.101)$$

is independent of N , and the limit $N \rightarrow \infty$ can be taken. The left-hand side of equation (23.101) is obviously 1 when $\theta = n\pi$, which implies that the coefficient of 2θ in the right-hand side must be an integer, and shows that both sides are consistent.

The variation of $\ln \det(\not{D} + m)$ at first order in θ ,

$$\ln \det \left[(e^{i\gamma_5\theta}(\not{D} + m) e^{i\gamma_5\theta}) (\not{D} + m)^{-1} \right] = 2i\theta(n_+ - n_-),$$

is related to the variation of the action (23.78), and thus to the expectation value of the integral of the divergence of the axial current $\langle \int d^4x \partial_\mu J_\mu^5(x) \rangle$. In the limit $m = 0$, it is thus related to the space integral of the chiral anomaly (23.95):

$$-\frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \int d^4x \operatorname{tr} \mathbf{F}_{\mu\nu}(x) \mathbf{F}_{\rho\sigma}(x) = n_+ - n_-. \quad (23.102)$$

Concerning this result, several comments can be made:

(i) at first order in θ , in the absence of a regularization, we have calculated ($\ln \det = \operatorname{tr} \ln$)

$$\ln \det [1 + i\theta (\gamma_5 + (\not{D} + m)\gamma_5(\not{D} + m)^{-1})] \sim 2i\theta \operatorname{tr} \gamma_5,$$

where we have used the cyclic property of the trace. Since the trace of the matrix γ_5 vanishes, naively, one could have expected a vanishing result. However, trace here means trace in γ space and in coordinate space, and γ_5 really stands here for $\gamma_5 \delta^{(d)}(x - y)$. The mode regularization yields a well-defined finite result for the undefined product $0 \times \delta^d(0)$ [261].

(ii) The quantization of the integral (23.102) implies that the value of the anomaly is related to topological properties of the gauge field [262], since the integral does not change when the gauge field is deformed continuously. The integral of the anomaly over the whole space thus depends only on the behaviour at large distances of the curvature tensor $\mathbf{F}_{\mu\nu}$, and the anomaly must be a total derivative as equation (23.96) confirms.

(iii) Gauge field configurations exist for which the right-hand side of equation (23.102) does not vanish, for example, instantons as we show in Section 39.6. We have shown previously that, if massless fermions are coupled to such gauge fields, the determinant resulting from the fermion integration necessarily vanishes. This has some physical implications which are examined in Sections 23.7.2 and 39.6.

(iv) One might be surprised that $\det \not{D}$ is not invariant under global chiral transformations. However, we have just established that when the integral of the anomaly does not vanish, $\det \not{D}$ vanishes. This explains that, to give a meaning to the right-hand side of equation (23.101), we have been forced to introduce a mass to obtain a non-trivial result. The determinant of \not{D} in the subspace orthogonal to eigenvectors with vanishing eigenvalue, even in presence of a mass, is chiral invariant by parity doubling, but for $n_+ \neq n_-$ not the determinant in the eigenspace of eigenvalue 0, because the trace of γ_5 does not vanish in the eigenspace (equation (23.100)). In the limit $m \rightarrow 0$, the complete determinant vanishes but not the ratio of determinants for different values of θ , because the powers of m cancel.

23.6.5 Non-Abelian anomaly: General axial current

We now discuss the problem of the conservation of a general axial current in the example of a fermion action that has a $G \times G$ chiral symmetry (subgroup of $U(N) \times U(N)$), in the background of non-Abelian vector gauge fields. The generators of the gauge group may or may not be related to the diagonal subgroup G of $G \times G$, which correspond to vector currents.

We denote by t^α the generators of G . The current then has the form

$$J_\mu^{5\alpha}(x) = -\bar{\psi}(x)\gamma_5\gamma_\mu t^\alpha\psi(x). \quad (23.103)$$

When the gauge group is connected with the chiral group, the current conservation equation involves the gauge covariant derivative ($\mathbf{D}_\mu = \partial_\mu + [A_\mu, \bullet]$):

$$\mathbf{D}_\mu J_\mu^{5\alpha}(x) = 0. \quad (23.104)$$

In the calculation of the contribution to the anomaly coming from terms quadratic in the gauge fields, the only modification in the previous results is the appearance of a different geometric factor. Then, the complete form of the anomaly is dictated by gauge covariance. One finds

$$\mathbf{D}_\lambda J_\lambda^{5\alpha}(x) = -\frac{i}{16\pi^2}\epsilon_{\mu\nu\rho\sigma} \text{tr } t^\alpha \mathbf{F}_{\mu\nu}(x) \mathbf{F}_{\rho\sigma}(x). \quad (23.105)$$

In particular, if the gauge group is disconnected from the chiral group, the anomaly is proportional to $\text{tr } t^\alpha$ and, therefore, only different from 0 for the Abelian factors of G .

23.6.6 Obstruction to gauge invariance

We consider a non-Abelian gauge field coupled only to left- or right-handed fermions:

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^4x \bar{\psi}(x) \frac{1}{2}(1 + \gamma_5) \not{D} \psi(x), \quad (23.106)$$

($\frac{1}{2}(1 - \gamma_5)$ is dealt with in the same way). We can construct a consistent gauge theory only if the partition function

$$\mathcal{Z}(\mathbf{A}_\mu) = \int [d\psi d\bar{\psi}] \exp [-\mathcal{S}(\psi, \bar{\psi})], \quad (23.107)$$

is gauge invariant.

If we introduce the generators \mathbf{t}^α of the gauge group in the fermion representation, we can write the corresponding current \mathbf{J}_μ as

$$J_\mu^\alpha(x) = -\bar{\psi}(x) \frac{1}{2}(1 + \gamma_5)\gamma_\mu \mathbf{t}^\alpha \psi(x). \quad (23.108)$$

The invariance of $\mathcal{Z}(\mathbf{A})$ under an infinitesimal gauge transformation again leads to a covariant conservation equation for the current:

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu(x) \rangle = 0.$$

The calculation of the term of degree 2 in the gauge field of the anomaly is straightforward with the regularization adopted in Section 23.6.2. The group structure yields a simple geometric factor. The global factor can be taken from the Abelian calculation. It differs from the result (23.88) by a factor 1/2 which comes from the projector $\frac{1}{2}(1 + \gamma_5)$. The general form of the term of third degree in the gauge field can also easily be found, but the calculation of the global factor is somewhat tedious. We argue in the next section that it can be obtained from consistency conditions. The complete expression reads:

$$(\mathbf{D}_\mu \mathbf{J}_\mu(x))^\alpha = -\frac{i}{24\pi^2} \partial_\mu \epsilon_{\mu\nu\rho\sigma} \text{tr } \mathbf{t}^\alpha (\mathbf{A}_\nu(x) \partial_\rho \mathbf{A}_\sigma(x) + \frac{1}{2} \mathbf{A}_\nu(x) \mathbf{A}_\rho(x) \mathbf{A}_\sigma(x)). \quad (23.109)$$

If the projector $\frac{1}{2}(1 + \gamma_5)$ is replaced by $\frac{1}{2}(1 - \gamma_5)$ the sign of the anomaly changes.

Unless this term vanishes identically, there is an obstruction to the construction of the gauge theory. It is easy to verify, taking into account the antisymmetry of the ϵ tensor, that the group factor is

$$d_{\alpha\beta\gamma} = \frac{1}{2} \text{tr} [\mathbf{t}^\alpha (\mathbf{t}^\beta \mathbf{t}^\gamma + \mathbf{t}^\gamma \mathbf{t}^\beta)]. \quad (23.110)$$

For a unitary representation, the generators \mathbf{t}^α are, with our conventions, anti-Hermitian. Therefore, the coefficients $d_{\alpha\beta\gamma}$ are purely imaginary:

$$d_{\alpha\beta\gamma}^* = \frac{1}{2} \text{tr} [\mathbf{t}^\alpha (\mathbf{t}^\beta \mathbf{t}^\gamma + \mathbf{t}^\gamma \mathbf{t}^\beta)]^* = -d_{\alpha\beta\gamma}. \quad (23.111)$$

For all real (the \mathbf{t}^α are antisymmetric) or ‘pseudo-real’ ($\mathbf{t}^\alpha = -S^T \mathbf{t}^\alpha S^{-1}$) representations, these coefficients vanish. It follows that the only non-Abelian groups which can lead to anomalies in four dimensions are: $SU(N)$ for $N \geq 3$, $SO(6)$ and E_6 .

23.6.7 Wess-Zumino consistency conditions

In Section 23.6.6, we have calculated the part of the anomaly which is quadratic in the gauge field, and asserted that the remaining part could then be inferred from geometric arguments. Indeed, the anomaly is the variation of a functional under an infinitesimal gauge transformation. As we argue in Section 29.3, this implies compatibility conditions [263], which are here constraints on the form of the anomaly. One convenient way to express these constraints is to express the nilpotency of BRST transformations (for details, see Chapter 26).

In a BRST transformation, the variation of the gauge field \mathbf{A}_μ takes the form (equation (22.46)):

$$\delta \mathbf{A}_\mu(x) = \mathbf{D}_\mu \mathbf{C}(x) \bar{\varepsilon}, \quad (23.112)$$

where \mathbf{C} is a non-physical ‘ghost field’ and $\bar{\varepsilon}$ an anticommuting constant. The corresponding variation of $\ln \mathcal{Z}(\mathbf{A})$ is

$$\delta \ln \mathcal{Z}(\mathbf{A}) = - \int d^4x \langle \mathbf{J}_\mu(x) \rangle \mathbf{D}_\mu \mathbf{C}(x) \bar{\varepsilon}. \quad (23.113)$$

We write the anomaly equation as,

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu(x) \rangle = \mathcal{A}(\mathbf{A}, x). \quad (23.114)$$

Equation (23.113), after an integration by parts, can then be rewritten as

$$\delta \ln \mathcal{Z}(\mathbf{A}) = \int d^4x \mathcal{A}(\mathbf{A}, x) \mathbf{C}(x) \bar{\varepsilon}. \quad (23.115)$$

Since \mathcal{AC} is a BRST variation, it is invariant under BRST transformation, where the gauge field transforms as in equations (22.46) and (23.112), and the fermion ghost $\mathbf{C}(x)$ as (equation (22.44)):

$$\delta \mathbf{C}(x) = \bar{\varepsilon} \mathbf{C}^2(x). \quad (23.116)$$

By expressing that \mathcal{AC} is BRST invariant, one obtains a constraint on the possible form of anomalies. It is simple to verify that this condition determines the term cubic in \mathbf{A} in the right-hand side of equation (23.109).

23.7 Anomalies: Applications in particle physics

We describe now three important physics applications of the anomaly analysis.

23.7.1 Weak-electromagnetic interactions: Anomaly cancellation

The condition of anomaly cancellation discussed in Section 23.6.6 constrains the model of weak-electromagnetic interactions. In the SM, for example, *the anomalous contributions of leptons cancels the quark contributions*. This cancellation occurs within each generation provided that for each flavour quarks exist in three states. In the weak-electromagnetic group $SU(2) \times U(1)$, $SU(2)$ alone is a safe group. Therefore, the problems come from the $U(1)$ factor. We expect a priori two conditions coming from the vertices with one $U(1)$ and two $SU(2)$ gauge fields and with three $U(1)$ gauge fields, but one verifies that the two conditions are equivalent. If one considers two $SU(2)$ and one $U(1)$ gauge fields, only $SU(2)$ doublets contribute and equation (23.109) leads to the condition:

$$\sum_{\text{all doublets}} Y_L \text{tr} \tau^\alpha \tau^\beta = 0,$$

in which Y_L is the $U(1)$ charge (see Sections 23.1–23.3). This condition reduces to

$$\sum_{\text{all doublets}} Y_L = 0. \quad (23.117)$$

The vertex with three $U(1)$ gauge fields yields the condition:

$$\sum_{\text{left-handed parts}} Y_L^3 - \sum_{\text{right-handed parts}} Y_R^3 = 0,$$

because the contributions to the anomaly of right-handed and left-handed couplings have opposite signs. In the SM, the left and right charges are related (equation (23.21)). Summing the charges of one doublet and the corresponding singlets, one obtains

$$\sum_{\text{all doublets}} (Y_L + 1)^3 + (Y_L - 1)^3 - 2Y_L^3 = 0,$$

a condition which reduces to equation (23.117).

In one generation, the lepton doublet has $Y_L = -1$ and the quark doublet $Y_L = 1/3$ (equations (23.22) and (23.34)). Therefore, a cancellation requires that the quarks exist in three states. These states are provided by the colour quantum number.

23.7.2 Electromagnetic π_0 decay

In an effective low-energy model for strong interactions (see Section 13.6), where hadrons are considered as elementary particles, based on a linearly broken $SU(2) \times SU(2)$ symmetry, the non-conservation of the axial current \mathbf{J}_μ^5 is expressed by the equation

$$\partial_\mu \mathbf{J}_\mu^5(x) = m_\pi^2 f_\pi \boldsymbol{\pi}(x), \quad (23.118)$$

where $\boldsymbol{\pi}(x)$ is the pion field. The third component $[J_\mu^5]_3$ of the current, corresponds to the neutral pion π_0 field in the right-hand side. After introduction of electromagnetic interactions in the model, the relation between the divergence of the axial current and the π_0 field makes it possible to calculate the electromagnetic decay rate of the neutral pion in a limit where the four-momentum \mathbf{k} of the pion vanishes. In the absence of anomalies, the expectation value of relation (23.118) multiplied by two photon fields implies that the decay rate vanishes for $\mathbf{k} = 0$, in contradiction with reasonable smoothness assumptions and experimental results. However, the anomaly equation (23.88) implies instead,

$$\partial_\mu [J_\mu^5]_3(x) = m_\pi^2 f_\pi \pi_0(x) - i \frac{\alpha}{8\pi} \epsilon_{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x). \quad (23.119)$$

Multiplying the equation by two photon fields, taking the expectation value and going to the limit $\mathbf{k} = 0$ to eliminate the left-hand side, one obtains a non-vanishing decay amplitude for a non-physical π_0 at zero total momentum. At leading order in the σ -model, one can extrapolate to the mass shell $k^2 = -m_\pi^2$. The theoretical decay rate becomes

$$\Gamma^{\text{theor.}} = \frac{\alpha^2 m_\pi^3}{64\pi^3 f_\pi^2} = 7.6 \text{ eV},$$

in very good agreement with the experimental value $\Gamma^{\text{exp.}} = (7.65 \pm 0.16)$ eV. A similar theoretical estimate is obtained in the quark model with massless quarks, for three colours.

A first non-vanishing estimate was derived by Steinberger, using straightforward Feynman graph calculation, long before the relation to anomalies had been discovered.

23.7.3 The solution of the $U(1)$ problem

In QCD with N_F quark flavours, in the limit of massless quarks the action has a symmetry corresponding to the $U(N_F) \times U(N_F)$ chiral group. The spontaneous breaking of the chiral symmetry to its diagonal subgroup $U(N_F)$ leads to expect N_F^2 Goldstone bosons associated with the axial currents.

From the preceding analysis, we know that the axial current corresponding to the $U(1)$ Abelian subgroup has an anomaly. However, the WT identities that imply the existence of Goldstone bosons correspond to constant group transformations and, therefore, involve only the space integral of the divergence of the current. Since the anomaly is a total derivative, one might have expected the integral to vanish.

However, the Euclidean field equations of non-Abelian gauge theories admit *instanton* solutions, which give a periodic structure to the vacuum (for details see Section 39.6). These instanton solutions correspond to gauge configurations which approach non-trivial pure gauges at infinity, and give the set of discrete non-vanishing values one expects from equation (23.102) to the space integral of the anomaly (23.95). This indicates (but no completely satisfactory calculation of the instanton contribution has been performed) that for small, but non-vanishing quark masses, the $U(1)$ axial current is far from being conserved and, therefore, no light would-be Goldstone boson is generated. This observation has resolved a long-standing puzzle, since experimentally no corresponding light pseudoscalar boson is found for $N_F = 2, 3$.

24 Large-momentum behaviour in quantum field theory (QFT)

So far, we have discussed mainly the infrared (IR) properties of quantum field theory (QFT). By contrast, in this chapter, we use renormalization group (RG) equations to characterize the large momentum behaviour of *renormalized* QFTs, assuming implicitly that a *universal* large momentum physics can be defined. Beyond perturbation theory, this is not obvious.

The initial effective QFT is valid only up to an energy-momentum scale much smaller than some cut-off Λ . Therefore, large momentum means much larger than the renormalization scale, but still much smaller than the cut-off scale. The existence of this large momentum physics also implies the existence of a crossover scale between low and large momentum physics (for an example, see also Section 15.10.3).

One theoretic reason for discussing the large momentum behaviour is the apparent connection between the existence of interacting renormalized QFTs and the presence of ultraviolet (UV) fixed points. The absence of identified UV fixed points in infrared-free QFTs, like the ϕ^4 field theory or quantum electrodynamics (QED), leads to the so-called *triviality issue* (see Section 9.12), which we examine again in Section 24.1.2.

The physics reason is that in collisions it is observed that quarks, fundamental particles of the Standard Model (SM) of particle physics, behave like free particles at the shortest distances presently accessible (the property of *asymptotic freedom*). This property can be explained by RG arguments when the RG β function is negative for small couplings (see Section 23.5.1). Therefore, the identification of QFTs that behave as free theories at short distance, or at large Euclidean momenta, is important, and this leads to examine the large momentum behaviour of all QFTs renormalizable in four dimensions.

First, we investigate scalar field theories, then add fermions and gauge fields and show that only theories having a non-Abelian gauge symmetry can be asymptotically free [264]. As an application, we calculate the total cross-section of electron–positron annihilation into hadrons at large momentum.

Note that, above two dimensions, the large momentum behaviour can entirely be discussed, using RG, in the framework of massless theories.

24.1 The $(\phi^2)^2$ Euclidean field theory: Large-momentum behaviour

We thus consider the effective (microscopic or bare) action (a momentum cut-off Λ is implied) for an N -component scalar field ϕ , in dimension d ,

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r_c \phi^2(x) + \frac{1}{4!} g_0 \Lambda^{4-d} (\phi^2(x))^2 \right], \quad (24.1)$$

where g_0 is dimensionless and r_c the critical value for which the physical mass vanishes.

However, the *perturbative massless* $(\phi^2)^2$ field theory exists only in dimension $d = 4$ and, in $d < 4$, in the sense of the $\varepsilon = 4 - d$ expansion.

In Section 15.4, we have derived the RG equations (equations (15.46)) for vertex functions in the Fourier representation, regularized by a momentum cut-off Λ .

The RG equations read

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0} - \frac{n}{2} \eta(g_0) \right) \tilde{\Gamma}^{(n)}(p_i; g_0, \Lambda) = 0, \quad (24.2)$$

where perturbative corrections decreasing like powers of Λ have been neglected.

At order g_0^2 , the RG functions are given by

$$\beta(g_0) = (d-4)g_0 + \frac{N+8}{48\pi^2} g_0^2 + O(g_0^3), \quad \eta(g_0) = \frac{N+2}{72(8\pi^2)^2} g_0^2 + O(g_0^3).$$

24.1.1 Dimensions $d < 4$

For $\varepsilon = 4 - d$ small, the theory has an IR fixed point $g_0^* \sim 48\pi^2\varepsilon/(N+8)$. In Section 15.5, we have solved the RG equations by introducing scale-dependent coupling and field renormalization. There, we have studied the IR behaviour; here, the flow of the effective coupling constant is reversed. For $g_0 < g_0^*$, the effective coupling at scale λ , $g_0(\lambda)$ decreases, for $g_0 > g_0^*$ it increases, and if initially $g_0 = g_0^*$, it is invariant.

Since increasing momenta amounts to decreasing the cut-off, a question arises: does one find some universal (continuum), non-IR physics before momenta become of the order of Λ ? Additional information is provided by solving the RG equation. We set

$$\tilde{\Gamma}^{(n)}(p_i; g_0, \Lambda) = \zeta^{n/2}(g_0) [M(g_0)]^{d-n(d-2)/2} G^{(n)}(p_i/M(g_0)),$$

where $M(g_0)$ is an RG invariant mass scale, satisfying the RG equation

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0} \right) M(g_0) = 0.$$

For $g_0 < g_0^*$,

$$M(g_0) = \Lambda g_0^{1/(4-d)} \exp \left[\int_{g_0}^0 dg' \left(\frac{1}{\beta(g')} - \frac{1}{(d-4)g'} \right) \right]. \quad (24.3)$$

Moreover,

$$\ln \zeta(g_0) = \int_0^{g_0} dg' \frac{\eta(g')}{\beta(g')}.$$

Since correlation functions depend only on the mass scale M_0 , a continuum large momentum behaviour can only be found if one can find momenta of order $M(g_0)$, or larger, but still much smaller than Λ . For $g_0 < g_0^*$, this yields the condition

$$M(g_0) \ll \Lambda \Rightarrow g_0 \ll 1. \quad (24.4)$$

This condition can only be satisfied if the initial coupling constant g_0 is close to a UV fixed point value. In the $(\phi^2)^2$ theory, in dimensions $d < 4$, the Gaussian fixed point $g_0 = 0$ is a UV fixed point. Therefore, initially g_0 must be sufficiently close to $g_0 = 0$. Then, the mass scale $M(g_0)$ is a crossover scale between IR and UV behaviours.

The condition (24.4) is equivalent to demand that the effective coupling constant $g_0(\mu/\Lambda)$ at scale $\mu \ll \Lambda$, which can be considered as a renormalized coupling g defined at renormalization scale μ ,

$$g \sim g_0(\mu/\Lambda), \quad \int_{g_0}^{g_0(\mu/\Lambda)} \frac{dg'}{\beta(g')} = \ln(\mu/\Lambda), \quad (24.5)$$

should be sufficiently far from its IR fixed point value.

24.1.2 The renormalized $(\phi^2)^2$ field theory for $d = 4$: The triviality issue

The Gaussian fixed point. A first conclusion, relevant for particle physics, is clear: in four dimensions, because the β -function is positive for g small, $g = 0$ is an IR fixed point, and the $(\phi^2)^2$ field theory is not asymptotically free at large momenta. Therefore, alone, it is not a suitable candidate to describe the physics of strong interactions at experimentally accessible short distances.

The triviality issue. The problem of the triviality [265] of the renormalized $(\phi^2)^2$ field theory in dimension 4 has already been examined in Section 9.12.2. We discuss it again, using the arguments of Section 24.1.1. The question is whether it is possible to find, for arbitrarily large values of the cut-off Λ , an initial (bare) coupling constant g_0 which yields a given renormalized coupling constant g . This problem has always a formal perturbative solution, but we want to discuss this question beyond perturbation theory.

We solve the massless RG equations as in Section 24.1.1, and introduce the RG invariant scale

$$M(g_0) = \Lambda \exp \left[\int_{g_0}^a \frac{dg'}{\beta(g')} \right],$$

where a is a (small) fixed value.

The condition $M(g_0) \ll \Lambda$ now plays an even more important role. Because $g_0 = 0$ is an IR fixed point, the existence of a non-trivial continuum UV behaviour is also a condition for the existence of a non-free $(\phi^2)^2$ field theory in the infinite cut-off limit. The same condition is recovered from equation (24.5):

$$\int_{g_0}^g \frac{dg'}{\beta(g')} = \ln(\mu/\Lambda).$$

For $\mu/\Lambda \rightarrow 0$, either g goes to the IR fixed point $g = 0$ (equation (9.108)), and the renormalized theory is free, or g_0 goes to another value where the integral diverges, which then is a UV fixed point.

Since, for small values of g_0 , the β -function remains positive, the existence of such a fixed point cannot be investigated by perturbative methods. However, various numerical investigations strongly suggest that such a fixed point does not exist. For instance, if a UV fixed point g_0^* can be found, for $g_0 > g_0^*$ the IR behaviour of the model is no longer mean-field like, up to logarithmic corrections. In particular, one would think that if a UV fixed point exists in some $(\phi^2)^2$ field theory, the result would apply to the corresponding Ising model or other $O(N)$ spin models, which can be obtained from the $(\phi^2)^2$ field theory after lattice regularization by taking the large g_0 limit.

In the case of a one component system (Ising-like), we present conclusive numerical evidence in Section 16.9 that, in two and three dimensions, the Ising model and the ϕ^4 field theory belong to the same universality class. This means that no IR unstable fixed point exists in these dimensions. In four dimensions, the evidence is somewhat weaker. This is expected within the RG framework, since the approach to scaling is only logarithmic. In higher dimensions, again the Ising model and the ϕ^4 fall in the same universality class. It would thus be somewhat surprising if, only in four dimensions, these models would behave differently, with moreover rather close exponents. Finally, numerical studies of the ϕ^4 field theory on the lattice have found no evidence of additional fixed points and are consistent with triviality.

Some of these remarks also apply to the $(\phi^2)^2$ theory with a small number of components. Moreover, we have calculated the β -function in the large N limit in Chapter 18 and found that it is proportional to g_0^2 .

Therefore, no fixed point exists in the domain of the $1/N$ expansion, that is, for g_0 of order $1/N$. On the other hand, for large g_0 , the $(\phi^2)^2$ theory becomes the non-linear σ -model which we have examined in the large N limit in Section 18.6.1. Again, we have found that it becomes a free field theory in four dimensions. Therefore, we have strong evidence that no non-trivial *renormalized* $(\phi^2)^2$ theory exists in four dimensions. Rigorous results would also imply triviality if it were possible to prove the divergence of the field renormalization.

Triviality and the Higgs boson. That the renormalized $(\phi^2)^2$ field theory is, most likely, a free field theory in the infinite cut-off limit, seems to be a problem for the Higgs sector of the SM of weak and electromagnetic interactions, which involves a $(\phi^2)^2$ interaction. Of course, then other coupling constants contribute to the $(\phi^2)^2$ coupling β -function, and the RG flow may be different, as we discuss in the coming sections. However, the most likely conclusion is that, in the Standard Model, this problem cannot be avoided. Considering the phenomenological success of the model, the conclusion is somewhat surprising.

However, the model has only been tested in some limited ranges of energies. Therefore, demanding that the QFT should be consistent on all scales has no real physics justification. If we keep the cut-off large but fixed, at the cut-off scale the theory will break down but, at lower scales, it may give acceptable physics answers. Correspondingly, the renormalized coupling can only vary in a limited range, which goes to 0 (logarithmically) for infinite cut-off. As we have already emphasized, the SM is an *effective* low-energy field theory, in which the cut-off, reflection of a larger mass scale where new physics appears, cannot be completely eliminated. Although the field theory is perturbatively renormalizable, it is not necessarily consistent on all scales.

One may then wonder about the meaning of the correlation functions as calculated from renormalized perturbation theory, that is, after the infinite cut-off limit is taken. We argue in Chapter 40 that there are intrinsic difficulties in reconstructing correlation functions from the knowledge of their perturbative expansion. Moreover, it follows from RG arguments that to a finite renormalized coupling constant can only correspond a complex bare coupling constant, the imaginary part vanishing to all orders in perturbation theory. Therefore, correlation functions, depending on the summation procedure, will either be complex, or will not satisfy field equations beyond perturbation theory.

Before discovery: An upper bound on the Higgs mass [266]. In the SM, the Higgs field through its various couplings gives masses to all fields. The observed masses determine the corresponding couplings. Before the experimental discovery of the Higgs boson, the Higgs field expectation value $\langle \phi \rangle$ was known, since it could be inferred from G_F , the weak coupling constant (equation (23.28)): $\langle \phi \rangle \approx 246$ GeV. By contrast, the Higgs self-interaction and, thus, the Higgs mass, were unknown. However, it was likely that the renormalized $(\phi^2)^2$ coupling would be such that perturbation theory remains, at least semi-quantitatively, applicable. Otherwise, the success of the SM in describing physics would have been difficult to understand.

In the perturbative regime, the Higgs mass increases with g . To obtain an upper bound on the Higgs mass, one has to examine what happens when g increases. For g large enough, the Higgs mass is mainly determined by the Higgs self-coupling. Therefore, one can consider a pure $(\phi^2)^2$ field theory. In the perturbative regime, RG arguments are applicable. Using the expansion of the β -function,

$$\beta(g_0) = \beta_2 g_0^2 + \beta_3 g_0^3 + O(g_0^4),$$

one solves equation (24.5) for g small.

One finds

$$\ln(\Lambda/\mu) = \frac{1}{\beta_2 g} + \frac{\beta_3}{\beta_2^2} \ln g + K(g_0) + O(g), \quad (24.6)$$

where the function $K(g_0)$, according to the preceding discussion, is bounded but can only be determined by non-perturbative methods. For g small, at leading order, the Higgs mass is related to the $|\phi|$ (=Higgs) expectation value by (see Section 17.2),

$$m_H^2 = \frac{1}{3} g \langle \phi \rangle^2 + O(g^2). \quad (24.7)$$

To minimize higher order corrections, one can choose $\mu = \langle \phi \rangle$. Eliminating g between equations (24.7) and (24.6), one finds

$$\ln\left(\frac{\Lambda}{\langle \phi \rangle}\right) = \frac{1}{3\beta_2} \frac{\langle \phi \rangle^2}{m_H^2} + \frac{2\beta_3}{\beta_2^2} \ln\left(\frac{m_H}{\langle \phi \rangle}\right) + \tilde{K}(g_0) + O(g).$$

If one neglects $\tilde{K}(g_0)$ and higher order corrections, one obtains a relation between the two ratios $\Lambda/\langle \phi \rangle$ and $m_H/\langle \phi \rangle$. If the Higgs particle of the SM exist, its mass must be smaller than the cut-off (which at this point only represents the onset of some new physics beyond the SM). For the values of the two coefficients of the β -function corresponding to $N = 4$ (equation (16.130)), $8\pi^2\beta_2 = 2$, $\beta_3/\beta_2^2 = -13/24$, one obtains an upper bound for m_H that ranges between 600 GeV and 700 GeV.

The determination of the Higgs boson mass (about 125 GeV) has rendered such bounds, which were initially very important, obsolete. Conversely, the Higgs boson mass being known, from the RG equations of the whole SM, one finds upper bounds (presumably physically irrelevant) on the scale of new physics of the order of 10^{10} GeV.

24.1.3 The ϕ_4^4 field theory for negative renormalized coupling

So far, we have implicitly assumed that the *renormalized coupling constant* g is always positive. Although this is not very plausible, it has been suggested that g could also be negative. This would mean that the identification between renormalized coupling and effective bare coupling is impossible, and the bare RG is somewhat pathological. Therefore, the issue must be discussed entirely within the framework of the renormalized theory.

Then, a question immediately arises: does a negative renormalized coupling correspond to a Hamiltonian bounded from below? We show in Section 7.11 that the vertex (1PI) generating functional (or thermodynamic potential) is related to the expectation value of the Hamiltonian in states with fixed field expectation value. In Section 16.5, from RG equations we infer that the behaviour of the thermodynamic potential for small field-expectation values is governed by IR fixed points. Conversely, the behaviour of the vertex functional density $\mathcal{V}(\varphi)$ for large constant fields is governed by UV fixed points. Adapting the derivation of Section 16.5, one obtains the RG equation satisfied by $\mathcal{V}(\varphi)$, subtracted at $\varphi = 0$, in a massless theory renormalized at scale μ :

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \eta(g) \varphi \frac{\partial}{\partial \varphi} \right] \mathcal{V}(\varphi, g, \mu) = 0. \quad (24.8)$$

Solving the equation by the method of characteristics, and using dimensional analysis, one finds

$$\mathcal{V}(\lambda \varphi, g, \mu) = \lambda^4 \mathcal{V}(Z^{1/2}(\lambda) \varphi, g(\lambda), \mu). \quad (24.9)$$

The functions $g(\lambda)$ and $Z(\lambda)$ satisfy the flow equations,

$$\lambda \frac{d}{d\lambda} g(\lambda) = \beta(g(\lambda)), \quad \lambda \frac{d}{d\lambda} \ln Z(\lambda) = -\eta(g(\lambda)). \quad (24.10)$$

The solutions (with boundary conditions $g(1) = g$, $Z(1) = 1$) can be written as

$$\ln \lambda = \int_g^{g(\lambda)} \frac{dg'}{\beta(g')}, \quad Z(\lambda) = \exp \left[- \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} \eta(g') \right]. \quad (24.11)$$

Therefore, the large field behaviour of $\mathcal{V}(\varphi)$ can be obtained by increasing λ and, thus, by studying the UV limit of the theory. At leading order, the RG functions are (equations (16.130) and (16.131)):

$$\beta(g) = \frac{1}{8\pi^2} \frac{(N+8)}{6} g^2 + O(g^3), \quad \eta(g) = \frac{1}{(8\pi^2)^2} \frac{(N+2)}{72} g^2 + O(g^3).$$

If we start from $g < 0$ and small, $g(\lambda)$ approaches the origin for λ large. The field renormalization $Z(\lambda)$ then tends towards a positive constant, and $\mathcal{V}(\varphi)$ can be taken from perturbation theory. Therefore, in the case of the $O(N)$ invariant $(\phi^2)^2$ field theory, φ being then the length of the vector φ , one obtains

$$g(\lambda) \sim -\frac{48\pi^2}{(N+8)\ln\lambda}, \quad \mathcal{V}(\lambda\varphi, g, \mu) \sim \frac{1}{4!} Z^2(\infty) \lambda^4 g(\lambda) \varphi^4. \quad (24.12)$$

The result shows that by increasing λ , \mathcal{V} can be made arbitrarily large and negative. Therefore, the corresponding Hamiltonian is not bounded from below and, thus, $g < 0$ is non-physical.

24.2 General ϕ^4 -like field theories: d=4

We now study the large momentum behaviour in a general renormalized ϕ^4 -like field theory in four dimensions. We consider an action for an N -component massless scalar field $\phi_i(x)$, which in the tree approximation has the form

$$\mathcal{S}(\phi) = \int d^4x \left[\frac{1}{2} \sum_i (\nabla \phi_i(x))^2 + \frac{1}{4!} \sum_{i,j,k,l} g_{ijkl} \phi_i(x) \phi_j(x) \phi_k(x) \phi_l(x) \right]. \quad (24.13)$$

The renormalized n -point correlation functions satisfy a slight generalization of the RG equation (10.82):

$$\mu \frac{\partial}{\partial \mu} \Gamma_{i_1 \dots i_n}^{(n)} + \sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial}{\partial g_{i'j'k'l'}} \Gamma_{i_1 \dots i_n}^{(n)} - \frac{1}{2} \sum_{m=1}^n \sum_{j_m} \eta_{i_m j_m} \Gamma_{i_1 \dots j_m \dots i_n}^{(n)} = 0. \quad (24.14)$$

The expansion of β_{ijkl} at leading order is (equation (10.86))

$$\beta_{ijkl} = \frac{1}{16\pi^2} \sum_{m,n} (g_{ijmn} g_{mnkl} + g_{ikmn} g_{mnl} + g_{ilmn} g_{mjnk}) + O(|g|^3). \quad (24.15)$$

Moreover, we know that η_{ij} is of order g^2 . Following the arguments given in Chapter 16, it is simple to verify that the thermodynamic potential density in a constant field φ , and subtracted at $\varphi = 0$, satisfies

$$\left(\mu \frac{\partial}{\partial \mu} + \sum_{i',j',k',l'} \beta_{i'j'k'l'} \frac{\partial}{\partial g_{i'j'k'l'}} - \frac{1}{2} \sum_{i,j} \eta_{ij} \varphi_i \frac{\partial}{\partial \varphi_j} \right) \mathcal{V}(\varphi) = 0. \quad (24.16)$$

To solve the equation by the method of characteristics (see Chapters 15 and 16), we introduce a scale parameter λ , scale-dependent coupling constants $g_{ijkl}(\lambda)$, and field renormalization matrix $Z_{ij}(\lambda)$ defined by

$$\lambda \frac{d}{d\lambda} g_{ijkl}(\lambda) = \beta_{ijkl}(g(\lambda)), \quad g_{ijkl}(1) = g_{ijkl}, \quad (24.17)$$

$$\left(\lambda \frac{dZ^{1/2}}{d\lambda} Z^{-1/2} \right)_{ij} = -\frac{1}{2} \eta_{ij}(g(\lambda)), \quad Z_{ij}(1) = \delta_{ij}. \quad (24.18)$$

We obtain

$$\mathcal{V}(\lambda \varphi_i, g_{ijkl}, \mu) = \lambda^4 \mathcal{V}(\varphi_i(\lambda), g_{ijkl}(\lambda), \mu), \quad \text{with } \varphi_i(\lambda) = \sum_j Z_{ij}^{1/2}(\lambda) \varphi_j. \quad (24.19)$$

We know from the analysis of Section 24.1.3 that the large field behaviour of $\mathcal{V}(\varphi)$ is governed by the UV fixed points of the theory. About the zeros of the RG β -function, little is known in the general case. However, there is a problem of interest for particle physics one can investigate: the existence of asymptotically free-field theories. If, for some initial value g_{ijkl} of the renormalized coupling constant, the effective coupling constant $g_{ijkl}(\lambda)$ flows into the origin for large scale, then $\mathcal{V}(\lambda \varphi)$ can be calculated from perturbation theory for λ large:

$$\mathcal{V}(\lambda \varphi) \propto \lambda^4 \sum_{i,j,k,l} \frac{g_{ijkl}(\lambda)}{4!} \varphi'_i \varphi'_j \varphi'_k \varphi'_l, \quad \text{with } \varphi'_i = \sum_j Z_{ij}^{1/2}(\infty) \varphi_j, \quad (24.20)$$

because the renormalization matrix Z_{ij} goes to a constant. Therefore, the boundedness of the Hamiltonian implies that

$$G(\lambda) = \sum_{i,j,k,l} g_{ijkl}(\lambda) \varphi_i \varphi_j \varphi_k \varphi_l, \quad (24.21)$$

must be a non-negative quartic form. Using the explicit form of the β function given by equation (24.15), for g_{ijkl} small, one finds

$$\lambda \frac{d}{d\lambda} G(\lambda) = \frac{3}{16\pi^2} \sum_{i,j,k,l,m,n} \varphi_i \varphi_j g_{ijmn}(\lambda) \varphi_k \varphi_l g_{klmn}(\lambda). \quad (24.22)$$

The right-hand side of the equation is a sum of squares. Therefore, $G(\lambda)$ is a positive increasing function. This is clearly incompatible with the assumed property that, at least for λ large enough, all functions $g_{ijkl}(\lambda)$ go to zero. The only other possibility is that all terms in the right-hand side vanish:

$$\sum_{i,j} \varphi_i \varphi_j g_{ijmn}(\lambda) = 0, \quad \forall m, n. \quad (24.23)$$

However, $G(\lambda)$ then vanishes identically. The argument is valid for all vectors φ . This implies that the field theory is free.

We conclude that no stable ϕ^4 -like theory is asymptotically free. Since no other UV fixed point is known in the bare theory, we would be tempted to conclude that all ϕ^4 -like theories are trivial, or correspond to first-order transitions without a continuum limit.

24.3 Theories with scalar bosons and Dirac fermions

We now consider four-dimensional, renormalizable QFTs involving only scalar bosons and Dirac fermions.

A simple example. We first discuss a theory with one scalar boson ϕ and one spin 1/2 Dirac fermion $(\psi, \bar{\psi})$ field, for which the RG β -functions can be inferred from expressions (20.30, 20.31) in Section 20.2.1. We parametrize the interaction terms as

$$\mathcal{S}_{\text{Int.}}(\psi, \bar{\psi}, \phi) = \int d^4x [-g\bar{\psi}(x)\psi(x)\phi(x) + \frac{1}{4!}u\phi^4(x)]. \quad (24.24)$$

The β -functions at one-loop order (order u^2, ug^2, g^4) and in four dimensions are then

$$\beta_{g^2} = \frac{5}{8\pi^2}g^4 + O(\text{two loops}), \quad \beta_u = \frac{1}{8\pi^2}(\frac{3}{2}u^2 + 4ug^2 - 24g^4) + O(\text{two loops}). \quad (24.25)$$

The fermions generate negative contributions to the ϕ^4 coupling RG function, which is no longer obviously positive. By contrast, β_{g^2} now is strictly positive. Therefore, for g^2 small, the running coupling constant $g^2(\lambda)$ increases for large λ . Since g^2 is positive, it grows in absolute value and the theory cannot be asymptotically free. However, there remains one case which must be examined separately: if g is of order u , then the two-loop contribution of order u^2g^2 is comparable to the one-loop term. This two-loop term, which comes entirely from Z_ϕ , the ϕ -field renormalization in the purely ϕ^4 theory, is given by equation (10.58), and has to be added in equation (20.21). The function β_{g^2} then becomes

$$\beta_{g^2} = \frac{5}{8\pi^2}g^4 + \frac{1}{48(8\pi^2)^2}u^2g^2 + O(g^6, g^2u^3). \quad (24.26)$$

Therefore, this additional term is also positive, and the conclusion remains unchanged.

The general case. The most general interaction, renormalizable in four dimensions, has the form

$$\begin{aligned} \mathcal{S}_{\text{Int.}}(\phi, \bar{\psi}, \psi) = & \int d^4x \left[\frac{1}{4!} \sum_{i,j,k,l} u_{ijkl} \phi_i(x) \phi_j(x) \phi_k(x) \phi_l(x) \right. \\ & \left. - \sum_{i,a,b} \bar{\psi}_a(x) (g_{ab}^i + i\gamma_5 h_{ab}^i) \phi_i(x) \psi_b(x) \right], \end{aligned} \quad (24.27)$$

in which $(g^i)_{ab}$ and $(h^i)_{ab}$ are Hermitian matrices (see Section A12). Since the diagrams contributing to the RG functions have been calculated in the one component case in Section 20.1.2, we just have to take into account the additional geometric factors. It is convenient to set

$$z_{ab}^i = g_{ab}^i + ih_{ab}^i. \quad (24.28)$$

A remark simplifies the calculation. If one calculates the Feynman diagrams in massless theory, then each time a $\bar{\psi}\psi\phi$ vertex commutes with a fermion propagator, the matrix \mathbf{z}^i is changed into its Hermitian conjugate $(\mathbf{z}^i)^\dagger$. In dimensional regularization, the renormalization constants then are

$$\mathbf{Z}_\psi = \mathbf{I} - \frac{1}{16\pi^2\varepsilon} \sum_i (\mathbf{z}^i)^\dagger \mathbf{z}^i + O(2 \text{ loops}), \quad (24.29)$$

$$(\mathbf{Z}_\phi)_{ij} = \delta_{ij} - \frac{1}{8\pi^2\varepsilon} \text{tr} [(\mathbf{z}^i)^\dagger \mathbf{z}^j + (\mathbf{z}^j)^\dagger \mathbf{z}^i] + O(2 \text{ loops}), \quad (24.30)$$

(\mathbf{I} is the identity matrix) while the divergent part of the $\bar{\psi}\psi\phi$ three-point function is proportional to $\mathbf{z}^j(\mathbf{z}^i)^\dagger \mathbf{z}^j$.

A short calculation then leads to the expression of the RG β_z -function:

$$16\pi^2\beta_z^i = \sum_j \left\{ \frac{1}{2} [\mathbf{z}^j(\mathbf{z}^j)^\dagger \mathbf{z}^i + \mathbf{z}^i(\mathbf{z}^j)^\dagger \mathbf{z}^j] + \text{tr}[(\mathbf{z}^i)^\dagger \mathbf{z}^j] \mathbf{z}^j + \text{tr}[\mathbf{z}^i(\mathbf{z}^j)^\dagger] \mathbf{z}^j \right. \\ \left. + 2\mathbf{z}^j(\mathbf{z}^i)^\dagger \mathbf{z}^j \right\}. \quad (24.31)$$

The flow equation for the quantity $\sum_i \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^i$ follows:

$$8\pi^2\lambda \frac{d}{d\lambda} \sum_i \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^i = \sum_{i,j} \left\{ \frac{1}{2} [\text{tr} \mathbf{z}^j(\mathbf{z}^j)^\dagger \mathbf{z}^i(\mathbf{z}^i)^\dagger + \text{tr}(\mathbf{z}^j)^\dagger \mathbf{z}^j(\mathbf{z}^i)^\dagger \mathbf{z}^i] + \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j \right. \\ \left. + \text{tr} \mathbf{z}^i(\mathbf{z}^j)^\dagger \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j + 2 \text{tr} \mathbf{z}^j(\mathbf{z}^i)^\dagger \mathbf{z}^j(\mathbf{z}^i)^\dagger \right\}. \quad (24.32)$$

The matrices $\sum_i \mathbf{z}^i(\mathbf{z}^i)^\dagger$ and $\sum_i (\mathbf{z}^i)^\dagger \mathbf{z}^i$ are positive, therefore, the two first terms in the right-hand side of equation (24.32), being of the form of the trace of the square of a positive matrix, are positive. The fourth term is larger than the third one, since

$$\sum_{i,j} \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j = \sum_{i,j} \text{Re} [\text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j] \\ \leq \sum_{i,j} [\text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j]^* [\text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j] = \sum_{i,j} \text{tr} \mathbf{z}^i(\mathbf{z}^j)^\dagger \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j. \quad (24.33)$$

It follows that

$$8\pi^2\lambda \frac{d}{d\lambda} \sum_i \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^i \geq 2 \sum_{i,j} [\text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^j + \text{tr} \mathbf{z}^j(\mathbf{z}^i)^\dagger \mathbf{z}^j(\mathbf{z}^i)^\dagger]. \quad (24.34)$$

In terms of the four-point vertex $M_{abcd} = z_{ab}^i z_{cd}^i$, it is simple to verify that the inequality (24.34) can be rewritten as

$$8\pi^2\lambda \frac{d}{d\lambda} \sum_i \text{tr}(\mathbf{z}^i)^\dagger \mathbf{z}^i \geq \sum_{a,b,c,d} (M_{abcd} + M_{adbc}) (M_{abcd}^* + M_{adbc}^*), \quad (24.35)$$

which proves that the right-hand side is positive. We conclude that $\sum_i \text{tr}(\mathbf{z}^i(\lambda))^\dagger \mathbf{z}^i(\lambda)$ is a positive increasing function for λ large, except, as before, if the two-loop contribution coming from the ϕ -field renormalization can cancel the one-loop term we have just considered. This two-loop term is proportional to $u_{iklm} u_{klmj} z_{ab}^j$, with a positive coefficient, as we have noted before. It contributes to equation (24.32) by a term proportional to $\sum_{a,b,i,j,k,l} u_{iklm} u_{klmj} z_{ab}^j (z_{ba}^i)^\dagger$, which is positive:

$$\sum_{a,b,i,j,k,l} u_{iklm} u_{klmj} z_{ab}^j (z_{ab}^i)^\dagger = \sum_{a,b,i,j,k,l} (u_{iklm} z_{ab}^{i*}) (u_{klmj} z_{ab}^j). \quad (24.36)$$

Therefore, the inequality remains valid and the conclusion is unchanged: a field theory containing only scalar bosons and fermions *cannot be asymptotically free* in four dimensions.

Since such theories do not seem to have other UV fixed points, if they have a continuum limit they are presumably ‘trivial’ in the sense of the triviality of the ϕ^4 field theory.

24.4 Gauge theories

In Section 21.12.1, we have calculated the RG β -function for QED, at leading order. In four dimensions, in a theory with n_F fermions and n_B bosons of charge e , the β -function at leading order is given by

$$\beta(e^2) = (4n_F + n_B) \frac{1}{3} \frac{e^4}{8\pi^2} + O(e^6). \quad (24.37)$$

Therefore, QED is IR free in four dimensions, like the ϕ^4 theory, and it is doubtful that it exists as a theory consistent at all scales. Of course, since the physical coupling constant is very small, the predictions of QED are not affected by this possible inconsistency whose effects are much too small. QED is presumably consistent up to enormous energies larger than the energy associated with the Planck scale.

By contrast, the β -function for matterless non-Abelian gauge theories, which we have calculated at leading order in Section 23.5.1, is given by

$$\beta(g^2) = -\frac{g^4}{8\pi^2} \frac{11}{6} C(G) + O(g^6), \quad (24.38)$$

in which $C(G)$ is the Casimir of the group G . Due to the negative sign, non-Abelian gauge theories corresponding to semi-simple groups are *asymptotically free* in four dimensions. What we have learned, in addition, in this chapter is quite remarkable: only theories possessing a non-Abelian gauge symmetry may share this property. In the terminology of critical phenomena, these theories are the only ones for which dimension 4 is the lower-critical dimension, in the same sense as dimension 2 is the lower-critical dimension for theories which have a global continuous symmetry (see Chapter 19).

Gauge theories with fermions. In Section 23.5.1, we have also calculated the contribution of fermions to the β -function. If the fermions belong to the representation R , and $T(R)$ is the trace of the square of the generators of the Lie algebra in the representation, defined by

$$\text{tr } \mathbf{t}^a \mathbf{t}^b = -\delta_{ab} T(R), \quad (24.39)$$

the β -function reads (equation (23.73)) [252],

$$\beta(g^2) = -\frac{1}{3}(11C(G) - 4T(R)) \frac{g^4}{8\pi^2} + O(g^6). \quad (24.40)$$

This result is sometimes expressed in terms of the equivalent of the fine structure constant $\alpha_s = g^2/4\pi$.

Before even calculating it, we know that the contribution of fermions is positive, as in the Abelian case. Therefore, a gauge theory with enough fermions is no longer asymptotically free. Actually, with the conventional normalization used for this problem,

$$C(G) = N \quad \text{for } SU(N), \quad T(R) = \frac{1}{2}N_F,$$

in which N_F is the number of *flavours*, that is, the number of fermion multiplets belonging to the fundamental representation of $SU(N)$. For the physical *colour group* $SU(3)$, the condition for asymptotic freedom is $N_F < 33/2$, a condition satisfied by the present SM.

If this condition is satisfied, the corresponding non-Abelian gauge theory could be consistent on all scales but, for the existence of low-energy physics, one condition must

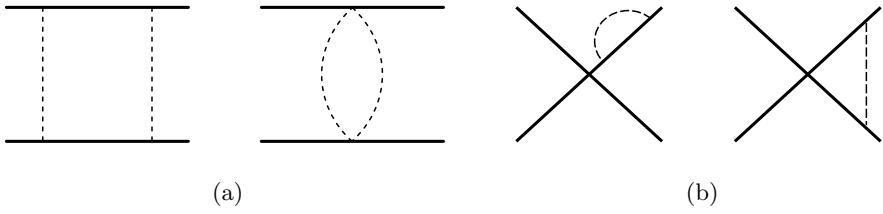


Fig. 24.1 Dashed lines correspond to gauge fields

be satisfied: the bare coupling constant must be close enough to the UV fixed point value, that is, small enough. Since, from experiments, we know that the effective coupling at low energy is large and, since the decrease in the RG flow is only logarithmic, this constraint is not very severe.

Gauge fields and scalar bosons. The situation is more complicated in the case of a theory also containing scalar bosons, like in the Higgs model. In general, the scalar fields have a tendency to destroy asymptotic freedom. First, they yield a positive contribution to the gauge coupling β -function which, as can be seen in equation (21.119), is $1/4$ of the fermion contribution corresponding to the same representation. However, more important, they introduce a ϕ^4 coupling which, as we know from the analysis of Section 24.2, by itself does not lead to asymptotic freedom. Let us assume, for simplicity, that we need only one ϕ^4 coupling constant u . The corresponding β -function has at one-loop order the form

$$\beta_u = au^2 + 2bug^2 + cg^4. \quad (24.41)$$

We know that a is positive. It is easy to verify that the contributions to c of the two diagrams (a) of Fig. 24.1 are positive. If the gauge group is $SU(N)$ and the scalar field belongs to the fundamental representation, a short calculation shows that b is negative (diagrams (b) of Fig. 24.1). From the form (23.77) of the β -function, we see that when g and u are small, if g^2 is much smaller than u or the converse, then u increases. Therefore, the only possibility for asymptotic freedom is that u and g^2 remain of the same order. It is then natural to introduce the ratio

$$v = u/g^2. \quad (24.42)$$

Setting

$$\beta_{g^2} = dg^4 + \dots, \quad (24.43)$$

and using equation (9.88), which relates the β -functions in different parametrizations:

$$\beta_j(g) \frac{\partial \tilde{g}_i}{\partial g_j} = \tilde{\beta}_i(\tilde{g}), \quad (24.44)$$

we can calculate

$$\beta_v = \beta_{g^2} \frac{\partial v}{\partial g^2} + \beta_u \frac{\partial v}{\partial u}. \quad (24.45)$$

We obtain

$$\beta_v = g^2 [av^2 + (2b - d)v + c]. \quad (24.46)$$

The theory can only be asymptotically free if the second degree polynomial in v has positive roots. This requires

$$(b - d/2)^2 > ac, \quad d - 2b > 0. \quad (24.47)$$

The two zeros v_1 and v_2 are then positive. Assuming

$$0 < v_1 < v_2 ,$$

we verify that v_1 is UV stable while v_2 is IR stable. Therefore, depending on the starting point in the (g, u) -plane, the effective coupling constants may be driven towards the origin. The conditions (24.47) are rather stringent. In particular, one can verify, by explicit calculation of the coefficients (a, b, c, d) , that it is impossible to add enough scalars belonging to the fundamental representation of $SU(N)$ to give masses to all vector bosons.

A general analysis of a system involving gauge fields, fermions and scalar bosons is rather complex. However, a few general results have been obtained:

(i) It is necessary to render the coefficient d of equation (24.43) small by adding enough fermions.

(ii) Generically, it is impossible to give a mass to all gauge fields through the Higgs mechanism without losing asymptotic freedom. However, theories exist in which one can find a manifold of zero measure in the space of coupling constants which leads to asymptotic freedom. This situation requires a fine tuning of the Yukawa-type interactions between scalars and fermions. Only in some supersymmetric theories is this fine tuning automatically realized and, therefore, natural.

The conclusion is that most probably the weak electromagnetic sector in particle physics is not asymptotically free. This is completely consistent with the observation that the gauge couplings are small, as one would expect in an IR free theory.

24.5 Applications: The theory of strong interactions

At the shortest distances presently experimentally accessible, strong interactions are well-described by a set of fermions, quarks, transforming under the fundamental representation of the group $SU(3)$, and interacting via $SU(3)$ gauge fields (gluons). The existence of six quarks (*six flavours*) has been established, organized by pairs and belonging to three generations. Moreover, experiments have shown that the number of light neutrinos (neutrinos with masses below 45 GeV) is exactly 3, a constraint on any additional generation. With six quarks, the theory is asymptotically free and asymptotic freedom alone leaves much room for additional quarks.

Asymptotic freedom has first emerged to provide an explanation for the experimental observations of point-like structure in deep inelastic scattering. One measures the inclusive cross section for the scattering of leptons (electrons, muons, or neutrinos) off nucleons to give leptons plus any number of unobserved hadrons at large momentum transfer. In this way, one probes the matrix elements between nucleon states of the product of two electromagnetic or weak currents near the light cone. We have shown in Chapter 11 that information about the behaviour near the light cone can be obtained from RG arguments. It is clear from the discussion of Section 17.2, which can be immediately transposed to UV stable fixed points, that this behaviour is characterized by logarithmic deviations from a free-field behaviour.

Asymptotic freedom thus provides a simple and elegant explanation to the results obtained in deep inelastic scattering experiments and, in particular, RG describes well the behaviour of structure functions. Here, we do not give here a detailed discussion of the theoretic predictions and refer the interested reader to the abundant literature. We rather examine, as an illustration, a somewhat simpler example: electron–positron annihilation.

Electron–positron annihilation [267, 79]. The total cross-section of annihilation of electron–positron pairs into hadrons is related, at leading order in the electromagnetic charge, to the expectation value of the product of two hadronic electromagnetic currents J_μ . Due to current conservation, in momentum space the expectation value takes the form

$$\langle \tilde{J}_\mu(q) \tilde{J}_\nu(-q) \rangle = (\delta_{\mu\nu} q^2 - q_\mu q_\nu) F(q^2). \quad (24.48)$$

The cross-section is proportional to $\text{Im } F(q^2)$ for $q^2 < 0$. At large values of q^2 , in an asymptotically free theory, the behaviour of $F(q^2)$ can be estimated from RG and perturbation theory. More precisely, one can only obtain the behaviour of $F(q^2)$ in the Euclidean region q^2 large and positive (corresponding to short-Euclidean distance). This behaviour extends to all directions in the complex q^2 -plane at the possible exception of the physical region $q^2 < 0$, because $F(q^2)$ has a cut on the negative real axis corresponding to the intermediate hadron states we are looking for. It is easy to construct analytic functions which decrease faster than any power in all directions of the complex plane and oscillate on the cut; an example is provided by the function $\exp(-\sqrt{|q^2|})$. Hereafter, we ignore this difficulty, but let us point out that the behaviour we obtain might only apply to some local average of $F(q^2)$, which suppresses such oscillations.

The electromagnetic current J_μ is exactly conserved and, thus, requires no renormalization and has no anomalous dimension (see Section A13.4). The two-point function of currents, which have dimension 3, has dimension 2. In equation (24.48), we have factorized a factor of dimension 2 and, thus, $F(q^2)$ is logarithmically divergent. The two-point function requires an additive renormalization, which induces an inhomogeneous term in the RG equations (see Section 9.7). Denoting by μ the renormalization scale, and associating a parameter m to the masses of the quarks, we can write the RG equation as

$$\left(\mu \frac{\partial}{\partial \mu} + \beta \frac{\partial}{\partial g} + \eta_m m \frac{\partial}{\partial m} \right) F(q^2, g, m, \mu) = B(g). \quad (24.49)$$

The solution of the analogous equation (17.22) is also discussed in Section 17.2. We denote by $C(g)$ a particular solution of the inhomogeneous equation:

$$\beta(g) \frac{\partial}{\partial g} C(g) = B(g). \quad (24.50)$$

Then, $F(q^2) - C$ satisfies a homogeneous equation, which can be solved in the usual way:

$$F((\lambda q)^2, g, m, \mu) - C(g) = F(q^2, g(\lambda), m(\lambda)/\lambda, \mu) - C(g(\lambda)). \quad (24.51)$$

Since the theory is asymptotically free, the effective coupling constant $g(\lambda)$ goes to zero for large λ . The left-hand side term $B(g)$ does not vanish for $g = 0$ (even in a free theory $F(q^2)$ is divergent),

$$B(g) = B_0 + B_1 g^2 + O(g^4), \quad (24.52)$$

and thus $C(g)$ behaves for g small like

$$C(g) = -\frac{B_0}{2\beta_2 g^2} + \left(\frac{B_1}{\beta_2} - \frac{B_0 \beta_3}{\beta_2^2} \right) \ln g + O(1). \quad (24.53)$$

Therefore, at large-scale λ , the right-hand side of equation (24.51) is dominated by the singular terms of $C(g(\lambda))$:

$$F(\lambda^2 q^2, g, m, \mu) = -C(g(\lambda)) + O(1). \quad (24.54)$$

From the definition of $g(\lambda)$ and equation (24.50), it follows that

$$\lambda \frac{d}{d\lambda} C(g(\lambda)) = B(g(\lambda)), \quad (24.55)$$

and, therefore, using the expansion of $g(\lambda)$ for λ large,

$$C(g(\lambda)) = B_0 \ln \lambda - \frac{B_1}{2\beta_2} \ln \ln \lambda + O(1). \quad (24.56)$$

The final result is

$$F(q^2) = -\frac{B_0}{2} \ln \left(\frac{q^2}{\mu^2} \right) + \frac{B_1}{2\beta_2} \ln \ln \left(\frac{q^2}{\mu^2} \right) + O(1) \quad \text{for } q^2 \rightarrow \infty. \quad (24.57)$$

It follows that

$$\text{Im}(F(q^2)) = -\frac{B_0}{2}\pi + \frac{B_1}{2\beta_2} \frac{\pi}{\ln(q^2/\mu^2)} + o\left(\frac{1}{\ln(q^2/\mu^2)}\right) \quad \text{for } q^2 M \rightarrow \infty. \quad (24.58)$$

One usually expresses this result in terms of the ratio $R(q^2)$ of the cross-section for $e_+ e_-$ into hadrons to $e_+ e_-$ into $\mu_+ \mu_-$. The latter cross-section is given by the imaginary part of the one-loop correction to the photon inverse propagator due to muons. The expression of the corresponding diagram has been given in Section 21.12.1 (equation (21.109)). It behaves for q^2 large as $\ln q^2$, its imaginary part is just a constant. In a free quark theory, the hadronic cross-section for q^2 is given in terms of the same diagram, the only difference being the coefficient that involves the charges Q_i of the quarks. Therefore, for q^2 large and negative, this ratio is simply the sum of the squares of the quark charges, the charge of the electron being taken as unit. In an asymptotically free theory, the result is the same at leading order. The gauge interaction between quarks leads to logarithmic corrections to the leading term:

$$R(q^2) = \sum_i Q_i^2 \left[1 - \frac{B_1}{B_0 \beta_2} \frac{1}{\ln(q^2/\mu^2)} + o\left(\frac{1}{\ln(q^2/\mu^2)}\right) \right]. \quad (24.59)$$

A two-loop calculation yields the coefficient B_1 . The final result is usually expressed in terms of the effective coupling constant $g(q/\mu)$ at scale q/μ :

$$R(q^2) = \sum_i Q_i^2 \left[1 + \frac{1}{4\pi^2} g^2(q/\mu) + O(g^4(q/\mu)) \right]. \quad (24.60)$$

For the $SU(3)$ colour group and for the six flavours already observed, the coefficient of the leading term is

$$\sum_i Q_i^2 = 3 \times \left(3 \times \left(\frac{2}{3} \right)^2 + 3 \times \left(\frac{1}{3} \right)^2 \right) = 5. \quad (24.61)$$

This result is valid when q is large compared to all quark masses. In fact, in experiments, one measures R for momenta large compared to some quark masses and comparable or smaller than others. If the masses are well separated, one expects, and indeed observes, that R varies slowly in intermediate regions and remains close to the value obtained by taking into account only the quarks of smaller masses. For example, below the top threshold, the largest relevant value is $11/3$.

25 Lattice gauge theories: Introduction

In this chapter, we describe a construction of gauge theories on a lattice [221, 241] (chosen hypercubic, for convenience) based on parallel transport, a notion that we have defined, in the continuum, in Sections 21.7 and 22.1: gauge fields are replaced by parallel transporters associated to links of the lattice, while gauge transformations correspond to independent group transformations on each lattice site (for early reviews see Ref. [268]).

Note that the problem of defining lattice fermions, which is discussed in Section 12.9, generalizes to gauge theories.

Lattice gauge theory also provides a non-perturbative regularization for the continuum gauge theories studied in Chapters 21–24 and 26: the low temperature or small coupling expansion of the lattice model is a regularized continuum perturbative expansion.

We study the properties of a lattice gauge theory from the point of view of phase transitions. We concentrate mainly on matterless lattice gauge theories (*i.e.*, without scalar bosons or fermions). This implies that we cannot investigate many properties, like the spectrum of a realistic theory like quantum chromodynamics (QCD) (for early calculations see Refs. [269]), where fermions are coupled through a gauged $SU(3)$ colour group, but we can still study one important question by lattice methods:

Does the theory generate *confinement*, that is, a force between charged particles increasing at large distances, so that heavy quarks in the fundamental representation cannot be separated? More generally, can one find charged (from the gauge group point of view) asymptotic states like massless vector particles in the theory?

Problems that we do not consider here, which can also be discussed in this framework, are the appearance of massive group singlet bound states in the spectrum (gluonium) and the question of a deconfinement transition at finite physical temperature [270].

Gauge theories have properties quite different from the ferromagnetic systems studied in Chapters 14–17. In particular, the absence of a local order parameter requires a study of the behaviour of a non-local quantity, a functional of loops generally called *Wilson's loop*, to distinguish between the confined and deconfined phases. Results can be obtained in the high temperature, or strong coupling limit and in the mean-field approximation.

However, we emphasize that, at present, the main physics results of lattice gauge theories, like the spectrum or the high temperature properties, are obtained, after dynamical quarks are added, from large-scale numerical simulations with realistic values of physical parameters (including small quark masses) [271]. One important remaining issue concerns the exploration of gauge theories at high density. Indeed, after the addition of a chemical potential, the measure in the field integral is no longer positive.

25.1 Gauge invariance on the lattice: Parallel transport

Hypocubic lattice. We consider a d -dimensional hypercubic lattice with a lattice spacing a , and sites at positions $x \equiv a(n_1, n_2, \dots, n_d)$, where $\mathbf{n} \in \mathbb{Z}^d$.

Scalar matter or spin models: The problem of gauge invariance. We consider a lattice model involving a set of complex scalar dynamic variables $\varphi(x)$, representing matter fields, located on sites of the lattice. On $\varphi(x)$ acts a unitary representation $\mathcal{D}(G)$ of a compact group G , and the lattice model is G -invariant.

We want to construct a gauge-invariant lattice model: a model is gauge invariant (local invariance) if it is invariant under independent group transformations on every lattice site. For the φ -measure of integration as well as for all the terms in the lattice action that depend only on one site, global invariance implies local invariance, as in the continuum. Problems only arise with terms which connect different lattice sites. If $\varphi(x)$ transforms under a representation $\mathcal{D}(G)$ of the group G as

$$\varphi_{\mathbf{g}}(x) = \mathbf{g}\varphi(x), \quad \mathbf{g} \in \mathcal{D}(G), \quad (25.1)$$

a term in the action containing products of the form $\varphi^*(x) \cdot \varphi(y)$, $x \neq y$, is invariant under global, but not local transformations:

$$\varphi^*(x) \cdot \varphi(y) \mapsto \varphi^*(x)\mathbf{g}^\dagger(x)\mathbf{g}(y)\varphi(y). \quad (25.2)$$

As we have shown in Sections 21.7 (equation (21.69)) and 22.1, to render it gauge invariant, one has to introduce new dynamic variables, which generate parallel transport between distinct sites x and y .

25.1.1 Parallel transport on the lattice

A parallel transporter is a unitary matrix $\mathbf{U}(C(x, y))$ belonging to the representation $\mathcal{D}(G)$, which depends on an oriented continuous curve C joining the sites x and y and, in a gauge transformation, transforming as (equation (22.2))

$$\mathbf{U}(C(x, y)) \mapsto \mathbf{g}(x)\mathbf{U}(C(x, y))\mathbf{g}^\dagger(y). \quad (25.3)$$

A gauge-invariant quantity on the lattice is then

$$\varphi^*(x)\mathbf{U}(C(x, y))\varphi(y). \quad (25.4)$$

A peculiarity of the lattice is that the curves can only join two lattice sites by following the *links*, the segments that connect adjacent sites, of the lattice.

Parallel transporter satisfy a composition rule: if $\mathbf{U}(C(x, y))$ and $\mathbf{U}(C(y, z))$ are two matrices transforming under the rule (25.3), the product of matrices

$$\mathbf{U}(C_1(x, y))\mathbf{U}(C_2(y, z)) = \mathbf{U}(C_1(x, y) \cup C_2(y, z)), \quad (25.5)$$

transforms like

$$\mathbf{U}(C_1(x, y) \cup C_2(y, z)) = \mathbf{g}(x)\mathbf{U}(C_1(x, y) \cup C_2(y, z))\mathbf{g}^\dagger(z). \quad (25.6)$$

In the continuum, in a local field theory, a parallel transport for any curve can be expressed in terms of transport for infinitesimal curves (equation (22.5)), and thus, in terms of a gauge field or connection element of the representation of the Lie algebra.

For a curve that follows the links of the lattice, the minimal curve is a link. As a consequence of the composition rule (25.5), one can thus take as dynamic variables elements of the group representation $\mathbf{U}(x, x + an_\mu)$, with $\mu = 1, \dots, d$, where n_μ is the unit vector in the μ direction, associated with parallel transport along oriented links of the lattice. They transform like,

$$\mathbf{U}(x, x + an_\mu) \mapsto \mathbf{g}(x)\mathbf{U}(x, x + an_\mu)\mathbf{g}^\dagger(x + an_\mu).$$

It is consistent with the transformation (25.3) to impose

$$\mathbf{U}(x, x + an_\mu) = \mathbf{U}^{-1}(x + an_\mu, x). \quad (25.7)$$

Then, parallel transport between the points x and y along a continuous oriented path $C(x, y)$ on the lattice can be generated by the product of link variables,

$$\mathbf{U}[C(x, y)] = \prod_{\text{links } \ell \in C(x, y)} \mathbf{U}_\ell,$$

where the product is ordered along the path.

25.2 The matterless gauge theory

We now define a matterless lattice gauge theory, and determine its formal continuum limit, as obtained from a low-temperature (in the sense of classical statistical physics), or strong-coupling expansion.

25.2.1 Action and partition function

We first need a gauge-invariant interaction for the gauge elements \mathbf{U} . It follows from the transformation (25.3) that only the traces of the products of \mathbf{U} 's on closed loops are gauge invariant. On a hypercubic lattice, the shortest loop is a square, called hereafter a *plaquette*. In what follows, we thus consider a pure gauge action of the form

$$\mathcal{S}(\mathbf{U}) = - \sum_{\text{plaquettes}} \text{tr } \mathbf{U}(x, y)\mathbf{U}(y, z)\mathbf{U}(z, w)\mathbf{U}(w, x), \quad (25.8)$$

where x, y, z , and w are four neighbouring lattice sites forming a square.

It is not surprising that the lattice action involves a product of parallel transporters along closed loops, since we know quite generally that the curvature tensor $\mathbf{F}_{\mu\nu}$, which appears in the pure gauge action of the continuum theory, is associated with infinitesimal transport around a closed loop. Note that each plaquette appears with both orientations in such a way that the sum is real when the group is unitary.

The partition function. We can then write the partition function corresponding to the action (25.8) as

$$\mathcal{Z} = \int \prod_{\text{links } \ell} d\mathbf{U}_\ell e^{-\beta_P \mathcal{S}(\mathbf{U})}, \quad (25.9)$$

in which β_P is the plaquette coupling constant (classically the inverse temperature). We integrate over \mathbf{U} with the invariant (de Haar) measure associated with the group G . In contrast to continuum gauge theories, the expression (25.9) is well-defined as long as the volume is finite, because the gauge group is compact. Therefore, gauge fixing is not required, and a *completely gauge-invariant* lattice model can be defined.

25.2.2 Low-temperature analysis

We first want to understand the precise connection between the lattice theory (25.9) and the continuum field theory. For this purpose, we investigate the lattice theory at low temperature, that is, for large-positive β_P . In this limit, the partition function is dominated by minimal-energy configurations.

Let us show that the minimum of the energy corresponds to matrices \mathbf{U} gauge transform of the identity. For link variables, we will use the symbolic notation $\mathbf{U}(x_1, x_2) = \mathbf{U}_{12}$.

We start from a first plaquette 1234. Without loss of generality, we can set

$$\mathbf{U}_{12} = \mathbf{g}_1^{-1}\mathbf{g}_2, \quad \text{with } \mathbf{g}_1, \mathbf{g}_2 \in \mathcal{D}(G). \quad (25.10)$$

The matrix \mathbf{g}_1 is arbitrary and \mathbf{g}_2 is calculated from \mathbf{U}_{12} and \mathbf{g}_1 . Then, we can also set

$$\mathbf{U}_{23} = \mathbf{g}_2^{-1}\mathbf{g}_3, \quad \mathbf{U}_{34} = \mathbf{g}_3^{-1}\mathbf{g}_4. \quad (25.11)$$

These relations define first \mathbf{g}_3 , then \mathbf{g}_4 . The minimum of the action is obtained when the real part of all traces is maximum, that is, when the products of the group elements on a plaquette are 1. (The trace of a unitary matrix \mathbf{U} is maximum when all its eigenvalues are 1.)

In particular,

$$\mathbf{U}_{12}\mathbf{U}_{23}\mathbf{U}_{34}\mathbf{U}_{41} = \mathbf{1}, \quad (25.12)$$

which yields

$$\mathbf{U}_{41} = \mathbf{g}_4^{-1}\mathbf{g}_1. \quad (25.13)$$

If we now consider an adjacent plaquette, the argument can be repeated for all links but one, which has already been fixed. In this way, one can show that the minimum of the action is a pure gauge. Thus, for large coupling constant β_P , all group elements are constrained to stay, up to a gauge transformation, close to the identity (in a finite volume with consistent boundary conditions). From the analysis, we learn that the minimum of the potential is highly degenerate at low temperature, since it is parametrized by a gauge transformation, which corresponds to a finite number of degrees of freedom per site. This unusual property of lattice gauge theories corresponds to the property that the gauge action in classical mechanics determines the motion only up to a gauge transformation. Therefore, to perform a low-temperature expansion, it becomes necessary to fix the gauge in order to sum over all minima.

Low-temperature expansion. We choose a gauge such that the minimum of the energy corresponds to all matrices $\mathbf{U} = \mathbf{1}$. At low temperature, the matrices \mathbf{U} are then close to the identity and we set (see equation (22.5))

$$\mathbf{U}(x, x + an_\mu) = 1 - a\mathbf{A}_\mu(x) + O(a^2), \quad (25.14)$$

in which a is the lattice spacing, x the position on the lattice, and n_μ the unit vector in the direction μ . From the discussion of Section 22.1, we know that the matrix $\mathbf{A}_\mu(x)$ is the connection or gauge field. We have already shown that the transformation (25.3) of the parallel transporter implies for $\mathbf{A}_\mu(x)$ at leading order in the lattice spacing:

$$\mathbf{A}_\mu(x) \mapsto \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x) + \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x),$$

which is the usual gauge transformation.

We now expand the lattice action for small fields. To simplify calculations, we parametrize the unitary matrix \mathbf{U} associated with the link $(x, x + an_\mu)$ in terms of the antisymmetric matrix $\mathbf{A}_\mu(x)$ as

$$\ln \mathbf{U}(x + an_\mu, x) = -a\mathbf{A}_\mu(x + \frac{1}{2}an_\mu) + O(a^3). \quad (25.15)$$

We verify below that we need \mathbf{U} up to order a^2 . With the parametrization (25.15), equation (25.7) implies that the term of order a^2 vanishes. We define the antisymmetric matrix $\mathbf{F}_{\mu\nu}(x)$ by

$$\begin{aligned} e^{-a^2\mathbf{F}_{\mu\nu}(x)} &= \mathbf{U}(x, x + an_\nu)\mathbf{U}(x + an_\nu, x + a(n_\mu + n_\nu)) \\ &\times \mathbf{U}(x + a(n_\mu + n_\nu), x + an_\mu)\mathbf{U}(x + an_\mu, x). \end{aligned} \quad (25.16)$$

To calculate $\mathbf{F}_{\mu\nu}(x)$, we introduce the expansion (25.15) and use repeatedly the Baker–Hausdorff formula:

$$\ln(e^A e^B) = A + B + \frac{1}{2}[A, B] + \dots \quad (25.17)$$

Applied to the product of several factors, it takes the form

$$\ln(e^{A_1} e^{A_2} \cdots e^{A_n}) = \sum_i A_i + \frac{1}{2} \sum_{i < j} [A_i, A_j] + \cdots \quad (25.18)$$

and, therefore,

$$a^2 \mathbf{F}_{\mu\nu}(x) = a [\mathbf{A}_\mu(x + \frac{1}{2}an_\mu) + \mathbf{A}_\nu(x + an_\mu + \frac{1}{2}an_\nu) - \mathbf{A}_\mu(x + an_\nu + \frac{1}{2}an_\mu) \\ - \mathbf{A}_\nu(x + \frac{1}{2}an_\nu)] + a^2 [\mathbf{A}_\mu(x), \mathbf{A}_\nu(x)] + O(a^3). \quad (25.19)$$

At leading order, the curvature tensor is recovered:

$$\mathbf{F}_{\mu\nu}(x) = \partial_\mu \mathbf{A}_\nu(x) - \partial_\nu \mathbf{A}_\mu(x) + [\mathbf{A}_\mu(x), \mathbf{A}_\nu(x)] + O(a). \quad (25.20)$$

We obtain one term in the plaquette action by taking the trace of expression (25.16). Since $\mathbf{F}_{\mu\nu}$ is an antisymmetric matrix, $\text{tr } \mathbf{F}_{\mu\nu}$ vanishes. Thus,

$$\text{tr } e^{-a^2 \mathbf{F}_{\mu\nu}(x)} = \text{tr } \mathbf{1} + a^4 \text{tr } \mathbf{F}_{\mu\nu}^2(x) + O(a^6). \quad (25.21)$$

This result shows that the leading term of the small field expansion of the plaquette action (25.8) is the standard gauge action studied in Chapter 22. The relation between β_P and the bare coupling constant e_0 of continuum gauge theories is thus,

$$a^4 \beta_P \sim e_0^{-2}. \quad (25.22)$$

As we have anticipated in Chapters 21 and 22, we conclude that the low-temperature expansion, in a fixed gauge, of lattice gauge theories indeed provides a lattice regularization of continuum gauge theories. We have only discussed the pure gauge action here, but the inclusion of matter fields is simple. Higher order terms in the small field expansion yield additional interactions needed to maintain gauge invariance on the lattice. This is not surprising: we have already shown that the gauge-invariant extension of Pauli–Villars’s regularization also introduces additional interactions.

25.3 Wilson’s loop and confinement

In Section 23.5.1, we have calculated the renormalization group (RG) β -functions for non-Abelian gauge theories, and shown that pure gauge theories are *asymptotically free* in four dimensions, because the origin in the coupling constant space is an ultraviolet (UV) fixed point. This property also implies that the effective interaction increases at large distance. Therefore, as in the case of the two-dimensional non-linear σ -model, the spectrum of a non-Abelian gauge theory cannot be determined from perturbation theory. To explain the non-observation of free quarks, it has been conjectured, quite early, that the spectrum of the symmetric phase contains only neutral states, that is, states that are singlets for the group transformations.

When discussing phase transitions, it is quite convenient to identify a local order parameter, that is, a local observable whose expectation value distinguishes between the phases. However, in gauge theories, no local order parameter can be found to distinguish between the quantum electrodynamics (QED) phase of Abelian gauge theories, in which charge states can be produced, and the so-called *confined* phase [272]. This property follows from the simple remark that physical observables correspond to gauge-invariant operators, which are neutral by construction. Moreover, we have seen in the study of continuum gauge theories (Chapters 21 and 22) that the only gauge-independent quantities corresponding to non-gauge-invariant operators are the S -matrix elements. Since it is very difficult to determine S -matrix elements beyond perturbation theory, Wilson has suggested to study, in pure gauge theories, a gauge-invariant non-local quantity, the energy of the vacuum in the presence of largely separated static charges.

25.3.1 Wilson's loop in continuum Abelian gauge theories

We first study the vacuum energy in pure Abelian gauge theories, in the continuum, because all calculations can be done explicitly. In continuum field theory, in order to calculate the average energy, it is necessary to introduce the gauge Hamiltonian and, therefore, convenient to work in the temporal gauge. We have constructed a wave function for two static point-like charges, in the temporal gauge, in Section 21.5.2 (equation (21.51)):

$$\psi(A) = \exp \left[-ie \oint_{C_0} \mathbf{A}(s) \cdot d\mathbf{s} \right], \quad (25.23)$$

in which the charges are located at both ends of the curve C_0 .

We consider the expectation value in the quantum state $\psi(A)$,

$$W(C_0) = \langle \psi | e^{-HT} | \psi \rangle, \quad (25.24)$$

in which H is the gauge Hamiltonian in the temporal gauge, and T a time. For large time T , we obtain the vacuum energy $E(C_0)$ in the presence of static electric charges:

$$W(C_0) \underset{T \rightarrow \infty}{\sim} e^{-TE(C_0)}. \quad (25.25)$$

If the charges are separated by a distance R , we expect E to depend only on R and not on C_0 .

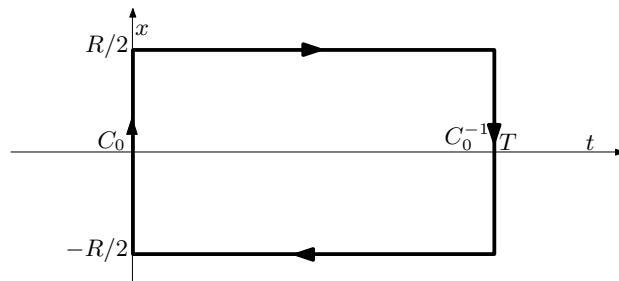


Fig. 25.1 The loop C

The loop functional $W(C_0)$ is given by the field integral

$$W(C_0) = \left\langle \exp \left[-ie \oint_{C'_0} \mathbf{A}(s) \cdot d\mathbf{s} \right] \right\rangle,$$

where C'_0 , which is now defined in space and time, is the union of two curves, which coincide with C_0 at time 0, and with C_0^{-1} at time T , respectively. The expectation value here means the average over gauge field configurations.

Since, in the temporal gauge, the time component of A_μ vanishes, we can add two straight lines in the time direction to C'_0 , which join the ends of the curves $C_0(t = 0)$ and $C_0(t = T)$. $W(C_0)$ then becomes a functional of a closed loop C (see Fig. 25.1):

$$W(C_0) \equiv W(C) = \left\langle \exp \left[-ie \oint_C \mathbf{A}(s) \cdot d\mathbf{s} \right] \right\rangle. \quad (25.26)$$

The representation (25.26) is now explicitly gauge invariant, since it is the expectation value of a parallel transporter corresponding to a closed loop in space-time.

The question of confinement is related to the behaviour of the energy E when the separation R between charges becomes large. In a pure Abelian gauge theory, in the continuum, which is a free-field theory, the expression (25.26) can be evaluated explicitly. To simplify calculations, we also take a straight line for C_0 and use Feynman's gauge. The quantity $W(C)$ then is given by

$$W(C) = \int [dA_\mu] \exp \left[-\mathcal{S}(A) + \int d^d x \mathbf{J}(x) \cdot \mathbf{A}(x) \right], \quad (25.27)$$

with

$$\mathcal{S}(A) = \frac{1}{2} \int d^d x \left[\sum_{\mu, \nu} \partial_\mu A_\nu(x) \right]^2, \quad J_\mu(x) = -ie \oint_C \delta^{(d)}(x - s) ds_\mu.$$

The result is

$$\ln W(C) = -\frac{\Gamma(d/2 - 1)}{8\pi^{d/2}} e^2 \oint_{C \times C} ds_1 \cdot ds_2 |\mathbf{s}_1 - \mathbf{s}_2|^{2-d}. \quad (25.28)$$

The integral in the right-hand side exhibits a short-distance singularity, and a short-distance cut-off thus is required. Moreover, to normalize the right-hand side of equation (25.28), we divide it by $W(C)$ taken for $R = a$, a being a fixed distance. We now write the integrals more explicitly:

$$\begin{aligned} \oint_{C \times C} \frac{ds_1 \cdot ds_2}{2|\mathbf{s}_1 - \mathbf{s}_2|^{d-2}} &= \int_0^T |u - t|^{2-d} du dt + \int_0^R |x - y|^{2-d} dx dy \\ &- \int_0^R [(x - y)^2 + T^2]^{1-d/2} dx dy - \int_0^T [(t - u)^2 + R^2]^{1-d/2} dt du. \end{aligned} \quad (25.29)$$

The first term in the right-hand side is cancelled by the normalization. The second term is independent of T and, therefore, negligible for large T . It is actually related to the scalar product of the wave function $\psi(A)$ and the ground state eigenfunction. For $d > 2$, the third term decreases with T , and only the last term increases with T . One finds,

$$\begin{aligned} \int_0^T \left\{ [(t - u)^2 + R^2]^{1-d/2} - [(t - u)^2 + a^2]^{1-d/2} \right\} dt du \\ \sim \sqrt{\pi} \frac{\Gamma((d-3)/2)}{\Gamma(d/2 - 1)} (R^{3-d} - a^{3-d}) T. \end{aligned} \quad (25.30)$$

Therefore, the vacuum energy $E(R)$ in presence of the static electric charges has the form

$$E(R) - E(a) = \frac{e^2}{4\pi^{(d-1)/2}} \Gamma((d-3)/2) (a^{3-d} - R^{3-d}). \quad (25.31)$$

We recognize the Coulomb potential between two charges.

For $d \leq 3$, the energy of the vacuum increases without bound when the charges are separated, and free charges cannot exist.

For $d = 3$, the potential increases logarithmically.

For $d = 2$, the Coulomb potential increases linearly with distance.

In more general situations, the method that we have used previously to determine the energy is complicated, because we have to take the large- T limit first, and then evaluate the large- R behaviour. It is more convenient to take a square loop, $T = R$, and evaluate the large- R behaviour of $W(C)$. Here, we obtain

$$\begin{aligned} & \ln W[C(R)] - \ln W[C(a)] \\ &= \frac{1}{2\pi^{d/2}} \Gamma(d/2 - 1) e^2 \left\{ \int_0^R [(u-t)^2 + R^2]^{1-d/2} du dt \right. \\ &\quad \left. - \int_0^a [(u-t)^2 + a^2]^{1-d/2} du dt - \int_a^R |u-t|^{2-d} du dt \right\}. \end{aligned} \quad (25.32)$$

For $d > 3$, dimensions in which the Coulomb potential decreases, the right-hand side is dominated by terms which correspond to the region $|s_1 - s_2| \ll R$ in equation (25.28):

$$\ln W[C(R)] - \ln W[C(a)] \sim \text{const.} \times R. \quad (25.33)$$

This is called the *perimeter law*, since $\ln W(C)$ is proportional to the length of C and is, therefore, relevant to the $d = 4$ Coulomb phase.

By contrast, for $d \leq 3$, $\ln W(C)$ increases as R^{4-d} . The reason is that two charges separated on C by a distance of order R , feel like a potential of order R^{d-3} .

In particular for $d = 2$, $\ln W(C)$ increases like R^2 , that is, like the area of the surface enclosed by C : this is the *area law* expected in confining situations.

25.3.2 Non-Abelian gauge theories

When discussing the Abelian example in Section 22.3, we have noticed that, in the temporal gauge, the wave function corresponding to two opposite point-like static charges is also related to a parallel transporter along a curve joining the charges. In a non-Abelian gauge theory the same arguments apply: the expectation value of the operator e^{-TH} in the corresponding state is given by the average, in the sense of the field integral, of the parallel transporter along a closed loop:

$$W(C) = \left\langle P \exp \left[-i \oint_C \mathbf{A}(s) \cdot d\mathbf{s} \right] \right\rangle, \quad (25.34)$$

in which the symbol P means path ordering, since the matrices $\mathbf{A}_\mu(s)$ at different points do not commute.

If we calculate $W(C)$ in perturbation theory, at leading order, we find the same results as in the Abelian case. However, we know from RG group arguments, that we cannot trust perturbation theory at large distances. Therefore, to obtain a qualitative understanding of the phase structure, we use the lattice model to first calculate $W(C)$ in the large-coupling or high-temperature limit $\beta_P \rightarrow 0$.

Strong coupling expansion for Wilson's loop. Here, we assume that the gauge group has a *non-trivial centre*. We take the explicit example of gauge-group elements on the lattice belonging to the fundamental representation of $SU(N)$ (whose centre is \mathbb{Z}_N , with elements roots z of unity ($z^N = 1$) multiplied by the identity).

We calculate $W(C)$ by expanding the integrand in expression (25.9) in powers of β_P . For the loop C , we choose, for simplicity a rectangle, although the generalization to other contours is simple.

Any non-vanishing contribution must be invariant by the change of variables $U_\ell \mapsto z_\ell U_\ell$, where z_ℓ belongs to the centre. Let us consider one link belonging to the loop and multiply the corresponding link variable $\mathbf{U}(x, x + an_\mu)$ by z_0 . We now multiply all link variables $\mathbf{U}(x + y, x + y + an_\mu)$, which are obtained by a translation y in the hyperplane perpendicular to n_μ , by z_y . Another link belonging to the loop belongs to the set but with opposite orientation. Plaquettes involving such variables involve them in pairs. For a result to be invariant and thus non-vanishing, we need that the number of times each link variable appears in the direction n_μ minus the number of times it appears in the direction $-n_\mu$ vanishes (mod N). Thus, we start adding plaquettes to satisfy this condition at point x . However, the addition of one plaquette does not change the total difference between the numbers of links in the $+n_\mu$ and $-n_\mu$ directions. Therefore, always at least one condition remains unsatisfied until the plaquettes reach the other side of the loop. We can then repeat the arguments for the remaining links of the loop and the new non-integrated remaining links of the plaquettes. The number of required plaquette variables to obtain a non-vanishing result, is at least equal to the area of the rectangle, the minimal-area surface having the loop as a boundary. We can then perform the integrations, which are just factorized group integrations.

In this way, we obtain a contribution to $W(C)$ proportional to $(\beta_P)^A$, in which A is the number of plaquettes. The largest contribution corresponds to plaquettes covering the minimal-area surfaces bounded by the loop. Indeed, it is obtained by covering the rectangle with plaquettes in such a way that each link variable appears only twice in either orientation. For a rectangular loop $R \times T$, we find

$$W(C) \sim e^{RT \ln \beta_P}. \quad (25.35)$$

This results indicates that the potential between static charges is linearly rising at large distance. Static charges creating the loop cannot simply be screened by the gauge field, otherwise we would again find a perimeter law.

Remarks.

(i) If the centre is trivial, it is possible to form a tube along the loop and this implies a perimeter law. For example, in the case of the $SO(3)$ group, in the decomposition of a product of two spin 1 representations, we again find a spin 1, which can be coupled to a third spin 1 to form a scalar. Thus, two plaquettes can be glued to the same link of the loop without constraint on the orientation of the plaquette.

(ii) The asymptotic form (25.35) is also valid for the Abelian $U(1)$ lattice gauge theory. Therefore, in four dimensions, Wilson's loop has a perimeter law at any order in the weak coupling expansion, and an area law at large coupling. We expect a phase transition between a low coupling Coulomb phase, described by a free-field theory, and a strong coupling confined phase. This phase transition has been observed in numerical simulations. It seems to be first order and, therefore, does not lead to a new continuum field theory. The existence of the transition is related to the compact nature of the $U(1)$ group, which is only relevant on the lattice (lattice QED based on group elements is also called compact QED) [273]. Defects in which the group element varies by a multiple of 2π around a plaquette govern the dynamics of the transition. They correspond in the continuum to magnetic monopoles.

In four dimensions, monopole loops yield, for dimensional reasons, logarithmic contributions to the action, a situation reminiscent of the two-dimensional Coulomb gas discussed in Chapter 31. The separation of vortices in the Kosterlitz–Thouless phase transition is replaced here by the separation of magnetic monopoles.

The string tension. The coefficient in front of the area

$$\sigma(\beta_P) \underset{\beta_P \rightarrow 0}{\sim} -\ln \beta_P \quad (25.36)$$

is called the *string tension*. If no phase transition occurs when β_P varies from 0 to infinity, the gauge theory leads to confinement. In this case, the behaviour of the string tension for β_P small is predicted by RG arguments. Since σ has the dimension of a mass squared, one finds

$$\sigma(e_0) \sim (e_0^2)^{-\beta_2/\beta_3} \exp(-1/\beta_2 e_0^2). \quad (25.37)$$

in which e_0^2 is related to β_P by equation (25.22) and β_2, β_3 are the two first coefficients of the RG β -function given in equation (23.75). A physical quantity relevant to the continuum limit can then be obtained by dividing $\sqrt{\sigma}$ by its asymptotic behaviour. If one defines the scale momentum by

$$\Lambda_L = a^{-1} (\beta_2 e_0^2)^{-\beta_3/2\beta_2} \exp(-1/2\beta_2 e_0^2), \quad (25.38)$$

then $\Lambda_L / \sqrt{\sigma}$ has a continuum limit. When one calculates σ by non-perturbative lattice methods, the verification of the scaling behaviour (25.37) indicates that the result is relevant to the continuum field theory and not simply a lattice artifact.

It is possible to systematically expand σ in powers of β_P . The possibility of investigating whether confinement is realized in the continuum limit, depends on the possibility of analytically continuing the strong coupling expansion up to the origin. Unfortunately, theoretical arguments lead one to believe that, independently of the group, the string tension is affected by a singularity associated with the roughening transition, a transition which, however, is not related to bulk properties. At strong coupling, the contributions to the string tension come from smooth surfaces. When e_0^2 decreases (β_P increases), one passes through a critical point e_{0R}^2 , after which the relevant surfaces become rough. At the singular coupling e_{0R}^2 , the string tension does not vanish, but has a weak singularity. Still, at this point, the strong coupling expansion diverges. Therefore, it is impossible to extrapolate to arbitrarily small coupling. The usefulness of the strong coupling expansion then depends on the position of the roughening transition with respect of the onset of weak coupling behaviour. Notice that, numerically in the neighbourhood of the roughening transition, rotational symmetry is approximately restored (at least at large enough distance).

One can also calculate other quantities which are associated to bulk properties and are, therefore, not affected by roughening singularities, such as the free energy (the connected vacuum amplitude) or the plaquette–plaquette correlation function. However, even for these quantities the extrapolation is not easy, because the transition between strong and weak coupling behaviours is in general very sharp. This is confirmed by results coming from Monte Carlo simulations, and is interpreted as indicating the presence of singularities in the complex β_P plane close to the real axis.

Flux tube. We note that the potential between static charges in the confined phase is linearly increasing, in the same way as the Coulomb potential in one space dimension. This leads to the following physical picture: in QED, the gauge field responsible of the potential has no charge and propagates essentially like a free-field isotropically in all space directions. Conservation of flux on a sphere then yields the R^{2-d} force between the charges. However, in the non-Abelian case, the attractive force between the gauge particles generates a flux tube between static charges in such a way that the force remains the same as in one space dimension [274].

Gauge symmetry breaking: Elitzur's theorem. Let us add a simple comment about the absence of a local order parameter in gauge theories. We have seen that, in the temporal gauge, the ground state is invariant under space-dependent gauge transformations. This property is incompatible with the existence of a local order parameter, which is necessarily non-gauge invariant. Therefore, the question is: can a phase transition on the lattice lead to a spontaneous breaking of gauge invariance? The answer to this question can be obtained by generalizing the arguments developed for ordinary symmetries in Sections 14.2 and 14.3 [272]. We consider the transition probability at low temperature between two states, concentrated one around the minimal-energy configuration $\mathbf{A}_\mu = 0$ and the other one around a pure non-trivial gauge. If the gauge function is only different from 0 in a finite space volume, the cost in energy is the same as in a one-dimensional system and, therefore, the transition probability always remains finite, independently of the number of space dimensions. Therefore, the quantum ground state is gauge invariant. Note that this argument does not apply to gauge transformations that do not vanish at large distances. Therefore, it does not forbid a spontaneous breaking of the global symmetry associated with the gauge group.

25.4 Mean-field approximation

We have shown that the pure gauge lattice model yields a continuum gauge theory at low temperature or coupling. In the continuum model, in perturbation theory, a separation of charges at large distances is possible. By contrast, at high temperature, charges are confined in the lattice model.

Therefore, it is necessary to investigate the possibility of phase transitions in lattice gauge theories. In the case of spin models, the mean-field approximation gives a semi-quantitative understanding of the phase structure, at least for $d > 2$. Therefore, it is natural to also study gauge theories in the mean-field approximation [275].

The general mean-field formalism is described in Section A14.2. We introduce two sets of real matrices, ϕ_ℓ and H_ℓ , in which the index ℓ corresponds to links. Then, we express the partition function

$$\mathcal{Z} = \int \prod_{\text{links } \ell} dU_\ell \exp [-\beta_P \mathcal{S}(U)], \quad (25.39)$$

in which $\mathcal{S}(U)$ is the lattice action (25.8), as

$$\mathcal{Z} = \int \prod_\ell d\phi_\ell dH_\ell dU_\ell \exp \left[-\beta_P \mathcal{S}(\phi) + \sum_{\text{links } \ell'} \text{tr } H_{\ell'} (\phi_{\ell'} - U_{\ell'}) \right]. \quad (25.40)$$

The introduction of the variables ϕ makes it possible to express the action in terms of an average link variable.

Since the average of a unitary matrix is not unitary, we have defined ϕ as an arbitrary complex matrix. The variables H_ℓ represent directly, at leading order, the mean field that approximates the effect of the plaquette interaction.

The integral over the matrices U factorizes into a product of integrals over the different link variables of the form,

$$\int dU e^{-\text{tr} HU} = e^{-\rho(H)}, \quad (25.41)$$

in which $\rho(H)$ is thus a $G \times G$ group-invariant function of H (H transforming under independent right and left multiplications).

The partition function becomes

$$\mathcal{Z} = \int \prod_\ell dH_\ell d\phi_\ell \exp \left\{ - \left[\beta_P \mathcal{S}(\phi) + \sum_{\ell'} (\rho(H_{\ell'}) - \text{tr } H_{\ell'} \phi_{\ell'}) \right] \right\}. \quad (25.42)$$

We look for saddle points in the variables H and ϕ . Since H and ϕ are general real or complex matrices, we expect to find many saddle points. However, both for simplicity and symmetry reasons, we look for solutions with H_ℓ and ϕ_ℓ constant on the lattice, and multiple of the identity (up to a gauge transformation):

$$\phi_\ell = \varphi I, \quad H_\ell = hI, \quad (25.43)$$

in which I is the identity matrix. Denoting by $\mathcal{S}(\varphi, h)$ the average lattice action per link, one finds

$$\mathcal{S}(h, \varphi) = \text{tr } I \left[-\frac{1}{2}(d-1)\beta_P \varphi^4 + V(h) - h\varphi \right], \quad (25.44)$$

in which $V(h)$ is defined by

$$V(h) = \frac{\rho(hI)}{\text{tr } I}, \quad \text{with } V(h) = -\frac{1}{4}h^2 + O(h^4), \text{ for } SU(2).$$

The saddle point equations are

$$\varphi = V'(h), \quad h = -2(d-1)\beta_P \varphi^3. \quad (25.45)$$

Eliminating φ , one obtains

$$h = -2(d-1)\beta_P [V'(h)]^3. \quad (25.46)$$

For h small, $V'(h)$ is at least linear in h (as in $SU(2)$). We realize immediately the essential difference with the spin models we had considered so far. The right-hand side of equation (25.46) is at least cubic in h , instead of being linear. Thus, the equation has never a non-trivial solution arbitrarily close to 0. For β_P small, there only exists the trivial solution $h = 0$ that, according to the strong coupling or high temperature analysis, corresponds to the confined phase, in which Wilson's loop follows an area law. For a critical value $\beta_{P,c}$, h jumps from 0 to a finite value, indicating a *first-order* phase transition. We recall that at a first-order transition the correlation length, at least above the transition, remains finite. Therefore, the neighbourhood of the transition temperature does not define a continuum field theory, in contrast with, for example, the non-linear σ -model.

Above $\beta_{P,c}$, the expectation value of Wilson's loop is given by

$$W(C) = \left\langle \text{tr} \prod_{\text{all links } \ell \in C} \phi_\ell \right\rangle \sim \phi^{P(C)}, \quad (25.47)$$

in which $P(C)$ is the perimeter of the loop. Therefore, Wilson's loop follows a perimeter law, and the phase is deconfined. Above $\beta_{P,c}$, the system is in the low-temperature phase, which can be described by a continuum field theory and perturbation theory.

Phase transition. As we have shown in Section A14.2, the mean-field approximation is valid in high dimensions. Continuum field theory tells us that the zero temperature ($\beta_P = \infty$) is an infrared stable fixed point for $d > 4$. Thus, the mean-field result can apply only for $d > 4$.

However, we would naively expect a second order phase transition in $4 + \varepsilon$ dimensions with a critical temperature of order ε , or $\beta_P \sim 1/\varepsilon$, in analogy with the non-linear σ -model. The open question is whether in any integer dimension $d > 4$ the transition is really second order.

For $d \leq 4$, the zero temperature is a UV fixed point. The simplest consistent scheme is one in which the critical temperature vanishes, and the model always remains in the confined high-temperature phase. For gauge theories, the dimension $d = 4$ plays the role of the dimension $d = 2$ for the non-linear σ -model. The large momentum behaviour of correlation functions can be determined from perturbation theory, but no analytical method yields directly their low-momentum behaviour and, therefore, for example, the spectrum of the theory. The only other analytical piece of information available is the small coupling expansion in a finite volume of the eigenstates of the quantum Hamiltonian, which one can try to extrapolate by numerical methods towards the infinite volume limit (see Chapter 32). However, again there is numerical evidence of a sharp transition between the finite volume and infinite volume results, making the extrapolation difficult. The most promising quantities seem to be ratios of masses. This lack of reliable analytical methods explains the popularity of numerical simulations, based on stochastic methods of Monte Carlo type, in this problem.

Monte Carlo methods. We do not describe here the numerical methods that are used in lattice gauge theories [276], and have become increasingly sophisticated. They are generalizations of the methods that we briefly explain in the example of simpler systems in Section A34.2.

In pure gauge theories, the existence of phase transitions has been investigated for many lattice actions. For the gauge group $SU(3)$, relevant to the physics of strong interactions, the string tension has been carefully measured, the plaquette-plaquette correlation function has been studied to determine the mass of low-lying gluonium states. Finally, calculations have been performed at finite physical temperature, that is, on a 3+1 dimensional lattice in the limit in which the size of the lattice remains finite in one dimension, this size being related to the temperature. In this way, the temperature of a deconfinement transition (most likely only a crossover, after the addition of quarks [277]) has been determined.

25.4.1 Fermions

We have discussed in Section 12.9 the construction of a lattice action for fermions, in particular, the fermion doubling problem, and its possible solutions. We do not repeat the discussion here.

A gauge-invariant lattice action for fermions has the form

$$\mathcal{S}(\bar{\psi}, \psi, \mathbf{U}) = \sum_{x, \mu} \bar{\psi}(x + an_\mu) \gamma_\mu \mathbf{U}(x + an_\mu, x) \psi(x), \quad (25.48)$$

where n_μ is the unit vector in μ direction and a the lattice spacing.

Fermions in numerical simulations. One important qualitative feature of strong interaction physics is the approximate spontaneous breaking of chiral symmetry (see Section 13.6). However, we have shown in Section 8.7 that non-trivial problems arise when one tries to construct a chiral-invariant lattice action.

One has the choice only between choosing an action that is not explicitly chiral symmetric, and in which one tries to restore chiral symmetry by adjusting the fermion mass term (Wilson's fermions), and choosing a chiral-symmetric action with too many fermions (staggered or Kogut–Susskind fermions), or, as it has been later discovered, choosing various Dirac operators satisfying the Ginsparg–Wilson relation [278, 279] (see Section 12.9.1).

In the latter solution, several implementations can be interpreted as adding an extra space dimension for the fermions. Although this method theoretically solves the doubling problem, from the point of view of numerical simulations the calculation of the corresponding lattice operator in a gauge background is extremely time consuming.

Indeed, independently of the choice of the lattice Dirac operator, an important practical difficulty also arises with fermions: because it is impossible to simulate numerically fermions, it is necessary to integrate over fermions explicitly. This generates an effective gauge field action which contains a contribution proportional to the logarithm of the fermion determinant and is, therefore, no longer local. The speed of numerical methods crucially depends on the locality of the action. This explains that most numerical simulations with fermions have been, for some time, performed in the so-called quenched approximation, in which the determinant is neglected. This approximation corresponds to the neglect of all fermion loops and bears some similarity with the eikonal approximation. In this approximation, the approximate spontaneous breaking of chiral symmetry has been verified by measuring the decrease of the pion mass for decreasing quark masses. Owing to the difficulty of the problem, a precise numerical study of the effect of dynamical fermions at realistic lattice sizes, and close enough to the chiral limit, has only started in the two last decades. The physical spectrum of hadrons has finally been reproduced, confirming once again the relevance of QCD [280]. Moreover, Lüscher has also proposed a clever method to calculate S -matrix elements [281] from Euclidean field theory.

A25 Gauge theory and confinement in two dimensions

Two dimensions is, from the point of view of gauge theories, peculiar in the sense that a gauge field really has no dynamical degrees of freedom. Still the gauge field generates a force between particles which, as we have seen in Section 25.3, leads to confinement, even in the Abelian case. An example of such a situation is discussed in the solvable Schwinger model in Section 30.5. This is a question that we examine here on the lattice for pure gauge theories [282].

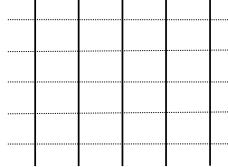


Fig. 25.2 The dotted lines correspond to $\mathbf{U} = 1$

Lattice gauge theories in two dimensions. We consider a pure gauge action (25.8) in a two-dimensional lattice Ω of size $aL_1 \times aL_2$, where a is the lattice spacing. The partition function is given by ($\mathbf{U}_{ij} \equiv \mathbf{U}(x_i, x_j)$)

$$\mathcal{Z} = \int \prod_{\text{links } \ell \in \Omega} d\mathbf{U}_\ell \exp \left(\beta \sum_{\text{plaquettes}} \mathcal{S}(\mathbf{U}) \right), \quad (A25.1)$$

with (n_1 and n_2 are the unit vectors in the time and space directions, resp.),

$$\begin{aligned} \mathcal{S}(\mathbf{U}) &= \sum_{x \in \Omega} \text{tr } \mathbf{U}(x, x + an_1) \mathbf{U}(x + an_1, x + a(n_1 + n_2)) \\ &\quad \times \mathbf{U}(x + a(n_1 + n_2), x + an_2) \mathbf{U}(x + an_2, x). \end{aligned}$$

We assume free boundary conditions in the n_1 direction (to avoid closed loop variables surviving due to the topological properties of the two-dimensional space manifold). We then use gauge invariance to set to 1 all link variables in the n_1 direction (the horizontal links in Fig. 25.2). This is the equivalent of the temporal gauge of the continuum theory. Then,

$$\begin{aligned} \mathcal{S}(\mathbf{U}) &= \sum_{x \in \Omega} \text{tr } \mathbf{U}(x + an_1, x + an_1 + an_2) \mathbf{U}(x + an_2, x) \\ &= \sum_x \text{tr } \mathbf{U}^{-1}(x + an_1 + an_2, x + an_1) \mathbf{U}(x + an_2, x). \end{aligned}$$

In two dimensions this has the remarkable effect of decoupling all links in the orthogonal, n_2 direction (the vertical direction in Fig 25.2). The partition function factorizes. If L_2 is the size in the n_2 direction, one finds

$$\mathcal{Z} = \mathcal{Z}_1^{L_2}, \quad (A25.2)$$

with

$$\mathcal{Z}_1 = \int \prod_{0 \leq m' < L_1} dU_{m'} \exp \left(\beta \sum_{0 \leq m < L_1} \text{tr } \mathbf{U}_{m+1}^{-1} \mathbf{U}_m \right), \quad (A25.3)$$

where $\mathbf{U}_m \equiv \mathbf{U}(man_1 + an_2, man_1)$.

We recognize in \mathcal{Z}_1 the partition function of a one-dimensional model with a nearest neighbour interactions. The partition function can be calculated by the transfer matrix method. In the case of free space boundary conditions, we also can set

$$\mathbf{V}_{m+1} = \mathbf{U}_{m+1}^{-1} \mathbf{U}_m. \quad (A25.4)$$

The integrals over the matrices \mathbf{V}_m then factorize. One finds

$$\mathcal{Z}(\beta) = z^\omega(\beta), \quad (A25.5)$$

in which $\omega = L_1 L_2$, $a^2 \omega$ is the area of the lattice, and $z(\beta)$ is given by

$$z(\beta) = \int d\mathbf{U} \exp(\beta \operatorname{tr} \mathbf{U}). \quad (A25.6)$$

Hereafter, we assume that the volume of the group has been normalized to 1. The function $z(\beta)$ is, as expected, a regular positive function: no phase transition occurs in two dimensions. With the same method, one can also calculate the expectation value of Wilson's loop. Denoting by aR and aT the sizes of the loop in space and time, one finds

$$W(C) = e^{-RT\sigma(\beta)}, \quad (A25.7)$$

with

$$\sigma(\beta) = -\ln \left(\frac{z'(\beta)}{Nz(\beta)} \right). \quad (A25.8)$$

We have again assumed that the centre of the group is non-trivial and called N the trace of the unit matrix. As expected, in two dimensions Wilson's loop follows an *area law* for all groups and all couplings.

For β small, $z(\beta)$ has the expansion

$$z(\beta) = 1 + z_2 \beta^2 + O(\beta^4). \quad (A25.9)$$

This yields for the *string tension*,

$$\sigma(\beta) \sim -\ln \beta, \quad (A25.10)$$

in agreement with expression (25.35). For β large, $z(\beta)$ can be calculated by the steepest descent method, which also generates perturbation theory, and one finds

$$\ln z(\beta) = N(\beta - K \ln \beta) + O(\beta^{-1}), \quad (A25.11)$$

in which K is a constant. Therefore,

$$\sigma(\beta) \sim K/\beta, \quad (A25.12)$$

in agreement with perturbation theory in the continuum.

Finally, since in two dimensions the pure gauge theory has no physical degrees of freedom, no particle propagates, and no gluonium state can be found.

26 Becchi–Rouet–Stora–Tyutin (BRST) symmetry. Gauge theories: Zinn–Justin equation (ZJ) and renormalization

In the first part of the chapter, we explain Faddeev–Popov’s quantization and the resulting Slavnov–Taylor (ST) identities in a simple context. After introduction of Faddeev–Popov ‘ghost’ fermions, the formalism automatically leads to a BRST symmetry [236, 237]. We explore some properties of the symmetry. The differential operator, of fermionic type, representing the BRST symmetry, with a proper choice of variables, has the form of a cohomology operator. It has a simple form in terms of *Grassmann coordinates*.

BRST symmetry plays an essential role in the quantization and renormalization of non-Abelian gauge theories, as discussed in the second part of the chapter, but also in the renormalization of generalized non-linear σ -models in Chapter 29, and in the study of stochastic differential equations and critical dynamics in Chapters 34–36.

Quantization of gauge theories requires a gauge-fixing procedure. Starting from the non-covariant temporal gauge, and using a simple identity, we show the equivalence with a quantization in a general class of gauges, including relativistic covariant gauges. Adapting the formalism developed in Sections 26.1–26.4, we derive ST identities and the corresponding BRST symmetry. We apply the formalism to renormalization. However, the explicit form of the BRST symmetry is not stable under renormalization. We show that the initial BRST symmetry implies a more general, quadratic *master equation*, also called ZJ equation, satisfied by the quantized action, equation in which gauge and BRST symmetries are no longer explicit. By contrast, in the case of renormalizable gauges, the ZJ equation is stable under renormalization. We solve it to determine the structure of the renormalized action. We verify that it leads to a renormalized form of the BRST symmetry and encodes, in a subtle way, the gauge properties of the quantized action.

Physical observables in a gauge theory should be independent of the specific gauge fixing, that is, should be gauge independent. We first prove that the vacuum amplitude (or partition function) is gauge independent. We argue that this property remains true for correlation functions of gauge-invariant operators of canonical dimension not larger than 4 in four dimensions. As a consequence, these correlation functions have a physical meaning. A similar property holds for all gauge-invariant operators, but the proof is more involved and is not given here. Finally, when a S -matrix can be defined, S -matrix elements also are gauge independent.

The appendix describes additional relations between BRST symmetry and ZJ equation, and between BRST symmetry and supersymmetry (see Section 34.7 and Chapter 27).

Notation. For a general discussion of non-Abelian gauge theories, it is useful to introduce an abstract formalism, which makes it possible to understand the algebraic structure and the renormalization procedure, without being overwhelmed by the notational complexity. In particular, for simplicity we give to variables and equations discrete indices.

However, the drawback of the notation is that the translation of the general identities, which then appear, to a usual, more concrete notation becomes a non-trivial exercise. For example, the translation may involve replacing summation by integration and summation, and differentiation by functional differentiation.

Finally, *implicit summation over repeated indices is always meant*.

26.1 ST identities: The origin

The ST identities in gauge theories [283, 284] owe less to gauge symmetry than to the gauge fixing procedure. Therefore, it is possible to explain their origin and properties in a rather general, but simple framework.

26.1.1 A simple identity

Let φ^α be a (finite or infinite) set of real (for simplicity) variables satisfying a system of real (algebraic, differential, functional differential) equations, which we denote

$$E_\alpha(\varphi) = 0, \quad (26.1)$$

where the functions $E_\alpha(\varphi)$ are expandable in powers of φ .

We assume that the functions $E_\alpha(\varphi)$ are smooth, and $E_\alpha = E_\alpha(\varphi)$ is a one-to-one mapping in some neighbourhood of $E_\alpha = 0$, which can be inverted in $\varphi^\alpha = \varphi^\alpha(E)$. In particular, this implies that the equation has a unique solution $\varphi_s^\alpha \equiv \varphi^\alpha(0)$. In the neighbourhood of φ_s , the determinant $\det \mathbf{E}$ of the matrix \mathbf{E} with elements

$$E_{\alpha\beta} \equiv \frac{\partial E_\alpha}{\partial \varphi^\beta},$$

does not vanish, and thus $E_\alpha(\varphi)$ can be chosen such that $\det \mathbf{E}$ is positive.

For any continuous function $F(\varphi)$, a simple formal expression for $F(\varphi_s)$ can be derived, which *does not involve solving the equation explicitly*. One starts from the obvious identity,

$$F(\varphi_s) = \int \left\{ \prod_\alpha dE^\alpha \delta(E_\alpha) \right\} F(\varphi(E)),$$

where $\delta(E)$ is Dirac's δ -function. One then changes variables $E \mapsto \varphi$. This generates the Jacobian $\det \mathbf{E} > 0$. Thus,

$$F(\varphi_s) = \int \left\{ \prod_\alpha d\varphi^\alpha \delta[E_\alpha(\varphi)] \right\} \det \mathbf{E}(\varphi) F(\varphi). \quad (26.2)$$

In the context of non-Abelian gauge theories, $\det \mathbf{E}$ is the Faddeev–Popov determinant (see Section 26.7) [219].

26.1.2 An invariant measure

Since $\prod_\alpha dE_\alpha$ is the invariant measure for the group of translations, $E_\alpha \mapsto E_\alpha + \nu_\alpha$ with ν_α constants, the measure

$$d\rho(\varphi) = \det \mathbf{E}(\varphi) \prod_\alpha d\varphi^\alpha, \quad (26.3)$$

is the invariant measure for the translation group realized non-linearly on the new coordinates φ_α (at least in some neighbourhood of $\nu_\alpha=0$):

$$\varphi^\alpha \mapsto \varphi'^\alpha, \quad \text{with} \quad E_\alpha(\varphi') - \nu_\alpha = E_\alpha(\varphi). \quad (26.4)$$

Expanding for ν_α small, one obtains the infinitesimal form of the transformation:

$$\varphi'^\alpha - \varphi^\alpha = [E^{-1}(\varphi)]^{\alpha\beta} \nu_\beta + O(|\nu|^2). \quad (26.5)$$

In the context of non-Abelian gauge theories, formally analogous equations can be derived, where the variables φ^α parametrize gauge transformations and equation (26.1) fixes the gauge. The corresponding infinitesimal transformations (26.5) are at the basis of the derivation by Slavnov and, independently, Taylor of a set of important identities satisfied by Green's functions, thus called *ST identities* [283, 284] (see Section 26.7.1). This method has then been extended to various other field theories, for example, to the dynamics generated by a Langevin equation [285] (Chapters 34, 35).

Reciprocal property. Conversely, we can characterize the general form of non-linear representations of the translation group. We write an infinitesimal group transformation of parameters ν_α as,

$$\delta\varphi^\alpha = [M^{-1}(\varphi)]^{\alpha\beta}\nu_\beta, \quad (26.6)$$

in which the matrix $M_{\alpha\beta}(\varphi)$ has to be determined. We impose, to the differential operators Δ^α ,

$$\Delta^\alpha = [M^{-1}(\varphi)]^{\beta\alpha}\frac{\partial}{\partial\varphi^\beta}, \quad (26.7)$$

to form a representation of the Lie algebra of the translation group, that is, to commute (see also Section 29.3):

$$[\Delta^\alpha, \Delta^\beta] = 0. \quad (26.8)$$

This implies

$$[M^{-1}]^{\gamma\alpha}[M^{-1}\partial_\gamma MM^{-1}]^{\delta\beta} = [M^{-1}]^{\gamma\beta}[M^{-1}\partial_\gamma MM^{-1}]^{\delta\alpha}.$$

Multiplying by $M_{\alpha\alpha'}M_{\beta\beta'}M_{\delta\delta'}$ and summing over α, β, δ , we find

$$\partial_\beta M_{\delta\alpha} - \partial_\alpha M_{\delta\beta} = 0. \quad (26.9)$$

We infer (for a simply connected φ -manifold, an implicit assumption throughout the whole chapter) that $M_{\alpha\beta}$ has the form

$$M_{\alpha\beta} = \partial_\beta E_\alpha. \quad (26.10)$$

We now characterize the invariant measure $\mathcal{J}(\varphi)d\varphi$ for these non-linear transformations. The variation of $\mathcal{J}(\varphi)$ has to cancel the Jacobian coming from the change of variables corresponding to the transformation (26.6):

$$\partial_\alpha \mathcal{J}[M^{-1}(\varphi)]^{\alpha\beta} + \mathcal{J}\partial_\alpha[M^{-1}(\varphi)]^{\alpha\beta} = 0. \quad (26.11)$$

This yields a system of partial differential equations for the function $\mathcal{J}(\varphi)$:

$$\partial_\alpha \ln \mathcal{J} = [M^{-1}]^{\gamma\beta}\partial_\gamma M_{\beta\alpha}. \quad (26.12)$$

Equation (26.9) is an integrability condition for equation (26.12). We can use it to rewrite equation (26.12) as

$$\partial_\alpha \ln \mathcal{J} = \partial_\alpha \ln \det \mathbf{M}, \quad (26.13)$$

which has the solution

$$\mathcal{J} \propto \det \mathbf{M}. \quad (26.14)$$

26.2 From ST symmetry to BRST symmetry

In quantum field theory (QFT), the non-linear and non-local character of the transformations (26.5) is the source of technical complications. Remarkably enough, the invariance under the infinitesimal transformations (26.5) can be replaced by an invariance under linear anticommuting-type transformations, at the expense of introducing additional Grassmann variables.

One again starts from the identity (26.2) and first replaces the δ -function by its Fourier representation:

$$\prod_{\alpha} \delta [E_{\alpha}(\varphi)] = \int \prod_{\alpha} \frac{d\lambda^{\alpha}}{2i\pi} \exp [-\lambda^{\alpha} E_{\alpha}(\varphi)],$$

where the λ integration runs along the imaginary axis. Moreover, the determinant can be written as an integral over Grassmann variables \bar{c}^{α} and c^{α} (see Section 1.7):

$$\det \mathbf{E} = \int \prod_{\alpha} (d\bar{c}^{\alpha} dc^{\alpha}) \exp [c^{\alpha} E_{\alpha\beta}(\varphi) \bar{c}^{\beta}].$$

The expression (26.2) then becomes

$$F(\varphi_s) = \mathcal{N} \int \prod_{\alpha} (d\varphi^{\alpha} d\lambda^{\alpha} d\bar{c}^{\alpha} dc^{\alpha}) F(\varphi) \exp [-S(\varphi, \lambda, c, \bar{c})], \quad (26.15)$$

in which \mathcal{N} is a constant normalization factor,

$$\mathcal{N}^{-1} = \int \prod_{\alpha} (d\varphi^{\alpha} d\lambda^{\alpha} d\bar{c}^{\alpha} dc^{\alpha}) \exp [-S(\varphi, \lambda, c, \bar{c})],$$

and $S(\varphi, \lambda, c, \bar{c})$ the function (and element of the Grassmann algebra),

$$S(\varphi, \lambda, c, \bar{c}) = \lambda^{\alpha} E_{\alpha}(\varphi) - c^{\alpha} E_{\alpha\beta}(\varphi) \bar{c}^{\beta}. \quad (26.16)$$

In the context of non-Abelian gauge theories, \mathbf{c} and $\bar{\mathbf{c}}$ are non-physical fermion fields, dubbed Faddeev–Popov ghosts (see Section 26.7).

26.2.1 BRST symmetry

Somewhat surprisingly, the function S has a new type of symmetry, which is an extension to Grassmann variables of the invariance of the measure $\det \mathbf{E}(\varphi) d\varphi$ under the infinitesimal transformations (26.5).

The BRST symmetry, first noticed, in the context of quantized gauge theories [236, 237], is a Grassmann symmetry, in the sense that the parameter ε of the transformation is an additional generator of the Grassmann algebra. The variations of the various integration variables are

$$\delta_{\text{BRST}} \varphi^{\alpha} = \varepsilon \bar{c}^{\alpha}, \quad \delta_{\text{BRST}} \bar{c}^{\alpha} = 0, \quad \delta_{\text{BRST}} c^{\alpha} = \varepsilon \lambda^{\alpha}, \quad \delta_{\text{BRST}} \lambda^{\alpha} = 0, \quad (26.17)$$

with

$$\varepsilon^2 = 0, \quad \varepsilon \bar{c}^{\alpha} + \bar{c}^{\alpha} \varepsilon = 0, \quad \varepsilon c^{\alpha} + c^{\alpha} \varepsilon = 0.$$

The transformation has a *vanishing square*:

$$\delta_{\text{BRST}}^2 = 0. \quad (26.18)$$

In the expression (26.15), the integration measure is also invariant.

Grassmann equations. Analogous expressions can be derived for a system of Grassmann equations $\mathcal{E}_\alpha(c) = 0$ involving Grassmann variables c^α . The Lagrange multiplier \bar{c} is then a Grassmann variable. The properties of change of variables in a Grassmann integral imply that the determinant, $\det \mathbf{E}$, is replaced by the inverse of a determinant. Therefore, the auxiliary variables, which have to be introduced to represent the determinant, are commuting complex variables φ, λ . The equation $\mathcal{E}_\alpha(c) = 0$ leads to

$$S(c, \bar{c}, \varphi, \lambda) = \mathcal{D}[\varphi^\alpha \mathcal{E}_\alpha(c)] = \bar{c}^\alpha \mathcal{E}_\alpha(c) + \varphi^\alpha \frac{\partial \mathcal{E}_\alpha}{\partial c^\beta} \lambda^\beta. \quad (26.19)$$

ST and BRST symmetries. A simple way of understanding the relation between the transformations (26.5) and (26.17) is to note that

$$\det \mathbf{E} (E^{-1})^{\alpha\beta} = \int d\bar{c} dc c^\beta \bar{c}^\alpha \exp(c^\gamma E_{\gamma\delta} \bar{c}^\delta). \quad (26.20)$$

Therefore, if we factorize the integral over c and \bar{c} , we can rewrite the transformation (26.5) inside the integral (26.15) as

$$\delta \varphi^\alpha = (E^{-1})^{\alpha\beta} \nu_\beta = -\bar{c}^\alpha (c^\beta \nu_\beta). \quad (26.21)$$

BRST transformations correspond to the mapping $-c^\alpha \nu_\alpha \mapsto \varepsilon$.

A form of this argument can be used to prove that as long as only functions of φ are concerned, the consequences of ST or BRST symmetry are the same. The BRST symmetry, which extends the transformations to functions of φ, λ, c , and \bar{c} , greatly simplifies the renormalization of various field theories, for two reasons:

(i) The transformations (26.17) are linear, while the transformation (26.5) is non-linear.

We have already seen, in the case of the non-linear σ -model, that non-linear transformations can sometimes be linearized at the price of introducing auxiliary fields.

(ii) More important, in field theories the transformations (26.17) are local, in contrast to the transformation (26.5).

26.2.2 BRST transformations: Differential operator representation

The BRST transformation δ_{BRST} , when acting on functions of $\{\varphi, \lambda, c, \bar{c}\}$, can be represented by the Grassmann differential operator

$$\mathcal{D} \equiv \bar{c}^\alpha \frac{\partial}{\partial \varphi^\alpha} + \lambda^\alpha \frac{\partial}{\partial c^\alpha}. \quad (26.22)$$

The property (26.18) is expressed by the identity (which one can verify directly)

$$\mathcal{D}^2 = 0. \quad (26.23)$$

The differential operator \mathcal{D} has the form of a *cohomology operator*, generalization of the exterior differentiation of differential forms (the first term $\bar{c}^\alpha \partial/\partial \varphi_\alpha$ in the BRST operator has the property of the differentiation of forms).

A quantity Q that satisfies $\mathcal{D}Q = 0$ is called BRST *closed*. The equation $\mathcal{D}^2 = 0$ implies that all quantities of the form $\mathcal{D}Q(\varphi, \lambda, c, \bar{c})$, called *BRST exact*, are BRST closed. One immediately verifies that the function S defined by equation (26.16) is BRST exact:

$$S = \mathcal{D}\Phi, \text{ with } \Phi = c^\alpha E_\alpha(\varphi), \quad (26.24)$$

and, thus, BRST closed,

$$\mathcal{D}S = 0. \quad (26.25)$$

The implications of the BRST symmetry, like the classification of exact contributions, rely on arguments based on BRST cohomology. The reduction of \mathcal{S} to the exact form (26.24) requires additional conditions like the degrees in c^α and λ^α , and an equal number of c^α and \bar{c}^α factors in products (fermion number conservation).

If one relaxes the degree conditions, other exact contributions can be added. For example, at degree 1 in λ^α , one can add to Φ the monomial,

$$\Phi_1 = \frac{1}{2} c^\alpha a_{\alpha\beta} \lambda^\beta, \quad \text{with } a_{\alpha\beta} \text{ constant symmetric positive matrix.} \quad (26.26)$$

This translates into the addition to \mathcal{S} of $\mathcal{S}_1 = \frac{1}{2} \lambda^\alpha a_{\alpha\beta} \lambda^\beta$. Integrating then expression (26.15) over λ^α , one verifies that now the strict condition (26.1) is replaced by a Gaussian average centred around $E_\alpha = 0$.

At degree 2 in c^α , one can add

$$\Phi_2 = g_{\alpha\beta\gamma\delta} c^\alpha c^\beta \bar{c}^\gamma \varphi^\delta,$$

with $g_{\alpha\beta\gamma\delta}$ antisymmetric in $\alpha \leftrightarrow \beta$ and $\gamma \leftrightarrow \delta$. This yields to the addition to \mathcal{S} of

$$\mathcal{S}_2 = g_{\alpha\beta\gamma\delta} (-c^\alpha c^\beta \bar{c}^\gamma \bar{c}^\delta + 2\lambda^\alpha \varphi^\delta c^\beta \bar{c}^\gamma). \quad (26.27)$$

26.3 BRST symmetry: More general coordinates. Group structure

Change of coordinates. Although the first term $\bar{c}^\alpha \partial_\alpha$ in the BRST operator (26.22) is identical to the differentiation of forms, here the Grassmann variables \bar{c}^α are not simply external variables introduced for convenience to exhibit the antisymmetry of the corresponding tensors; the \bar{c}^α are additional dynamical variables. In particular, it may be convenient, as we will show, to change the variables $\varphi^\alpha, \bar{c}^\alpha$. In a different set of variables, the transformations (26.17) take a more general form (see, e.g Section 29.3.2).

26.3.1 Group manifolds and gauge invariance

When the variables φ^α parametrize an element $\mathbf{g}(\varphi)$ of a Lie group in some matrix representation, it is convenient to express BRST transformations on $\mathbf{g}(\varphi)$ directly and to parametrize the variation of \mathbf{g} in terms of a Grassmann matrix $\bar{\mathbf{C}}$ belonging to the Lie algebra of the group:

$$\delta_{\text{BRST}} \mathbf{g} = \varepsilon \bar{\mathbf{C}} \mathbf{g}. \quad (26.28)$$

Thus,

$$\delta_{\text{BRST}} \mathbf{g} = \varepsilon \bar{\mathbf{c}}^\alpha \frac{\partial \mathbf{g}}{\partial \varphi^\alpha} = \varepsilon \bar{\mathbf{C}} \mathbf{g} \Rightarrow \bar{\mathbf{C}} = \bar{\mathbf{c}}^\alpha \frac{\partial \mathbf{g}}{\partial \varphi^\alpha} \mathbf{g}^{-1}.$$

The variation of $\bar{\mathbf{C}}$ then is given by

$$\begin{aligned} \delta_{\text{BRST}} \bar{\mathbf{C}} &= -\varepsilon \bar{\mathbf{c}}^\alpha \bar{\mathbf{c}}^\beta \frac{\partial^2 \mathbf{g}}{\partial \varphi^\alpha \partial \varphi^\beta} \mathbf{g}^{-1} - \bar{\mathbf{c}}^\alpha \frac{\partial \mathbf{g}}{\partial \varphi^\alpha} \mathbf{g}^{-1} \varepsilon \bar{\mathbf{c}}^\beta \frac{\partial \mathbf{g}}{\partial \varphi^\beta} \mathbf{g}^{-1} \\ &= -\bar{\mathbf{c}}^\alpha \frac{\partial \mathbf{g}}{\partial \varphi^\alpha} \mathbf{g}^{-1} \varepsilon \bar{\mathbf{c}}^\beta \frac{\partial \mathbf{g}}{\partial \varphi^\beta} \mathbf{g}^{-1} = \varepsilon \bar{\mathbf{C}}^2. \end{aligned} \quad (26.29)$$

This expression also appears in gauge theories, the group element being there associated with gauge transformations, and the transformation (26.28) being an infinitesimal gauge transformation (see Section 26.7.1).

One can verify directly that, as expected, the BRST variation of $\bar{\mathbf{C}}^2$ vanishes:

$$\delta_{\text{BRST}} \bar{\mathbf{C}}^2 = \bar{\mathbf{C}} \varepsilon \bar{\mathbf{C}}^2 + \varepsilon \bar{\mathbf{C}}^2 \bar{\mathbf{C}} = 0.$$

To render the relation with the BRST transformations of gauge theories even more explicit, it is instructive to introduce an element \mathbf{B} of the Lie algebra, not affected by BRST transformations, and assume that the function (26.1), which determines \mathbf{g} is a function only of the product

$$\mathbf{A} = \mathbf{g} \mathbf{B} \mathbf{g}^{-1}.$$

The BRST transformation of \mathbf{A} then is

$$\delta_{\text{BRST}} \mathbf{A} = \varepsilon \bar{\mathbf{C}} \mathbf{A} - \mathbf{A} \varepsilon \bar{\mathbf{C}} = \varepsilon [\bar{\mathbf{C}}, \mathbf{A}], \quad (26.30)$$

which is again an element of the Lie algebra. A short calculation again confirms that, as expected, $[\bar{\mathbf{C}}, \mathbf{A}]$ is BRST closed.

When $\bar{\mathbf{C}}$ and \mathbf{A} are expanded on a basis of generators \mathbf{t}_i of the Lie algebra, $\bar{\mathbf{C}} = \bar{c}^i \mathbf{t}_i$, $\mathbf{A} = a^i \mathbf{t}_i$, the transformations (26.30) and (26.29) become

$$\delta_{\text{BRST}} \bar{c}^i = \frac{1}{2} \varepsilon f_{jk}^i \bar{c}^j \bar{c}^k, \quad (26.31)$$

$$\delta_{\text{BRST}} a^i = \varepsilon f_{jk}^i \bar{c}^j a^k, \quad (26.32)$$

where the coefficients f_{jk}^i are the structure constants of the Lie algebra. These transformations are examples of the transformations (26.34), and the identity $\delta_{\text{BRST}}^2 = 0$ is satisfied, because f_{jk}^i is a structure constant.

Finally, \mathbf{C} and $\boldsymbol{\lambda}$ can also be chosen as belonging to the Lie algebra, $\mathbf{C} = c^i \mathbf{t}_i$, $\boldsymbol{\lambda} = \lambda^i \mathbf{t}_i$, and then,

$$\delta_{\text{BRST}} c^i = \varepsilon \lambda^i, \quad \delta_{\text{BRST}} \lambda^i = 0. \quad (26.33)$$

26.3.2 Direct construction

Section A26.1 describes a method to construct a large class of BRST transformations, starting from a ZJ equation. Starting from a function,

$$\mathcal{S}(\varphi, \bar{c}) = \frac{1}{2} F_{\alpha\beta}(\varphi) \bar{c}^\alpha \bar{c}^\beta,$$

with $F_{\alpha\beta} = -F_{\beta\alpha}$, one obtains (equation (A26.1))

$$\delta_{\text{BRST}} \varphi_\alpha = \varepsilon F_{\alpha\beta}(\varphi) \bar{c}^\beta, \quad \delta_{\text{BRST}} \bar{c}^\alpha = \frac{1}{2} \varepsilon \frac{\partial F_{\beta\gamma}(\varphi)}{\partial \varphi_\alpha} \bar{c}^\beta \bar{c}^\gamma, \quad (26.34)$$

where the functions $F_{\alpha\beta}(\varphi)$ must satisfy the consequence of a ZJ equation (A26.5),

$$\{F_{\delta\alpha} \partial F_{\beta\gamma} / \partial \varphi_\delta\}_{\alpha\beta\gamma} = 0,$$

where the global subscript $\alpha\beta\gamma$ means antisymmetrized in the three indices.

In Section A26.1, we have indicated that the transformations (26.34) leave the integration measure invariant. When $F^{\alpha\beta}$ is linear, $F_{\alpha\beta} = f_{\alpha\beta}^\gamma \varphi_\gamma$, one recognizes the Jacobi identity for the structure constants $f_{\alpha\beta}^\gamma$ of a Lie algebra, and φ transforms by the adjoint representation, as in equations (26.31, 26.32).

26.3.3 Generalization and compatibility conditions

In Section 26.6, we consider situations in which physics is described in terms of some variables A^i , themselves functions of the smaller number of variables φ^α . Then BRST transformations have the form

$$\delta A^i = \varepsilon D_\beta^i(A) \bar{c}^\beta, \quad \delta \bar{c}^\alpha = -\frac{1}{2} f_{\beta\gamma}^\alpha(A) \varepsilon \bar{c}^\beta \bar{c}^\gamma. \quad (26.35)$$

The property (26.23), $\mathcal{D}^2 = 0$, implies for D and f the two equations,

$$\{f_{\alpha\beta}^\delta f_{\gamma\delta}^\epsilon + D_\alpha^i \partial_i f_{\beta\gamma}^\epsilon\}_{\alpha\beta\gamma} = 0, \quad (26.36)$$

$$D_\alpha^j \partial_j D_\beta^i - D_\beta^j \partial_j D_\alpha^i = f_{\alpha\beta}^\delta D_\delta^i, \quad (26.37)$$

where the subscript $\alpha\beta\gamma$ in equation (26.36) means antisymmetrized in the three indices.

The extension we find here can be understood in the following way. We consider the set of first-order partial differential equations,

$$\Delta_\alpha S(A) = 0, \quad (26.38)$$

where the Δ_α are the differential operators,

$$\Delta_\alpha = D_\alpha^i(A) \frac{\partial}{\partial A^i}.$$

As discussed in Sections 13.1 and 29.3, the system (26.38) is called compatible if the equations $[\Delta_\alpha, \Delta_\beta]S = 0$ are linear combinations of the initial equations (26.38). Then,

$$[\Delta_\alpha, \Delta_\beta] = f_{\alpha\beta}^\gamma(A) \Delta_\gamma. \quad (26.39)$$

We have only encountered examples where the structure constants $f_{\alpha\beta}^\gamma$ were A -independent up to now, but this is not the general situation.

Equation (26.39) itself has an integrability condition, the Jacobi identities, since the left-hand side is a commutator. A short calculation yields the condition (26.36). Therefore, the nilpotency of the BRST operator encodes the compatibility of the linear system (26.38), and the Lie algebra structure. Finally, note that equation (26.38) is equivalent to $\mathcal{D}S(A) = 0$. If it has non-trivial solutions, these solutions cannot be cast into the form $S(A) = \mathcal{D}\Phi(A, c)$, that is, they are not BRST exact.

26.4 Stochastic equations

We now assume that equation (26.1) depends on a set of stochastic variables ν_a , the ‘noise’, with normalized probability measure $d\rho(\nu)$:

$$E_\alpha(\varphi, \nu) = 0. \quad (26.40)$$

The solution φ^α of the equation becomes a stochastic variable. Quantities of interest are then expectation values of functions of φ :

$$\langle F(\varphi) \rangle = \int d\rho(\nu) \prod_\alpha d\varphi^\alpha \delta [E_\alpha(\varphi, \nu)] \det \mathbf{E} F(\varphi). \quad (26.41)$$

After the set of transformations described in Section 26.2, this representation becomes

$$\langle F(\varphi) \rangle \propto \int d\rho(\nu) \prod_\alpha (d\varphi^\alpha d\lambda^\alpha d\bar{c}^\alpha dc^\alpha) F(\varphi) \exp [-S(\varphi, \lambda, c, \bar{c}, \nu)], \quad (26.42)$$

with S given by equation (26.16):

$$S(\varphi, \lambda, c, \bar{c}, \nu) = \lambda^\alpha E_\alpha(\varphi, \nu) - c^\alpha E_{\alpha\beta}(\varphi, \nu) \bar{c}^\beta. \quad (26.43)$$

We introduce the function $\Sigma(\varphi, \lambda, c, \bar{c})$ obtained by noise averaging:

$$\exp[-\Sigma(\varphi, \lambda, c, \bar{c})] = \int d\rho(\nu) \exp[-S(\varphi, \lambda, c, \bar{c}, \nu)]. \quad (26.44)$$

Then,

$$\langle F(\varphi) \rangle \propto \int \prod_{\alpha} (d\varphi^{\alpha} d\lambda^{\alpha} d\bar{c}^{\alpha} dc^{\alpha}) F(\varphi) \exp[-\Sigma(\varphi, \lambda, c, \bar{c})]. \quad (26.45)$$

We have shown that S has a BRST symmetry. Applying the BRST operator (26.22) on both sides of equation (26.44), we conclude that $\Sigma(\varphi, \lambda, c, \bar{c})$ is still BRST symmetric,

$$\mathcal{D}\Sigma = 0, \quad (26.46)$$

although it no longer has the simple form (26.43), that is, a function linear in λ and $c\bar{c}$. Moreover, because S is BRST exact, the function Σ is also BRST exact. This can be shown by simple algebraic manipulations based on the identity,

$$f(\mathcal{D}X) = f(0) + \mathcal{D}[Xg(\mathcal{D}X)], \quad \text{with } g(x) = \frac{f(x) - f(0)}{x},$$

Grassmann number conservation. The function S only involves a Grassmann combination of the form $c\bar{c}$. Therefore, if we multiply c by a phase $e^{i\zeta}$ and \bar{c} by the complex conjugated phase $e^{-i\zeta}$, S is invariant. This symmetry is also a symmetry of Σ and has implications for expectation values of elements of the Grassmann algebra, as Wick's theorem (1.78) also shows.

26.4.1 A simple example: Stochastic equations linear in the noise

In the several places in this work, stochastic equations of the simple algebraic form

$$E_{\alpha}(\nu, \varphi) \equiv E_{\alpha}(\varphi) - \nu_{\alpha}, \quad (26.47)$$

are met. Introducing the Laplace transform of the measure $d\rho(\nu)$,

$$e^{w(\lambda)} = \int d\rho(\nu) e^{\lambda^{\alpha} \nu_{\alpha}}, \quad (26.48)$$

we obtain for the function Σ defined by equation (26.44),

$$\Sigma(\varphi, \lambda, c, \bar{c}) = -w(\lambda) + \lambda^{\alpha} E_{\alpha}(\varphi) - c^{\alpha} E_{\alpha\beta} \bar{c}^{\beta}. \quad (26.49)$$

As explained above, $\Sigma(\varphi, \lambda, c, \bar{c})$ is BRST exact. Explicitly one finds,

$$\Sigma(\varphi, \lambda, c, \bar{c}) = \mathcal{D}\Phi, \quad \text{with } \Phi(\varphi, \lambda, c) = c^{\alpha} \left[E_{\alpha}(\varphi) - \frac{\partial}{\partial \lambda^{\alpha}} \int_0^1 ds w(s\lambda) \right]. \quad (26.50)$$

This is a minimal modification of expression (26.16).

When the noise distribution is given by a Gaussian measure, and thus $w(\lambda)$ has the form,

$$w(\lambda) = \frac{1}{2} w_{\alpha\beta} \lambda^{\alpha} \lambda^{\beta},$$

it is also possible to integrate explicitly over the λ variables. The resulting integrand corresponds to

$$\Sigma(\varphi, c, \bar{c}) = \frac{1}{2} E_{\alpha}(\varphi) [w^{-1}]^{\alpha\beta} E_{\beta}(\varphi) - c^{\alpha} E_{\alpha\beta}(\varphi) \bar{c}^{\beta}.$$

The BRST transformation of c is then no longer linear:

$$\delta_{\text{BRST}} c^\alpha = \varepsilon [w^{-1}]^{\alpha\beta} E_\beta(\varphi).$$

In this form, the BRST transformation only has a vanishing square when φ is a solution of the equation $E(\varphi) = 0$. We conclude that the property $\mathcal{D}^2 = 0$ of BRST transformations is not true in all formulations, and may be satisfied only after the introduction of some auxiliary variables.

The general linear case. A slightly more general form is also encountered:

$$E_\alpha(\varphi, \nu) = E_\alpha(\varphi) - e_\alpha^a(\varphi) \nu_a. \quad (26.51)$$

Then,

$$\partial_\beta E_\alpha(\varphi, \nu) = \partial_\beta E_\alpha(\varphi) - \partial_\beta e_\alpha^a(\varphi) \nu_a.$$

Here, we only consider the special example of Gaussian stochastic variables ν with probability distribution,

$$d\rho(\nu) = \left(\prod_a d\nu_a \right) \exp(-\frac{1}{2} \nu_a [\Omega^{-1}]^{ab} \nu_b). \quad (26.52)$$

After integration over ν and some algebra, we find that Σ can be written as

$$\Sigma(\varphi, \lambda, c, \bar{c}) = \mathcal{D}\tilde{\Sigma}(\varphi, \lambda, c, \bar{c}), \quad (26.53a)$$

$$\tilde{\Sigma}(\varphi, \lambda, c, \bar{c}) = c^\alpha E_\alpha - \frac{1}{2} c^\alpha (w_{\alpha\beta}(\varphi) \lambda^\beta - w_{\alpha\beta,\gamma}(\varphi) c^\beta \bar{c}^\gamma), \quad (26.53b)$$

with the definitions,

$$w_{\alpha\beta} = e_\alpha^a \Omega_{ab} e_\beta^b, \quad w_{\alpha\beta,\gamma} = e_\alpha^a \Omega_{ab} \partial_\gamma e_\beta^b.$$

Σ differs from S by the addition of a function quadratic in both λ and $c\bar{c}$ (see equations (26.26) and (26.27)).

26.5 BRST symmetry, Grassmann coordinates, and gradient equations

Remark. This section is intended to provide an introduction to Chapters 34 and 35 where stochastic differential equations are considered.

26.5.1 BRST symmetry and Grassmann coordinates

A more compact representation of the BRST transformations (26.17) is obtained by introducing a Grassmann coordinate θ , and then the following two functions of θ :

$$\phi^\alpha(\theta) = \varphi^\alpha + \theta \bar{c}^\alpha, \quad C^\alpha(\theta) = c^\alpha + \theta \lambda^\alpha. \quad (26.54)$$

With this notation, the transformations (26.17) simply become a translation of θ :

$$\begin{aligned} \delta_{\text{BRST}} \phi^\alpha(\theta) &= \varepsilon \frac{\partial \phi^\alpha}{\partial \theta} = \phi^\alpha(\theta + \varepsilon) - \phi^\alpha(\theta), \\ \delta_{\text{BRST}} C^\alpha(\theta) &= \varepsilon \frac{\partial C^\alpha}{\partial \theta} = C^\alpha(\theta + \varepsilon) - C^\alpha(\theta). \end{aligned} \quad (26.55)$$

In particular, the BRST operator \mathcal{D} (equation (26.22)), which is a cohomology operator, is represented by $\partial/\partial\theta$,

$$\mathcal{D} \mapsto \frac{\partial}{\partial\theta} \Rightarrow \mathcal{D}^2 = 0.$$

In Section 26.2, we enforce the constraint

$$E_\alpha(\varphi) = 0, \quad (26.56)$$

by the integral (26.15). The integrand involves an effective action of the general form (equation (26.16))

$$S(\varphi, \lambda, c, \bar{c}) = \lambda^\alpha E_\alpha(\varphi) - c^\alpha \partial_\beta E_\alpha(\varphi) \bar{c}^\beta. \quad (26.57)$$

We note the expansion

$$C^\alpha(\theta) E_\alpha(\phi(\theta)) = c^\alpha E_\alpha(\varphi) + \theta \left(\lambda^\alpha E_\alpha(\varphi) - c^\alpha \frac{\partial E_\alpha}{\partial \varphi^\beta} \bar{c}^\beta \right). \quad (26.58)$$

Thus,

$$S(\varphi, \lambda, c, \bar{c}) = \frac{\partial}{\partial\theta} [C^\alpha(\theta) E_\alpha(\phi(\theta))], \quad (26.59)$$

which is equation (26.24) in a different notation, and shows again that S is BRST exact.

Because, in the case of Grassmann variables, integration and differentiation are identical operations, an integration over θ also selects the coefficient of θ . Therefore, $S(\varphi, \lambda, c, \bar{c})$ can be rewritten as an integral over θ :

$$S(\varphi, \lambda, c, \bar{c}) = \int d\theta C^\alpha(\theta) E_\alpha(\phi(\theta)). \quad (26.60)$$

In this expression, the BRST symmetry is manifest: the integrand does not depend on θ explicitly.

Note that, since the function S involves only a Grassmann combination of the form $c\bar{c}$ in a representation in terms of the functions (26.54), as in equation (26.60), each integration over θ is associated with a factor $C^\alpha(\theta)$.

BRST symmetry and group manifolds. If the variables φ^α parametrize a group element $g(\varphi)$ in some matrix representation, it is convenient to rewrite BRST transformations on $g(\varphi)$ directly (see also Section 26.3.1). This can be most easily done by noting that, with $\phi(\theta)$ (defined by equation (26.54)), we can associate a group element $\mathcal{G}(\theta)$, on which BRST transformations according to equation (26.55) act as

$$\delta_{\text{BRST}} \mathcal{G}(\theta) = \varepsilon \frac{\partial \mathcal{G}}{\partial \theta}. \quad (26.61)$$

However, since $\mathcal{G}(\theta)$ is a group element, it is natural to parametrize it in the form

$$\mathcal{G}(\theta) = \exp(\theta \bar{c}) g = (1 + \theta \bar{c}) g, \quad (26.62)$$

in which \bar{c} now is a Grassmann matrix belonging to the Lie algebra of the group. In component form, the transformation (26.61) then becomes (equations (26.28, 26.29))

$$\{\delta_{\text{BRST}} g = \varepsilon \bar{c} g, \delta_{\text{BRST}} \bar{c} g = 0\} \Rightarrow \delta_{\text{BRST}} \bar{c} = \varepsilon \bar{c}^2.$$

Introducing matrices \mathbf{t}_α , generators of the Lie algebra, and parametrizing \bar{c} as

$$\bar{c} = \bar{c}^\alpha \mathbf{t}_\alpha, \quad (26.63)$$

we can rewrite the transformation of \bar{c} (equations (26.31) and (29.45)) as

$$\delta_{\text{BRST}} \bar{c}^\alpha = \frac{1}{2} \varepsilon f_{\beta\gamma}^\alpha \bar{c}^\beta \bar{c}^\gamma. \quad (26.64)$$

26.5.2 Gradient equations and Grassmann coordinates

The two Grassmann variables, \bar{c}^α and c^α , which appear in equation (26.17), play, in general, different roles. However, there is one special situation in which a symmetry is established between them—when the equation (26.16) is such that $\partial_\beta E_\alpha$ is symmetric in $\alpha \leftrightarrow \beta$:

$$E_{\alpha\beta} = E_{\beta\alpha} \iff \partial_\beta E_\alpha = \partial_\alpha E_\beta. \quad (26.65)$$

Hence, assuming that the φ manifold is simply connected, $E_\alpha(\varphi)$ is itself a gradient; there exists a function $A(\varphi)$ such that

$$E_\alpha(\varphi) = \partial_\alpha A(\varphi). \quad (26.66)$$

The symmetry between c and \bar{c} generates an additional independent BRST symmetry, of the form

$$\delta_{\text{BRST}} \varphi^\alpha = c^\alpha \bar{\varepsilon}, \quad \delta_{\text{BRST}} c^\alpha = 0, \quad \delta_{\text{BRST}} \bar{c}^\alpha = \bar{\varepsilon} \lambda^\alpha, \quad \delta_{\text{BRST}} \lambda^\alpha = 0, \quad (26.67)$$

where $\bar{\varepsilon}$ is an additional element of the Grassmann algebra. In differential form, the generator of this second BRST symmetry reads

$$\bar{\mathcal{D}} = -c^\alpha \frac{\partial}{\partial \varphi^\alpha} + \lambda^\alpha \frac{\partial}{\partial \bar{c}^\alpha} \Rightarrow \bar{\mathcal{D}}^2 = 0, \quad \bar{\mathcal{D}}\mathcal{D} + \mathcal{D}\bar{\mathcal{D}} = 0.$$

It is then convenient to introduce two Grassmann coordinates $\bar{\theta}$, θ , and the function

$$\phi^\alpha(\bar{\theta}, \theta) = \varphi^\alpha + \theta \bar{c}^\alpha + c^\alpha \bar{\theta} + \theta \bar{\theta} \lambda^\alpha. \quad (26.68)$$

The operator $\bar{\mathcal{D}}$ is then represented by $\partial/\partial\bar{\theta}$.

In terms of ϕ , the expression (26.59) quite generally can be written as,

$$S(\phi) = \int d\bar{\theta} d\theta \bar{\theta} \frac{\partial \phi^\alpha}{\partial \bar{\theta}} E_\alpha [\phi(\bar{\theta}, \theta)]. \quad (26.69)$$

When the function $E(\varphi)$ has the special form (26.66), it is possible to integrate by parts over $\bar{\theta}$, and the function $S(\phi)$ then takes the remarkable form,

$$S(\phi) = \int d\bar{\theta} d\theta A [\phi(\bar{\theta}, \theta)] = \bar{\mathcal{D}}\mathcal{D}A(\varphi). \quad (26.70)$$

The two symmetries, which correspond to independent translations of θ and $\bar{\theta}$, are explicit here.

Gaussian noise. In Section 26.4.1, we have considered a stochastic equation with linearly-coupled Gaussian noise. After integration over the noise, this has resulted in the addition in equation (26.57) of a term of the form $w(\lambda) = -\frac{1}{2}w_{\alpha\beta}\lambda^\alpha\lambda^\beta$. In the notation (26.68), the addition leads to

$$S(\phi) = \int d\bar{\theta} d\theta \left[\frac{1}{2} \frac{\partial \phi^\alpha}{\partial \bar{\theta}} w_{\alpha\beta} \frac{\partial \phi^\beta}{\partial \theta} + A(\phi) \right]. \quad (26.71)$$

26.6 Gauge theories: Notation and algebraic structure

In what follows, we restrict the discussion to gauge fields denoted A_μ^α and scalar fields denoted φ^a , the addition of fermions being straightforward. In a concrete notation, we write an infinitesimal gauge transformation corresponding to a compact Lie group G as

$$\left\{ \begin{array}{l} \delta_\omega A_\mu^\alpha(x) = - \int d^d y \left[\partial_\mu^x \delta^{(d)}(x-y) \delta_\beta^\alpha \right. \\ \quad \left. + \int d^d z f_{\beta\gamma}^\alpha \delta^{(d)}(x-y) \delta^{(d)}(y-z) A_\mu^\gamma(z) \right] \omega^\beta(y), \\ \delta_\omega \varphi^a(x) = \int d^d y d^d z \delta^{(d)}(x-y) \delta^{(d)}(y-z) [\tau_\beta]^a_c (\varphi^c(y) + v^c) \omega^\beta(z), \end{array} \right. \quad (26.72)$$

where \mathbf{v} is a constant vector, $f_{\beta\gamma}^\alpha$ are the structure constants of the Lie algebra of G , the matrices τ_β are generators of a (in general reducible) representation of the Lie algebra and ω^α parametrizes the gauge transformation.

Abstract notation. We now combine all gauge fields and all scalar fields into one vector denoted by A^i , in which the index i stands for space variables x , Lorentz index μ , and group indices a, b :

$$A^i \equiv \{A_\mu^\alpha(x), \varphi^a(x)\}.$$

We write the corresponding infinitesimal gauge transformation (26.72) as

$$\delta_\omega A^i = D_\alpha^i(A) \omega^\alpha, \quad (26.73)$$

in which ω^α are the parameters of the gauge transformation. We specialize Greek indices correspond to the adjoint representation, but note that the index α now includes group indices and space variables.

The function $D_\alpha^i(A)$ is an affine function of the fields of the form

$$D_\alpha^i(A) = \Lambda_\alpha^i + [\mathbf{t}_\alpha]^i_j A^j, \quad (26.74)$$

where the matrices \mathbf{t}_α are generators of the Lie algebra.

Finally, the action $\mathcal{S}(A)$ is gauge invariant, and thus satisfies

$$D_\alpha^i(A) \frac{\delta \mathcal{S}}{\delta A^i} = 0. \quad (26.75)$$

Lie algebra. The differential operators

$$\mathcal{D}_\alpha = D_\alpha^i(A) \frac{\delta}{\delta A^i}, \quad (26.76)$$

form an affine representation of the Lie algebra. Therefore (see also equation (13.7)),

$$[\mathcal{D}_\alpha, \mathcal{D}_\beta] = f_{\alpha\beta}^\gamma \mathcal{D}_\gamma. \quad (26.77)$$

More explicitly,

$$D_\alpha^j \frac{\delta D_\beta^i}{\delta A^j} - D_\beta^j \frac{\delta D_\alpha^i}{\delta A^j} = f_{\alpha\beta}^\gamma D_\gamma^i, \quad (26.78)$$

where $f_{\alpha\beta}^\gamma$ is proportional to the numerical structure constant $f_{\beta\gamma}^\alpha$.

These equations can also be interpreted as compatibility conditions for the system (26.75) considered as a set of functional differential equations for $\mathcal{S}(A)$. They are formally identical to equations (26.37) of Section 26.3.2, and equations (29.42) of Section 29.3.

Since here, the functional $D_\alpha^i(A)$ is only affine, it is also simple to verify the identities (26.78) by a direct calculation, using the commutation relations of the Lie algebra of G . However, note that, because $D_\alpha^i(A)$ is a differential operator, as equation (26.72) shows it is necessary to carefully keep track of the δ -functions. In particular, the structure constants $f_{\alpha\beta}^\gamma$ have a non-trivial dependence on space variables (equation (26.72))

$$f_{\alpha\beta}^\gamma \equiv f_{\alpha\beta}^\gamma \delta^{(d)}(x-y)\delta^{(d)}(y-z). \quad (26.79)$$

As a consequence, it is actually convenient to rewrite the identity (26.78) with the help of two functions $\omega_1^\alpha \equiv \omega_1^a(x)$ and $\omega_2^\alpha \equiv \omega_2^a(x)$ and the operator

$$\Delta(\omega) = \omega^\alpha \mathcal{D}_\alpha. \quad (26.80)$$

The identity (26.78) takes the form

$$[\Delta(\omega_1), \Delta(\omega_2)] = \Delta(\omega_{12}), \quad \text{with} \quad \omega_{12}^\gamma = f_{\alpha\beta}^\gamma \omega_1^\alpha \omega_2^\beta. \quad (26.81)$$

Let us, for instance, indicate here how this identity can be recovered for the gauge field part. We write $\Delta(\omega)$ in explicit notation as,

$$\Delta(\omega) = \int d^d x [\delta_\alpha^\beta \partial_\mu \omega^\alpha(x) + f_{\alpha\gamma}^\beta \omega^\alpha(x) A_\mu^\gamma(x)] \frac{\delta}{\delta A_\mu^\beta(x)}.$$

One then calculates the commutator explicitly. Using the antisymmetry of $f_{\alpha\gamma}^\beta$ and the Jacobi identity, one indeed recovers equation (26.81) in explicit form, with $\omega_{12}^\alpha(x) = f_{\gamma\beta}^\alpha \omega_1^\beta(x) \omega_2^\gamma(x)$.

26.7 Gauge theories: Quantization

To quantize the gauge theory, we generalize the strategy followed in Sections 22.3 and 22.4. We begin with a gauge-invariant action $\mathcal{S}(A)$. We quantize the field theory first in the temporal gauge, and then pass to a covariant gauge. For this purpose, we introduce an equation that determines a space-dependent group element \mathbf{g} ,

$$E_\alpha(\mathbf{g}) \equiv F_\alpha(A^\mathbf{g}) - \nu_\alpha = 0, \quad (26.82)$$

where $A^\mathbf{g}$ is the gauge transform of A by \mathbf{g} , $F_\alpha(A)$ is a polynomial and ν_α an arbitrary field belonging to the adjoint representation of the Lie algebra.

Gauge transformations define classes in field space, corresponding to fields and all their gauge transforms, that is, orbits of the gauge group. We assume that the equation intersects all gauge orbits once and thus has a unique solution for \mathbf{g} (at least for A small so that the equation can be solved perturbatively).

The variation of the equation in an infinitesimal gauge transformation of parameters ω^α takes the form

$$\delta E_\alpha(\mathbf{g}) = M_{\alpha\beta}(A^\mathbf{g}) \omega^\beta, \quad \text{with} \quad M_{\alpha\beta}(A) = \mathcal{D}_\beta F_\alpha(A). \quad (26.83)$$

26.7.1 Quantization and ST symmetry

Before introducing fermion fields \bar{C}, C , an auxiliary field λ , and for the Gaussian ‘noise’ measure

$$d\rho(\nu) = d\nu e^{-\nu_\alpha^2/2\xi},$$

one finds the partition function

$$\mathcal{Z} = \int [dA] \det[M(A)] \exp[-\mathcal{S}(A) - F_\alpha^2(A)/2\xi]. \quad (26.84)$$

In this representation, the field integral is invariant under the infinitesimal gauge transformations that translate $F_\alpha(A)$ (see Section 26.1.2):

$$\delta A^i = D_\alpha^i [M^{-1}(A)]^{\alpha\beta} \mu_\beta \quad (26.85)$$

and, thus,

$$\delta [F_\alpha(A)] = \frac{\delta F_\alpha}{\delta A^i} D_\beta^i [M^{-1}(A)]^{\beta\gamma} \mu_\gamma = \mu_\alpha. \quad (26.86)$$

This is the basis of the ST identities [283] in gauge theories and of the *first proof, as given by Lee and Zinn-Justin* [41], *of the renormalizability of non-Abelian gauge theories in the spontaneously broken phase.* However, the non-local character of the transformation explains the complexity of the proof.

26.7.2 Quantization and BRST symmetry

From the results derived in Sections 26.1 and 26.2, one infers the identity

$$1 = \int [dg d\bar{C} dC d\lambda] \exp[-\mathcal{S}_{\text{gauge}}(A^g, C, \bar{C}, \lambda, \nu)], \quad (26.87)$$

with

$$\mathcal{S}_{\text{gauge}}(A, C, \bar{C}, \lambda, \nu) = \lambda^\alpha (F_\alpha(A) - \nu_\alpha) - C^\alpha M_{\alpha\beta}(A) \bar{C}^\beta. \quad (26.88)$$

The fields \bar{C} and C are spinless fermion fields (‘ghost’ fields) introduced to represent $\det M$, and transforming under the adjoint representation of the group G [5].

We insert the identity (26.87) into the representation of the partition function in the temporal gauge, and change variables $A^g \mapsto A$. The change of variables has the form of a gauge transformation. The action $\mathcal{S}(A)$ is thus invariant. The dependence of $\mathcal{S}_{\text{gauge}}$ in g disappears, and only the temporal gauge condition remains g dependent. We have shown in Section 22.4 that the integration over g yields a constant.

In the new gauge, the partition function has the integral representation

$$\mathcal{Z} = \int [dA d\bar{C} dC d\lambda] \exp[-\mathcal{S}(A) - \mathcal{S}_{\text{gauge}}(A, C, \bar{C}, \lambda, \nu)]. \quad (26.89)$$

Since the result does not depend on the field ν (the noise in the terminology of Section 26.4.1), we integrate over ν with a measure $d\rho(\nu)$, which we do not need specifying more precisely here, but which, eventually, will be chosen local and Gaussian. Introducing $w(\lambda)$, the generating functional of connected ν correlation functions,

$$w(\lambda) = \ln \int [d\rho(\nu)] \exp(\lambda^\alpha \nu_\alpha), \quad (26.90)$$

one finally obtains

$$\mathcal{Z} = \int [dA d\bar{C} dC d\lambda] \exp[-\mathcal{S}(A, C, \bar{C}, \lambda)], \quad (26.91)$$

with

$$\mathcal{S}(A, C, \bar{C}, \lambda) = \mathcal{S}(A) - w(\lambda) + \lambda^\alpha F_\alpha(A) - C^\alpha M_{\alpha\beta}(A) \bar{C}^\beta. \quad (26.92)$$

The interpretation of the expression is the following. Because the action is gauge invariant, the degrees of freedom associated with gauge transformations have no dynamics. To quantize these degrees of freedom and to give a meaning to the field integral in the continuum, it is necessary to provide one. For the field and group element \mathbf{g} , we thus have introduced a random distribution. This distribution is somewhat arbitrary, and we must eventually prove that ‘physical results’ (we shall clarify later what we mean by physical results) do not depend on its choice.

26.7.3 BRST transformations

It follows from the general analysis of 26.2 that the final action (26.92) has a BRST symmetry [236]. Because the gauge condition (26.82) determines a group element, the BRST variation of the element and of \bar{C} take the forms (26.28) and (26.29). Since the infinitesimal gauge transformation of A^i is given by equation (26.73), we infer that the BRST transformation of $[A^{\mathbf{g}}]^i$ is

$$\delta_{\text{BRST}}[A^{\mathbf{g}}]^i = \varepsilon D_\alpha^i(A^{\mathbf{g}})\bar{C}^\alpha. \quad (26.93)$$

Changing $[A^{\mathbf{g}}]^i$ to A^i , we conclude that the complete BRST transformations have the form

$$\delta_{\text{BRST}}A^i = \varepsilon D_\alpha^i(A)\bar{C}^\alpha, \quad \delta_{\text{BRST}}\bar{C}^\alpha = \tfrac{1}{2}\varepsilon f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma, \quad (26.94a)$$

$$\delta_{\text{BRST}}C^\alpha = \varepsilon \lambda^\alpha, \quad \delta_{\text{BRST}}\lambda^\alpha = 0, \quad (26.94b)$$

where we recognize the transformations (26.29) and (26.30), except that the transformation (26.94a) of A^i has now the form of an infinitesimal gauge transformation. Using a more explicit notation, one verifies that these transformations are analogous to the BRST transformations described in Section 22.4. Again, as expected, the double BRST transforms δ_{BRST}^2 vanish: the BRST transform of $D_\alpha^i(A)\bar{C}^\alpha$ vanishes as a consequence of equation (26.78) and the transform of $f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma$ as a consequence of the Jacobi identity for structure constants.

Notation. For later purpose, it is convenient to set

$$D^i = D_\alpha^i(A)\bar{C}^\alpha. \quad (26.95)$$

26.7.4 BRST differential operator and BRST symmetry

The anticommuting differential operator

$$\mathcal{D}_0 = D^i \frac{\delta}{\delta A^i} + \tfrac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma \frac{\delta}{\delta \bar{C}^\alpha} + \lambda^\alpha \frac{\delta}{\delta C^\alpha}, \quad (26.96)$$

acting on functions of A, C, \bar{C}, λ , implements the BRST transformations δ_{BRST} . Since δ_{BRST}^2 vanishes, the BRST operator has also a vanishing square:

$$[\mathcal{D}_0]^2 = 0. \quad (26.97)$$

As expected from equation (26.50), the quantized action (26.92) has the decomposition,

$$\mathcal{S}(A, C, \bar{C}, \lambda) = \mathcal{S}(A) + \mathcal{D}_0 \Phi, \quad \text{with} \quad \Phi = C^\alpha \left[F_\alpha(A) - \frac{\delta}{\delta \lambda^\alpha} \int_0^1 ds w(s\lambda) \right]. \quad (26.98)$$

The gauge-invariant action $\mathcal{S}(A)$ is BRST symmetric (closed), because the BRST transformation of A^i is also a gauge transformation. The contribution $\mathcal{D}_0\Phi(A, C, \lambda)$, which has been generated by gauge fixing, is BRST closed because it is BRST exact. Therefore, $\mathcal{S}(A, C, \bar{C}, \lambda)$ is BRST closed:

$$\mathcal{D}_0\mathcal{S}(A, C, \bar{C}, \lambda) = 0. \quad (26.99)$$

λ integration. In the examples we meet in this chapter, the function $w(\lambda)$, as defined by equation (26.90), is quadratic in λ and thus the corresponding Gaussian integral can be performed. After integration, the new action is still BRST symmetric, the variation of C^α takes the form

$$\delta C^\alpha = \varepsilon \tilde{\mathbf{a}}^{\alpha\beta} F_\beta(A),$$

where $\tilde{\mathbf{a}}$ is a constant matrix. The BRST operator is still nilpotent, but its square only vanishes when the equation $F(A) = 0$ is satisfied. Therefore, the property (26.97), which simplifies all proofs, is not shared by all realizations of BRST transformations, and may require the introduction of additional auxiliary variables.

26.8 WT identities and ZJ equation

Although we are interested only in A -field correlation functions, to study the consequences of the BRST symmetry (26.94), it is necessary to introduce sources not only for fields, but also for the composite operators (quadratic functions of the fields) that appear in the BRST transformations, because we expect that they will be, in general, renormalized. We thus consider the generating functional

$$\begin{aligned} \mathcal{Z}(J, \eta, \bar{\eta}, K, L, \mu) = & \int [dAd\bar{C}dCd\lambda] \exp [-\mathcal{S}_{\text{qu}}(A, \bar{C}, C, \lambda, K, L, \mu) \\ & + J_i A^i + \bar{\eta}_\alpha C^\alpha + \bar{C}^\alpha \eta_\alpha], \end{aligned} \quad (26.100)$$

where (K_i , $\bar{\eta}_\alpha$ and η_α are anticommuting elements)

$$\begin{aligned} \mathcal{S}_{\text{qu}}(A, \bar{C}, C, \lambda, K, L, \mu) = & \mathcal{S}(A) - w(\lambda) + \lambda^\alpha F_\alpha(A) - C^\alpha M_{\alpha\beta} \bar{C}^\beta \\ & - K_i D^i - \frac{1}{2} L_\alpha f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma - \mu_\alpha \lambda^\alpha. \end{aligned} \quad (26.101)$$

Since

$$K_i D^i + \frac{1}{2} L_\alpha f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma + \mu_\alpha \lambda^\alpha = \mathcal{D}_0 [K_i A^i + L_\alpha \bar{C}^\alpha + \mu_\alpha C^\alpha],$$

\mathcal{S}_{qu} is also BRST closed. Note also that,

$$D_i = -\frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i}, \quad \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma = -\frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha}, \quad \lambda^\alpha = -\frac{\delta \mathcal{S}_{\text{qu}}}{\delta \mu_\alpha}. \quad (26.102)$$

Therefore, \mathcal{D}_0 can be rewritten as

$$\mathcal{D}_0 = - \left(\frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i} \frac{\delta}{\delta A^i} + \frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha} \frac{\delta}{\delta \bar{C}^\alpha} + \frac{\delta \mathcal{S}_{\text{qu}}}{\delta \mu_\alpha} \frac{\delta}{\delta C^\alpha} \right). \quad (26.103)$$

Actually, because λ^α is an integration field, the introduction of $\delta \mathcal{S}_{\text{qu}} / \delta \mu_\alpha$ is not required, as the discussion will show, but convenient.

The equation expressing the BRST symmetry of \mathcal{S}_{qu} then takes the quadratic form (the ZJ equation [40, 286, 287]),

$$-\mathcal{D}_0 \mathcal{S}_{\text{qu}} \equiv \frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i} \frac{\delta \mathcal{S}_{\text{qu}}}{\delta A^i} + \frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha} \frac{\delta \mathcal{S}_{\text{qu}}}{\delta \bar{C}^\alpha} + \frac{\delta \mathcal{S}_{\text{qu}}}{\delta \mu_\alpha} \frac{\delta \mathcal{S}_{\text{qu}}}{\delta C^\alpha} = 0. \quad (26.104)$$

The ZJ equation is the basic equation for the discussion of the renormalization of quantized non-Abelian gauge theories.

WT identities. We consider the expression

$$\begin{aligned} X &= \int [\mathrm{d}A \mathrm{d}\bar{C} \mathrm{d}C \mathrm{d}\lambda] \mathcal{D}_0 \exp [-\mathcal{S}_{\text{qu}} + J_i A^i + \bar{C}^\alpha \eta_\alpha + \bar{\eta}_\alpha C^\alpha] \\ &\equiv - \int [\mathrm{d}A \mathrm{d}\bar{C} \mathrm{d}C \mathrm{d}\lambda] \left(J_i \frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i} + \eta_\alpha \frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha} + \bar{\eta}_\alpha \frac{\delta \mathcal{S}_{\text{qu}}}{\delta \mu_\alpha} \right) \\ &\quad \times \exp [-\mathcal{S}_{\text{qu}} + J_i A^i + \bar{\eta}_\alpha C^\alpha + \bar{C}^\alpha \eta_\alpha]. \end{aligned} \quad (26.105)$$

One verifies that the expression can be rewritten as

$$X = \left(J_i \frac{\delta}{\delta K_i} + \eta_\alpha \frac{\delta}{\delta L_\alpha} + \bar{\eta}_\alpha \frac{\delta}{\delta \mu_\alpha} \right) \mathcal{Z}, \quad (26.106)$$

by differentiating explicitly the integrand in the field integral (26.100).

In expression (26.105), we then let \mathcal{D}_0 act on the left-hand side, integrating by parts. The result vanishes provided that, in \mathcal{D}_0 , one can commute all differentiations to the left. Since

$$\frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i} \frac{\delta}{\delta A^i} = \frac{\delta}{\delta A^i} \frac{\delta \mathcal{S}_{\text{qu}}}{\delta K_i} - \frac{\delta^2 \mathcal{S}_{\text{qu}}}{\delta A^i \delta K_i}, \quad \frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha} \frac{\delta}{\delta \bar{C}^\alpha} = \frac{\delta}{\delta \bar{C}^\alpha} \frac{\delta \mathcal{S}_{\text{qu}}}{\delta L_\alpha} - \frac{\delta^2 \mathcal{S}_{\text{qu}}}{\delta L_\alpha \delta \bar{C}^\alpha},$$

this condition implies that (definitions (26.74)),

$$\frac{\delta^2 \mathcal{S}_{\text{qu}}}{\delta A^i \delta K_i} = \mathrm{tr} t^\alpha, \quad \frac{\delta^2 \mathcal{S}_{\text{qu}}}{\delta L_\alpha \delta \bar{C}^\alpha} = f_{\alpha\beta}^\alpha,$$

which are traces of the generators of the Lie algebra, vanish, a condition satisfied by orthogonal representations and all compact groups with semi-simple algebra (but excludes some representations of Abelian group $U(1)$ group).

If these conditions are satisfied, \mathcal{Z} satisfies the equation

$$\left(J_i \frac{\delta}{\delta K_i} + \eta_\alpha \frac{\delta}{\delta L_\alpha} + \bar{\eta}_\alpha \frac{\delta}{\delta \mu_\alpha} \right) \mathcal{Z} = 0, \quad (26.107)$$

which, expanded, yields WT identities for correlation functions.

Similarly, the generating functional $\mathcal{W} = \ln \mathcal{Z}$ of connected correlation functions satisfies,

$$\left(J_i \frac{\delta}{\delta K_i} + \eta_\alpha \frac{\delta}{\delta L_\alpha} + \bar{\eta}_\alpha \frac{\delta}{\delta \mu_\alpha} \right) \mathcal{W} = 0. \quad (26.108)$$

We then perform the Legendre transformation,

$$\Gamma(A, \bar{C}, C, \lambda, K, L) + W(J, \eta, \bar{\eta}, \mu, K, L) = J_i A^i + \bar{\eta}_\alpha C^\alpha + \bar{C}^\alpha \eta_\alpha + \mu_\alpha \lambda^\alpha, \quad (26.109)$$

with

$$J_i = \frac{\delta \Gamma}{\delta A^i}, \quad \bar{\eta}_\alpha = -\frac{\delta \Gamma}{\delta C^\alpha}, \quad \eta_\alpha = \frac{\delta \Gamma}{\delta \bar{C}^\alpha}, \quad \mu_\alpha = \frac{\delta \Gamma}{\delta \lambda^\alpha}. \quad (26.110)$$

These relations imply

$$A^i = \frac{\delta\Gamma}{\delta K_i}, \quad \bar{C}_\alpha = -\frac{\delta\Gamma}{\delta\eta^\alpha}, \quad \eta_\alpha = \frac{\delta\Gamma}{\delta C^\alpha}, \quad \mu_\alpha = \frac{\delta\Gamma}{\delta\lambda^\alpha}. \quad (26.111)$$

As usual for the sources coupled to composite fields (see equation (7.65)),

$$\frac{\delta\Gamma}{\delta K_i} + \frac{\delta\mathcal{W}}{\delta K_i} = 0, \quad \frac{\delta\Gamma}{\delta L_\alpha} + \frac{\delta\mathcal{W}}{\delta L_\alpha} = 0. \quad (26.112)$$

Therefore, Γ , the generating functional of vertex (or one-particle-irreducible) functions, satisfies a first form of the ZJ equation,

$$\frac{\delta\Gamma}{\delta A^i} \frac{\delta\Gamma}{\delta K_i} + \frac{\delta\Gamma}{\delta \bar{C}^\alpha} \frac{\delta\Gamma}{\delta L_\alpha} - \lambda^\alpha \frac{\delta\Gamma}{\delta C^\alpha} = 0. \quad (26.113)$$

For later purpose, it is convenient to add to Γ a source term for λ^α :

$$\Gamma \mapsto \Gamma - \mu_\alpha \lambda^\alpha.$$

Then,

$$\lambda^\alpha = -\frac{\delta\Gamma}{\delta\mu_\alpha},$$

and the ZJ equation takes the equivalent homogeneous quadratic form (see equation (26.104)),

$$\frac{\delta\Gamma}{\delta A^i} \frac{\delta\Gamma}{\delta K_i} + \frac{\delta\Gamma}{\delta \bar{C}^\alpha} \frac{\delta\Gamma}{\delta L_\alpha} + \frac{\delta\Gamma}{\delta\mu_\alpha} \frac{\delta\Gamma}{\delta C^\alpha} = 0. \quad (26.114)$$

Because the BRST transformations (26.94) involve composite operators, which are renormalized, their form is not preserved by renormalization. By contrast, the WT identity (26.114), as well as the ZJ equation (26.104), are quadratic, and the BRST transformations are only implicit (a property shared by the non-linear σ -model, see Chapter 19). As we will prove, these equations are stable under renormalization.

Regularization. We have already discussed the regularization of non-Abelian gauge theories (Section 22.5). While it is always possible to find gauge-invariant regularizations for theories with gauge and scalar fields, in the case of fermions and chiral gauge transformations, the corresponding property is not necessarily true, and this may be the source of anomalies (Section 23.6), and obstruction to gauge symmetry.

26.9 Renormalization: General considerations

We now prove that the ZJ equation (26.104), unlike the structure (26.101), is stable under renormalization.

We assume that the local action (26.101) has been regularized in a way consistent with gauge invariance. Perturbation theory then exhibits UV divergences that have to be removed by adding counter-terms to the action. The identities (26.113) imply relations among divergences. We use them to prove that counter-terms can be chosen in a way that preserves equation (26.104). In the next section, we then solve equation (26.104) to determine the most general form of the renormalized action (for other articles or lectures dealing with renormalization see *e.g.* Refs. [288]).

26.9.1 Counter-terms and ZJ equation

Our analysis is based on a loop expansion of the regularized generating functional Γ of vertex functions (defined in equation (26.109)),

$$\Gamma = \sum_{\ell=0}^{\infty} \Gamma_\ell. \quad (26.115)$$

For Γ , we use the homogeneous form (26.114) of the ZJ equation, which we write, symbolically, as

$$\Gamma * \Gamma = 0. \quad (26.116)$$

Expanding equation (26.116), we infer that the first term Γ_0 , which is the tree approximation, identical to the classical action \mathcal{S} , satisfies also the ZJ equation (26.104),

$$\Gamma_0 * \Gamma_0 = \mathcal{S} * \mathcal{S} = 0. \quad (26.117)$$

Renormalization. We assume, as an induction hypothesis, that we have been able to define a renormalized action $\mathcal{S}_{\ell-1}$ that satisfies equation (26.117), and renders Γ finite up to $(\ell - 1)$ loop order. We then express the consequences of equation (26.116) at loop order ℓ :

$$\mathcal{S} * \Gamma_\ell + \Gamma_\ell * \mathcal{S} = - \sum_{m=1}^{\ell-1} \Gamma_m * \Gamma_{\ell-m}. \quad (26.118)$$

The right-hand side of the equation is finite by induction. We define $\Gamma_\ell^{\text{div.}}$ as the sum of the divergent terms in the asymptotic expansion of Γ_ℓ in terms of the regularizing parameter. The divergent part of Γ_ℓ , which is local because all Feynman sub-diagrams have been renormalized, then satisfies (see also the example of the non-linear σ -model, Section 19.6).

$$\mathcal{S} * \Gamma_\ell^{\text{div.}} + \Gamma_\ell^{\text{div.}} * \mathcal{S} = 0. \quad (26.119)$$

Equation (26.119) shows that by defining the ℓ -loop renormalized action \mathcal{S}_ℓ by

$$\mathcal{S}_\ell = \mathcal{S}_{\ell-1} - \Gamma_\ell^{\text{div.}} + \text{higher orders}, \quad (26.120)$$

and adjusting the higher orders, it is possible to render Γ ℓ -loop finite with a renormalized action still satisfying equation (26.117). Therefore, we can construct a renormalized action that satisfies equation (26.117).

26.10 The renormalized gauge action

We have shown that the renormalized action satisfies the ZJ equation (26.104). The renormalized action is thus the most general solution of equation (26.104), local in the fields and sources, and consistent with symmetries and power counting. We work in four dimensions and assume a renormalizable gauge.

Ghost number conservation. In the original action, only the product $C\bar{C}$ appears. This implies a property of ghost number conservation. If we assign a ghost number $+1$ to \bar{C} and -1 to C , then K_i has a ghost number -1 , and $L_\alpha - 2$.

26.10.1 General renormalizable gauges: Power counting

In four dimensions, the Lagrangian density has dimension 4. We choose the gauge fixing term in such a way that the field A has the minimal dimension, equal to 1, and $F(A)$ has dimension 2 (we have exhibited such gauges in Chapters 21 and 22):

$$[A] = 1, \quad [F(A)] = 2.$$

We impose $[\lambda] = 2$, by choosing $w(\lambda)$ such λ has a constant propagator:

$$w(\lambda) = \frac{1}{2}a_{\alpha\beta}\lambda^\alpha\lambda^\beta, \quad \text{with } a_{\alpha\beta} \text{ constants.} \quad (26.121)$$

The other dimensions follow. Since $D_\alpha^i(A)$ has a term linear in A and a constant part with one derivative, it has dimension 1. The operator $M_{\alpha\beta}$ then has dimension 2. This implies that

$$[C] + [\bar{C}] = 2.$$

By convention, we can choose $[\bar{C}] = 0$, $[C] = 2$. This implies

$$[K] = 3, \quad [L] = 4.$$

Summarizing, we find

$$[A] = 1, \quad [\bar{C}] = 0, \quad [C] = 2, \quad [\lambda] = 2, \quad [K] = 3, \quad [L] = 4. \quad (26.122)$$

Since K and L have dimensions 3 and 4, respectively, the renormalized action can contain at most terms linear in K and L . Since λ has dimension 2, the renormalized action is at most a second-degree polynomial in λ , and the coefficient of $\lambda^\alpha\lambda^\beta$ is a constant matrix.

26.10.2 From the ZJ equation back to BRST symmetry

To solve equation (26.104), we now parametrize the solution in a way reminiscent of the initial action (26.101), but it should be kept in mind that the action now includes counter-terms, and the parameters that appear may be different from those parametrizing the action (26.101).

Power counting implies that \mathcal{S} is an affine function of K and L , and ghost number conservation implies that only the combinations $K\bar{C}$ and $L\bar{C}^2$ can appear. We thus set

$$\mathcal{S}(A, \bar{C}, C, \lambda, K, L) = -K_i D_\alpha^i(A) \bar{C}^\alpha - \frac{1}{2}L_\alpha f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma + \mathcal{S}(A, \bar{C}, C, \lambda).$$

Moreover, power counting implies that $f_{\beta\gamma}^\alpha$ has dimension 0, and is thus field independent, and D_α^i has dimension 1 and can, therefore, only be an affine function of A .

The terms linear in L and K in equation (26.104) yield, respectively:

$$f_{\alpha\gamma}^\beta f_{\delta\varepsilon}^\alpha \bar{C}^\gamma \bar{C}^\delta \bar{C}^\varepsilon = 0, \quad (26.123)$$

$$\left(\frac{\delta D_\alpha^i}{\delta A^j} D_\beta^j - \frac{1}{2}f_{\alpha\beta}^\gamma D_\gamma^i \right) \bar{C}^\alpha \bar{C}^\beta = 0. \quad (26.124)$$

The first equation implies that the constants $f_{\beta\gamma}^\alpha$ satisfy a Jacobi identity, because the product $\bar{C}^\gamma \bar{C}^\delta \bar{C}^\varepsilon$ is antisymmetric in $(\gamma, \delta, \varepsilon)$. Therefore, they are structure constants of a Lie algebra. It remains to show that the Lie algebra has not changed. This is straightforward when the gauge condition does not break the global symmetry. Otherwise, when the original algebra is semi simple, this follows from a continuity argument. Finally, in the general case, one can still use the gauge independence of physical observables as derived in Section 26.11.

The structure constants are thus linearly related to the structure constants appearing in the initial action (26.101).

Equation (26.123) is also an integrability condition for equation (26.124), which implies the commutation relations,

$$\frac{\delta D_\alpha^i}{\delta A^j} D_\beta^j - \frac{\delta D_\beta^i}{\delta A^j} D_\alpha^j = f_{\alpha\beta}^\gamma D_\gamma^i. \quad (26.125)$$

Therefore, the Lie algebra structure is recovered. In terms of the BRST operator,

$$\mathcal{D}_+ \equiv D_\alpha^i(A) \bar{C}^\alpha \frac{\delta}{\delta A_i} - \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma \frac{\delta}{\delta \bar{C}^\alpha}, \quad (26.126)$$

the two equations can be combined into a unique equation $\mathcal{D}_+^2 = 0$. The ZJ equation then reduces to a condition of BRST symmetry. Defining also

$$\mathcal{D}_- \equiv \lambda^\alpha \frac{\delta}{\delta C^\alpha}, \quad (26.127)$$

we can write it as

$$\mathcal{D}\mathcal{S}(A, \bar{C}, C, \lambda) \equiv (\mathcal{D}_+ + \mathcal{D}_-) \mathcal{S}(A, \bar{C}, C, \lambda) = 0, \quad (26.128)$$

with

$$\mathcal{D}_+^2 = \mathcal{D}_-^2 = 0, \quad \mathcal{D}_+ \mathcal{D}_- + \mathcal{D}_- \mathcal{D}_+ = 0. \quad (26.129)$$

26.10.3 BRST symmetry and renormalized action

Because \mathcal{D}_+ has degree 1 in \bar{C} and \mathcal{D}_- of degree 0, it is natural to expand $\mathcal{S}(A, \bar{C}, C, \lambda)$ in powers of \bar{C} , a situation we also meet in Section A26.3. Since the product $\bar{C}C$ has dimension 2, \mathcal{S} is a polynomial of degree 2 in $\bar{C}C$, and we can write it as,

$$\mathcal{S}(A, \bar{C}, C, \lambda) = \mathcal{S}^{(0)}(A, \lambda) + \mathcal{S}^{(1)}(A, \bar{C}, C, \lambda) + \mathcal{S}^{(2)}(\bar{C}, C).$$

Moreover, λ has dimension 2 and, therefore, $\mathcal{S}^{(0)}$ is a polynomial of degree 2 in λ , $\mathcal{S}^{(1)}$ of degree 1 and $\mathcal{S}^{(2)}$ is both λ and A independent. The $\bar{C}C$ independent term $\mathcal{S}^{(0)}$ can be parametrized as

$$\mathcal{S}^{(0)}(A, \lambda) = -\frac{1}{2} a_{\alpha\beta} \lambda^\alpha \lambda^\beta + \lambda^\alpha F_\alpha(A) + \mathcal{S}(A). \quad (26.130)$$

The equation $\mathcal{D}_- \mathcal{S}^{(0)} = 0$ is automatically satisfied. We note that $\mathcal{S}^{(0)}$ can be written as

$$\mathcal{S}^{(0)}(A, \lambda) = \mathcal{D}_- \Phi^{(0)} + \mathcal{S}(A), \quad (26.131)$$

with

$$\Phi^{(0)} = -\frac{1}{2} C^\alpha a_{\alpha\beta} \lambda^\beta + C^\alpha F_\alpha(A). \quad (26.132)$$

More generally, it is possible to verify for any function Σ polynomial in λ ,

$$\mathcal{D}_- \Sigma(A, \bar{C}, C, \lambda) = 0 \Rightarrow \Sigma(A, \bar{C}, C, \lambda) = \mathcal{D}_- \tilde{\Sigma}(A, \bar{C}, C, \lambda) + \Sigma(A, 0, 0, 0).$$

The other equations can be written as

$$\mathcal{D}_- \mathcal{S}^{(n+1)} = -\mathcal{D}_+ \mathcal{S}^{(n)}.$$

The equation $n = 0$ for $\lambda = 0$ reduces to

$$\bar{C}^\alpha D_\alpha^i(A) \frac{\delta \mathcal{S}(A)}{\delta A^i} = 0. \quad (26.133)$$

This equation implies that $\mathcal{S}(A)$ is gauge invariant.

Then, from equation (26.131) follows,

$$\mathcal{D}_- \mathcal{S}^{(1)} = -\mathcal{D}_+ \mathcal{D}_- \Phi^{(0)} = \mathcal{D}_- \mathcal{D}_+ \Phi^{(0)}.$$

The general solution is the sum of a special solution, and a \mathcal{D}_- exact term

$$\mathcal{S}^{(1)} = \mathcal{D}_+ \Phi^{(0)} + \mathcal{D}_- \Phi^{(1)}, \quad (26.134)$$

where $\Phi^{(1)}$ is proportional to $C^2 \bar{C}$:

$$\Phi^{(1)} = \frac{1}{2} g_{\alpha\beta\gamma} C^\alpha C^\beta \bar{C}^\gamma, \quad g_{\alpha\beta\gamma} = -g_{\beta\alpha\gamma}. \quad (26.135)$$

Power counting implies that $g_{\beta\gamma\delta}$ is a constant.

The next equation becomes

$$\mathcal{D}_- \mathcal{S}^{(2)} = -\mathcal{D}_+ \mathcal{D}_- \Phi^{(1)} = \mathcal{D}_- \mathcal{D}_+ \Phi^{(1)},$$

whose solution can be written as

$$\mathcal{S}^{(2)} = \mathcal{D}_+ \Phi^{(1)}, \quad (26.136)$$

because power counting prevents the addition of another \mathcal{D}_- exact term. Finally, we note that the last equation $\mathcal{D}_+ \mathcal{S}^{(2)}$ is automatically satisfied. Summing all contributions, we find that the renormalized action can be written as (for an interpretation of this representation see also equation (A26.17))

$$\mathcal{S}(A, \bar{C}, C, \lambda) = \mathcal{S}(A) + \mathcal{D}\Phi(A, \bar{C}, C, \lambda), \quad (26.137)$$

where $\mathcal{S}(A)$ is gauge invariant and

$$\Phi(A, \bar{C}, C, \lambda) = \Phi^{(0)} + \Phi^{(1)} = -\frac{1}{2} C^\alpha a_{\alpha\beta} \lambda^\beta + C^\alpha F_\alpha(A) + \frac{1}{2} g_{\alpha\beta\gamma} C^\alpha C^\beta \bar{C}^\gamma. \quad (26.138)$$

The renormalized action has a form similar to the bare action (26.101) except for the possible additional BRST exact term,

$$\mathcal{S}_4(\lambda, \bar{C}, C) = \mathcal{D}\Phi^{(1)} = g_{\alpha\beta\gamma} \left(-\frac{1}{4} C^\alpha C^\beta f_{\delta\varepsilon}^\gamma \bar{C}^\delta \bar{C}^\varepsilon + \lambda^\alpha C^\beta \bar{C}^\gamma \right). \quad (26.139)$$

This term has exactly the form given in equation (A26.15) and corresponds to the shift $L_\alpha \mapsto L_\alpha + \frac{1}{2} g_{\beta\gamma\alpha} C^\beta C^\gamma$.

26.10.4 Comments

The quartic ghost term. A comment now is in order: since in general the renormalized action is quartic in the ghost terms [40], in contrast to the initial action, the direct interpretation of the ghost integral as representing a determinant in local form is lost. However, the following result can be proven: if one adds a term depending on an auxiliary scalar field, transforming non-trivially under the gauge group, to the gauge function F_α , then the integration over this auxiliary field with an appropriate Gaussian weight generates the quartic ghost terms in their most general form [289].

Renormalization of gauge-invariant operators [290–293]. To generate correlation functions with operator insertions, one can add sources for them in the action. If the dimension of the gauge-invariant operators is at most 4, the new action is still renormalizable. The general analysis is not modified; the only difference is that some coupling constants are now space-dependent. In the case of operators of higher dimensions, the action with sources is no longer renormalizable. It is still possible to renormalize it at any finite order by introducing enough renormalization constants. The determination of the general form of the renormalized action, that is, the solution of equation (26.104) is a non-trivial problem and requires more work, based on cohomology techniques [290, 292]. In the case of compact Lie groups with semi-simple Lie algebras, the most general solution of equation (A26.7) is the sum of gauge-invariant terms and BRST exact contributions, that is, of the form $\mathcal{D}\Phi$. This result, first conjectured, has been rigorously proven. The part concerning C, λ is simple but the difficulties come from the set $\{A, \bar{C}, K, L\}$. Note that the form of the renormalized operators, when inserted in field correlation functions, depends on the explicit gauge. Only the expectation values of products of gauge-invariant operators, or the S -matrix elements between physical states, as we show in Section 26.11, are gauge independent, and thus physical.

26.10.5 Linear gauges

For the special class of gauges in which the gauge-fixing function $F_\alpha(A)$ is linear in A ,

$$F_\alpha(A) = F_{\alpha i} A^i, \quad (26.140)$$

the general analysis can be simplified. Then, the operator $F_\alpha(A)$ is in general still of dimension 2, but its correlation functions are now directly related to the correlation functions of the A^i field and, therefore, introduce no new independent renormalization.

To derive the consequences of equation (26.140), we use the λ -field equation of motion. Again, we parametrize explicitly $w(\lambda)$ as

$$w(\lambda) = \frac{1}{2} a_{\alpha\beta} \lambda^\alpha \lambda^\beta. \quad (26.141)$$

Then, the λ -field equation of motion, which relies on the identity

$$\int [d\lambda] \frac{\delta}{\delta \lambda^\alpha} [\exp(-\mathcal{S}(A, C, \bar{C}, \lambda, K, L) + \text{sources})] = 0,$$

with \mathcal{S} given by equation (26.101), for \mathcal{Z} and $\mathcal{W} = \ln \mathcal{Z}$ reads

$$\left(-a_{\alpha\beta} \frac{\delta}{\delta l_\beta} + F_{\alpha i} \frac{\delta}{\delta J_i} - \mu_\alpha \right) \mathcal{Z} = 0 \quad \Rightarrow \quad \left(a_{\alpha\beta} \frac{\delta}{\delta l_\beta} + F_{\alpha i} \frac{\delta}{\delta J_i} \right) \mathcal{W} = \mu_\alpha. \quad (26.142)$$

Since $F(A)$ is linear in A , the λ -field equation of motion is a first order differential equation and its implication for the generating functional of vertex functions Γ is simply,

$$\frac{\delta \Gamma}{\delta \lambda^\alpha} = -a_{\alpha\beta} \lambda^\beta + F_{\alpha i} A^i. \quad (26.143)$$

Equation (26.143) is satisfied by the action \mathcal{S} , and clearly is stable under renormalization. It implies that the quadratic and linear parts in λ of the action are unrenormalized. In particular, no term of the form $g_{\alpha\beta\gamma} C^\alpha C^\beta \bar{C}^\gamma$ is generated in expression (26.138). The action remains quadratic in the ghost fields. The renormalized action takes the simple form

$$\mathcal{S}(A, \bar{C}, C, \lambda, K, L) = \mathcal{S}(A) + \mathcal{D} [C^\alpha F_{\alpha i} A^i - \frac{1}{2} a_{\alpha\beta} C^\alpha \lambda^\beta + K_i A_i + L_\alpha \bar{C}^\alpha], \quad (26.144)$$

(\mathcal{D} being defined by equation (26.128)) where, in addition, as explained previously, $a_{\alpha\beta}$ and $F_{\alpha i}$ are unrenormalized.

The C-ghost equation of motion. In the case of linear gauges, another equation can be used to prove that the gauge function is unrenormalized, the C -ghost equation of motion. The equation

$$\frac{\delta \mathcal{S}}{\delta C^\alpha} = F_{\alpha i} D_\beta^i \bar{C}^\beta \equiv F_{\alpha i} \frac{\delta \mathcal{S}}{\delta K_i}, \quad (26.145)$$

implies

$$\left(F_{\alpha i} \frac{\delta}{\delta K_i} + \eta_\alpha \right) \mathcal{Z} = 0, \quad (26.146)$$

and, after Legendre transformation,

$$\left(F_{\alpha i} \frac{\delta}{\delta K_i} - \frac{\delta}{\delta C^\alpha} \right) \Gamma = 0. \quad (26.147)$$

Equation (26.145) thus is stable under renormalization.

This alternative proof, in the case of linear gauges, can be used even when the λ -field has been integrated out. Then equation (26.143) disappears, while equation (26.147) remains, and can be used to show that the gauge-fixing term $F(A)$ is not renormalized.

26.11 Gauge independence: Physical observables

General correlation functions are gauge dependent and, therefore, do not correspond directly to physical observables. By contrast, we now show that expectation values of products of gauge-invariant operators (local polynomials in the fields) and S -matrix elements are unaffected by infinitesimal changes of gauges. This establishes gauge independence, at least for gauges which can be continuously connected, and confirms that expectation values of gauge-invariant operators and S -matrix elements are physical observables.

Gauge-invariant operators. To generate correlation functions of gauge-invariant operators, we add source terms for them to the action. The action with these sources is still gauge invariant (equation (26.75)), and the effective action BRST symmetric. Assuming a gauge-invariant regularization, we examine how the partition function (or vacuum amplitude) is affected by an infinitesimal change of gauge δF_α , before renormalization.

Since the part of the action generated by the quantization procedure is BRST exact, that is, has the general form $\mathcal{D}_0 \Phi$, in which \mathcal{D}_0 has been defined by equation (26.96) (see equation (26.98)), the variation $\delta \mathcal{S}$ of the action due to a change in the gauge-fixing function takes the form $\delta \mathcal{S} = \mathcal{D}_0(\delta \Phi)$, and thus the variation of the partition function is

$$\delta \mathcal{Z} = - \int [dA d\bar{C} dC d\lambda] \mathcal{D}_0(\delta \Phi) e^{-\mathcal{S}}. \quad (26.148)$$

The operator \mathcal{D}_0 is a differential operator and, therefore, we can integrate by parts. Using again the property that the traces $f_{\alpha\beta}^\alpha$ and $\delta D_\alpha^i / \delta A^i$ vanish, we obtain

$$\delta\mathcal{Z} = \int [dAd\bar{C}dCd\lambda]\delta\Phi(\mathcal{D}_0\mathcal{S})e^{-S}, \quad (26.149)$$

an expression that vanishes as a consequence of the BRST symmetry.

Therefore, we have shown that the bare correlation functions of gauge-invariant operators are gauge independent, at least within a class of gauges that can be continuously connected. Moreover, a renormalization procedure exists which produces gauge-independent renormalized correlation functions of these operators.

These correlation functions contain the complete information about the physical properties of the gauge theory.

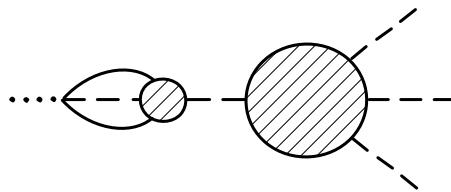


Fig. 26.1 One-line reducible contribution to the S -matrix

S-matrix elements. We now want to prove the gauge independence of the perturbative S -matrix, when it exists. We first calculate the variation of renormalized correlation functions under an infinitesimal change of gauge. We assume that we have renormalized the theory in a given gauge, but not yet eliminated the regularizing parameter.

We then proceed as previously mentioned. The variation of the action in an infinitesimal change of gauge has exactly the form $\mathcal{D}\delta\Phi$ (\mathcal{D} being now the renormalized BRST operator) considered in Section 26.10. We can still integrate by parts, but the resulting integrand does not vanish identically, because we have introduced sources for non-gauge-invariant fields:

$$\delta\mathcal{Z}(J) = - \int [dAd\bar{C}dCd\lambda]\delta\Phi J_i D_\alpha^i(A)\bar{C}^\alpha \exp(-S(A, C, \bar{C}, \lambda) + J_i A_i). \quad (26.150)$$

When we apply this result to correlation functions, we obtain a sum of contributions in which one A_i field has been replaced by $\delta\Phi J_i D_\alpha^i(A)\bar{C}^\alpha$, which is a linear combination of composite operators.

In the Fourier representation, we denote by p_i the external momenta and m_i the corresponding particle masses. If we multiply a contribution by $\prod_i(p_i^2 + m_i^2)$ and take the mass-shell limit $p_i^2 \rightarrow -m_i^2$, we find a non-vanishing result only if there is a pole in each external momentum squared. For a composite operator, this happens only if the line is one-particle reducible (see Fig. 26.1). Then, on the mass-shell, we obtain a contribution proportional to the matrix element of the field itself. This argument was presented for the first time in Section A7.2. The final result is that an infinitesimal change of gauge renormalizes multiplicatively the S -matrix elements. This corresponds to a field renormalization. Therefore, the S -matrix, suitably normalized, is gauge independent (see also Ref. [294]).

A26 BRST symmetry and ZJ equation: Additional remarks

We first exhibit a few general properties of the BRST symmetry and the ZJ equation (see also Ref. [287]). Moreover, we show in Section A26.3, with a simple example, how to solve the equation $\mathcal{D}_{\text{BRST}} = 0$, by simple BRST cohomology techniques.

A26.1 BRST symmetry and ZJ equation

A large class of BRST transformations can be generated in the following way: we consider a set of commuting variables x_i , and to each variable x_i we associate a Grassmann generator θ_i . We introduce a function $\mathcal{S}(\theta_i, x_i)$, a commuting element of the corresponding Grassmann algebra (sufficiently differentiable in the variables x_i). We define the BRST-like transformations,

$$\delta_{\text{BRST}} x_i = \varepsilon \frac{\partial \mathcal{S}}{\partial \theta_i}, \quad \delta_{\text{BRST}} \theta_i = \varepsilon \frac{\partial \mathcal{S}}{\partial x_i}, \quad (\text{A26.1})$$

where ε is an additional generator of the Grassmann algebra. Acting on functions of the set $\{\theta_i, x_i\}$, the transformations can be implemented by the differential operator,

$$\mathcal{D} = \frac{\partial \mathcal{S}}{\partial \theta_i} \frac{\partial}{\partial x_i} + \frac{\partial \mathcal{S}}{\partial x_i} \frac{\partial}{\partial \theta_i}. \quad (\text{A26.2})$$

We impose that δ_{BRST}^2 vanishes, a condition that translates into $\mathcal{D}^2 = 0$.

Since \mathcal{D} is of anticommuting type, only the terms generated by the non-commutation of $\partial \mathcal{S}/\partial \theta_i$ and $\partial \mathcal{S}/\partial x_i$ with the differential operators $\partial/\partial \theta_i$ and $\partial/\partial x_i$ survive in \mathcal{D}^2 . Then,

$$\begin{aligned} \mathcal{D}^2 &= \left[\left(\frac{\partial \mathcal{S}}{\partial \theta_i} \frac{\partial^2 \mathcal{S}}{\partial x_i \partial \theta_j} + \frac{\partial \mathcal{S}}{\partial x_i} \frac{\partial^2 \mathcal{S}}{\partial \theta_i \partial \theta_j} \right) \frac{\partial}{\partial x_j} \right. \\ &\quad \left. + \left(\frac{\partial \mathcal{S}}{\partial \theta_i} \frac{\partial^2 \mathcal{S}}{\partial x_i \partial x_j} + \frac{\partial \mathcal{S}}{\partial x_i} \frac{\partial^2 \mathcal{S}}{\partial \theta_i \partial x_j} \right) \frac{\partial}{\partial \theta_j} \right] = 0. \end{aligned} \quad (\text{A26.3})$$

The condition can be rewritten as

$$\mathcal{D}^2 = \left[-\frac{\partial}{\partial \theta_j} \left(\frac{\partial \mathcal{S}}{\partial \theta_i} \frac{\partial \mathcal{S}}{\partial x_i} \right) \right] \frac{\partial}{\partial x_j} + \left[\frac{\partial}{\partial x_j} \left(\frac{\partial \mathcal{S}}{\partial \theta_i} \frac{\partial \mathcal{S}}{\partial x_i} \right) \right] \frac{\partial}{\partial \theta_j} = 0. \quad (\text{A26.4})$$

The condition is satisfied, and \mathcal{D} then is a *cohomology operator*, if \mathcal{S} satisfies a general ZJ equation [40],

$$\frac{\partial \mathcal{S}}{\partial x_i} \frac{\partial \mathcal{S}}{\partial \theta_i} = 0. \quad (\text{A26.5})$$

We assume the θ_i 's exhaust all the generators of the Grassmann algebra. Then, a possible constant, which is an odd element, vanishes. Thus, \mathcal{S} itself is *BRST closed*.

Equation (A26.1) is the general form of BRST transformations when some integrability condition is satisfied. This condition is not satisfied by the transformation (26.17), but by all other implementations in the chapter. For example, the transformations (26.31), (26.29) and (26.33) can be generated by

$$\mathcal{S}(c_i, \bar{c}_i, a_i, \lambda_i) = \frac{1}{2} \lambda_i \lambda_i + \frac{1}{2} a_i f_{jk}^i \bar{c}_j \bar{c}_k, \quad (\text{A26.6})$$

because the f_{jk}^i are structure constants.

Equation (A26.5) is the general form of the ZJ equation.

The integration measure. The transformations (A26.1) leave the integration measure $\prod_i d\theta_i dx_i$ invariant. Indeed, the Jacobian factorizes into the product of two Jacobians (see equation (1.63)),

$$J_x = \det \left(\delta_{ij} + \varepsilon \frac{\partial \mathcal{S}}{\partial \theta_i x_j} \right) = 1 + \varepsilon \frac{\partial \mathcal{S}}{\partial \theta_i \partial x_i}, \quad J_\theta = \det^{-1} \left(\delta_{ij} + \varepsilon \frac{\partial \mathcal{S}}{\partial \theta_j x_i} \right) = 1 - \varepsilon \frac{\partial \mathcal{S}}{\partial \theta_i \partial x_i},$$

(equation (1.59)) where the identity $\det(1 + \varepsilon X) = 1 + \varepsilon \text{tr } X$ is used. Thus, $J_x J_\theta = 1$.

Correspondence with Section 26.8. We denote by θ_i the set of all anticommuting fields K_i , \bar{C}^α , C^α , and x_i all commuting fields A^i , L_α , μ_α . Using the explicit expression (26.104) of the action, we note that the ZJ equation for the action \mathcal{S} (and thus Γ) has the expression (A26.5). Equation (26.119) reads

$$\mathcal{D}\Gamma_l^{\text{div.}} = 0, \quad (\text{A26.7})$$

where \mathcal{D} is the differential operator (A26.2), or, in more explicit notation,

$$\mathcal{D} = \frac{\delta \mathcal{S}}{\delta A^i} \frac{\delta}{\delta K_i} + \frac{\delta \mathcal{S}}{\delta K_i} \frac{\delta}{\delta A^i} + \frac{\delta \mathcal{S}}{\delta \bar{C}^\alpha} \frac{\delta}{\delta L_\alpha} + \frac{\delta \mathcal{S}}{\delta L_\alpha} \frac{\delta}{\delta \bar{C}^\alpha} + \frac{\delta \mathcal{S}}{\delta \mu_\alpha} \frac{\delta}{\delta C^\alpha} + \frac{\delta \mathcal{S}}{\delta C^\alpha} \frac{\delta}{\delta \mu_\alpha}. \quad (\text{A26.8})$$

A26.2 Canonical invariance of the ZJ equation

Equation (A26.5) has properties analogous to the symplectic form $dp \wedge dq$ of classical mechanics, it is invariant under a different form of canonical transformations.

We change variables $(\theta, x) \mapsto (\theta', x')$, setting

$$x_i = \frac{\partial \varphi}{\partial \theta_i}(\theta, x'), \quad (\text{A26.9})$$

$$\theta'_i = \frac{\partial \varphi}{\partial x'_i}(\theta, x'), \quad (\text{A26.10})$$

in which $\varphi(x', \theta)$ is an odd element of the Grassmann algebra. We first eliminate x_i in equation (A26.5) using equation (A26.9),

$$\frac{\partial \mathcal{S}}{\partial \theta_i} \left[\frac{\partial \varphi}{\partial \theta_i \partial x'_j} \right]^{(-1)} \frac{\partial \mathcal{S}}{\partial x'_j} = 0. \quad (\text{A26.11})$$

We then eliminate θ_i using equation (A26.10). It is simple to verify that one recovers (A26.5) equation in the new variables:

$$\frac{\partial \mathcal{S}}{\partial \theta'_i} \frac{\partial \mathcal{S}}{\partial x'_i} = 0. \quad (\text{A26.12})$$

Infinitesimal canonical transformations. We consider a function φ of the transformations (A26.9) and (A26.10) close to the identity of the form

$$\varphi(\theta, x') = \theta_i x'_i + \kappa \psi(\theta, x'), \quad (\text{A26.13})$$

where κ is a real infinitesimal parameter. At order κ , the transformations (A26.9) and (A26.10) become,

$$x'_i = x_i - \kappa \frac{\partial \psi}{\partial \theta_i}(\theta, x) + O(\kappa^2), \quad \theta'_i = \theta_i + \kappa \frac{\partial \psi}{\partial x_i}(\theta, x) + O(\kappa^2). \quad (\text{A26.14})$$

Then, $\mathcal{S}(\theta', x')$ is given by

$$\mathcal{S}(\theta', x') - \mathcal{S}(\theta, x) = \kappa \frac{\partial \psi}{\partial x_i} \frac{\partial \mathcal{S}}{\partial \theta_i} - \kappa \frac{\partial \psi}{\partial \theta_i} \frac{\partial \mathcal{S}}{\partial x_i} + O(\kappa^2) = -\kappa \tilde{\mathcal{D}}\psi + O(\kappa^2), \quad (\text{A26.15})$$

where the definition (A26.2) has been used.

We infer that any infinitesimal addition to \mathcal{S} of a BRST exact term can be generated by an infinitesimal canonical transformation acting on \mathcal{S} . *The effect of the quantization procedure can be understood as resulting from an infinitesimal canonical transformation.*

More explicitly, in the action (26.101), the dependence on μ_α , which is an artificial variable, cannot change. This imposes the dependence on μ_α of the function φ in equations (A26.9) and (A26.10):

$$\varphi(A, \bar{C}, C, \lambda, K, L, \mu) = C^\alpha \mu_\alpha + \tilde{\varphi}(A, \bar{C}, C, \lambda, K, L).$$

It follows that the general change of variables is equivalent to a change induced by the function $\tilde{\varphi}$ on the restricted set $\{A, \bar{C}, K, L\}$ with, in addition the translation,

$$\mathcal{S} \mapsto \mathcal{S} + \lambda^\alpha \frac{\delta \tilde{\varphi}}{\delta C^\alpha}. \quad (A26.16)$$

It can be verified that the gauge-invariant action $\mathcal{S}(A)$ to which K and L source terms are added, and the renormalized quantized action (26.137) are related by such a transformation, with

$$\begin{aligned} \tilde{\varphi}(A, \bar{C}, C, \lambda, K, L) &= C^\alpha F_\alpha(A) - \frac{1}{2} a_{\alpha\beta} C^\alpha \lambda^\beta + A^i K_i + L_\alpha \bar{C}^\alpha \\ &\quad + \frac{1}{2} g_{\alpha\beta\gamma} C^\alpha C^\beta \bar{C}^\gamma. \end{aligned} \quad (A26.17)$$

A26.3 Elements of BRST cohomology

We have seen in Section 26.4.1 that, for a general stochastic equation linear in a Gaussian noise, the weight function Σ obtained after noise averaging (equations (26.53)) is BRST exact and quadratic in $\{\lambda, c\bar{c}\}$. Conversely, one may ask the question: what is the most general form of a function Σ quadratic in $\{\lambda, c\bar{c}\}$ and BRST symmetric? (A simpler version of the problem solved in Section 26.10.3.) We consider the equation

$$\mathcal{D}\Sigma(\varphi, \lambda, c, \bar{c}) = 0. \quad (A26.18)$$

We expand Σ in powers of $c\bar{c}$,

$$\Sigma(\varphi, \lambda, c, \bar{c}) = \Sigma_0(\varphi, \lambda) + \Sigma_1(\varphi, \lambda, c, \bar{c}) + \Sigma_2(\varphi, \lambda, c, \bar{c}), \quad (A26.19)$$

where Σ_0 is quadratic in λ , Σ_1 linear in $c\bar{c}$ and of first degree in λ , and Σ_2 quadratic in $c\bar{c}$ (and independent of λ).

We now express the BRST operator \mathcal{D} as the sum $\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_-$, with the definitions

$$\mathcal{D}_+ \equiv \bar{c}^\alpha \partial_\alpha, \quad \mathcal{D}_- \equiv \lambda^\alpha \frac{\partial}{\partial c^\alpha}. \quad (A26.20)$$

The operator \mathcal{D}_+ corresponds to form differentiation. One verifies

$$\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_-, \quad \mathcal{D}_+^2 = \mathcal{D}_-^2 = 0, \quad \mathcal{D}_+ \mathcal{D}_- + \mathcal{D}_- \mathcal{D}_+ = 0. \quad (A26.21)$$

Since \mathcal{D}_+ and \mathcal{D}_- differ by a factor $c\bar{c}$, the equation $\mathcal{D}\Sigma = 0$ decomposes into $\mathcal{D}_-\Sigma_0 = 0$, which is automatically satisfied, and

$$\mathcal{D}_- \Sigma_1 = -\mathcal{D}_+ \Sigma_0, \quad (A26.22a)$$

$$\mathcal{D}_- \Sigma_2 = -\mathcal{D}_+ \Sigma_1, \quad (A26.22b)$$

$$\mathcal{D}_+ \Sigma_2 = 0. \quad (A26.22c)$$

We parametrize Σ_0 as

$$\Sigma_0 = A(\varphi) + \lambda^\alpha E_\alpha(\varphi, \lambda) = A(\varphi) + \mathcal{D}_- [c^\alpha E_\alpha(\varphi, \lambda)].$$

Setting $\lambda = 0$ in equation (A26.22a), we first find

$$\partial_\alpha A(\varphi) = 0 \Rightarrow A(\varphi) = \text{const.}.$$

Then,

$$\mathcal{D}_- \Sigma_1 = -\mathcal{D}_+ \mathcal{D}_- [c^\alpha E_\alpha(\varphi, \lambda)] = \mathcal{D}_- \mathcal{D}_+ [c^\alpha E_\alpha(\varphi, \lambda)],$$

which has a solution

$$\Sigma_1 = \mathcal{D}_+ [c^\alpha E_\alpha(\varphi, \lambda)] + \mathcal{D}_- \tilde{\Sigma}_1,$$

up to possible terms that are not \mathcal{D}_- exact.

We examine this possibility, taking into account that Σ_1 has the general form,

$$\Sigma_1 = w_{\alpha\beta} c^\alpha \bar{c}^\beta + w_{\alpha,\beta\gamma} \lambda^\alpha c^\beta \bar{c}^\gamma.$$

The equation $\mathcal{D}_- \Sigma_1 = 0$ implies $w_{\alpha\beta} = 0$ and $w_{\alpha,\beta\gamma} = -w_{\beta,\alpha\gamma}$, which are the conditions for Σ_1 to be \mathcal{D}_- exact, with

$$\tilde{\Sigma}_1 = \frac{1}{2} w_{\alpha\beta,\gamma} c^\alpha c^\beta \bar{c}^\gamma.$$

Then equation (A26.22b) becomes

$$\mathcal{D}_- \Sigma_2 = -\mathcal{D}_+ \mathcal{D}_- \tilde{\Sigma}_1 = \mathcal{D}_- \mathcal{D}_+ \tilde{\Sigma}_1,$$

which, taking into account that Σ_2 does not depend on λ , has the solution

$$\Sigma_2 = \mathcal{D}_+ \tilde{\Sigma}_1,$$

which automatically satisfies the last equation $\mathcal{D}_+ \Sigma_2 = 0$.

Summing all contributions, we find that Σ is BRST exact, up to an additive constant:

$$\Sigma = \mathcal{D} [c^\alpha E_\alpha(\varphi, \lambda) + \frac{1}{2} w_{\alpha\beta,\gamma} c^\alpha c^\beta \bar{c}^\gamma] + \text{const..} \quad (\text{A26.23})$$

More general solutions. One can look for general solutions Σ of equation (A26.18), polynomial in $\bar{c}c$:

$$\Sigma = \sum_{p=0}^n \Sigma_p, \quad (\text{A26.24})$$

where Σ_p has the form

$$\Sigma_p = c^{\alpha_1} \bar{c}^{\beta_1} \cdots c^{\alpha_p} \bar{c}^{\beta_p} [\Sigma_p(\varphi, \lambda)]_{\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_p}.$$

In particular, Σ_p being proportional to p factors $\bar{c}c$, can be considered as a p -form.

The equation $\mathcal{D}\Sigma = 0$ decomposes into

$$\mathcal{D}_+ \Sigma_p = -\mathcal{D}_- \Sigma_{p+1} \quad \text{for } p \leq n-1, \quad \text{and} \quad \mathcal{D}_+ \Sigma_n = 0. \quad (\text{A26.25})$$

We now assume that, in the φ manifold, any closed form is exact (*i.e.* that the φ -manifold is simply connected). These equations can then be solved for decreasing values of p . First,

$$\Sigma_n = \mathcal{D}_+ \tilde{\Sigma}_{n-1},$$

where $\tilde{\Sigma}_{n-1}$ is a $(n-1)$ -form. For $p = n-1$,

$$\mathcal{D}_+ \Sigma_{n-1} = -\mathcal{D}_- \Sigma_n = -\mathcal{D}_- \mathcal{D}_+ \tilde{\Sigma}_{n-1} = \mathcal{D}_+ \mathcal{D}_- \tilde{\Sigma}_{n-1}$$

and, therefore,

$$\Sigma_{n-1} = \mathcal{D}_- \tilde{\Sigma}_{n-1} + \mathcal{D}_+ \tilde{\Sigma}_{n-2}.$$

Then, assuming by induction,

$$\Sigma_{p+1} = \mathcal{D}_- \tilde{\Sigma}_{p+1} + \mathcal{D}_+ \tilde{\Sigma}_p,$$

we find

$$\mathcal{D}_+ \Sigma_p = -\mathcal{D}_- \Sigma_{p+1} = -\mathcal{D}_- \mathcal{D}_+ \tilde{\Sigma}_p = \mathcal{D}_+ \mathcal{D}_- \tilde{\Sigma}_p,$$

whose general solution agrees with the induction hypothesis. The last equation is special and its solution is, We conclude that Σ can be written as

$$\Sigma = \mathcal{D} \tilde{\Sigma} + \text{const.}, \quad (A26.26)$$

with

$$\tilde{\Sigma} = \sum_{p=0}^{n-1} \tilde{\Sigma}_p.$$

This analysis proves that, in the case of simply connected manifolds, any BRST-symmetric function is BRST exact, up to a constant:

$$\mathcal{D}\Sigma = 0 \Rightarrow \Sigma = \mathcal{D} \tilde{\Sigma} + \text{const.}. \quad (A26.27)$$

An important consequence. To illustrate the importance of the property that the weight function Σ is not only BRST closed but also BRST exact, we consider the following integral:

$$Z = \langle 1 \rangle = \int \prod_\alpha (d\varphi^\alpha d\lambda^\alpha d\bar{c}^\alpha dc^\alpha) \exp [-\Sigma(\varphi, \lambda, c, \bar{c})], \quad (A26.28)$$

and calculate the variation of δZ of Z induced by an infinitesimal variation $\delta E_\alpha(\varphi, \nu)$ of the function $E_\alpha(\varphi, \nu)$ in equation (26.40) (the gauge condition in gauge theories).

To this variation corresponds a variation $\delta\Sigma$ of Σ which, according to our preceding analysis, can be written as

$$\delta\Sigma = \mathcal{D} [\delta \tilde{\Sigma}(\varphi, \lambda, c, \bar{c})]. \quad (A26.29)$$

Therefore, δZ has the form

$$\delta Z = \int \prod_\alpha (d\varphi^\alpha d\lambda^\alpha dc^\alpha d\bar{c}^\alpha) \mathcal{D}[\delta \tilde{\Sigma}] \exp(-\Sigma). \quad (A26.30)$$

Since \mathcal{D} is a differential operator, we can integrate by parts:

$$\delta Z = \int \prod_{\alpha} (d\varphi^{\alpha} d\lambda^{\alpha} dc^{\alpha} d\bar{c}^{\alpha}) \delta \tilde{\Sigma} \mathcal{D}\Sigma \exp(-\Sigma). \quad (A26.31)$$

Then, using the BRST symmetry condition $\mathcal{D}\Sigma = 0$, we conclude that

$$\delta Z = 0. \quad (A26.32)$$

Of course, this result is not surprising, since by construction Z is a constant independent of $E_{\alpha}(\varphi)$. However, we infer from it that equation (A26.27) implies this independence by itself without any further assumption on the explicit form of Σ .

The generalization of this result has important implications in quantized gauge theories, as we have argued in Section 26.11.

A26.4 From BRST symmetry to supersymmetry

A trivial extension of BRST transformations can be defined by considering a function of several Grassmann variables $f(\theta_1, \dots, \theta_n)$, and performing independent translations on each variable. The generators of these translations are the differential operators $\partial/\partial\theta_i$. The equivalent of the usual commutation relations between Lie algebra generators are replaced by anticommutation relations:

$$\frac{\partial}{\partial\theta_i} \frac{\partial}{\partial\theta_j} + \frac{\partial}{\partial\theta_j} \frac{\partial}{\partial\theta_i} = 0. \quad (A26.33)$$

Translations on Grassmann variables form the equivalent of usual Abelian Lie groups; non-trivial extensions ('non-Abelian') correspond to non-vanishing anticommutators.

Below, we denote the anticommutators by $\{\bullet, \bullet\}$.

First, a few general remarks are in order:

(i) If Q is the unique generator of this generalized structure called 'supergroup', then exponentiation implies

$$e^{\varepsilon_1 Q} e^{\varepsilon_2 Q} = e^{(\varepsilon_1 + \varepsilon_2)Q},$$

in which ε_1 and ε_2 are Grassmann variables. Expanding both sides, one obtains the condition

$$-\varepsilon_1 \varepsilon_2 Q^2 = 0 \Rightarrow Q^2 = 0, \quad (A26.34)$$

which is indeed satisfied by the generators of translations. If this condition is not satisfied, the supergroup has two generators Q and Q^2 , in which Q^2 is a commuting differential operator.

(ii) More generally, the anticommutators of generators are even elements of the Grassmann algebra. They form an ordinary Lie algebra. If we denote by Q_i , L_a the anticommuting and commuting elements, respectively, the general structure of a Lie superalgebra is

$$\{Q_i, Q_j\} = c_{ija} L_a, \quad [Q_i, L_a] = d_{iaj} Q_j, \quad [L_a, L_b] = f_{abc} L_c.$$

(iii) Triplets of generators thus satisfy a mixed Jacobi identity of the form:

$$[Q_i, \{Q_j, Q_k\}] + \text{cyclic permutations } (ijk) = 0. \quad (A26.35)$$

A number of properties of Lie groups and algebras can be extended to supergroups and super or graded Lie algebras. A discussion of these topics goes much beyond the scope of this work. We describe only a few simple examples.

Supersymmetry and time translation. We want to realize our algebra under the form of differential operators. From the preceding considerations, we conclude that we need at least two Grassmann variables, $\bar{\theta}$ and θ , and one real variable, which we call time and denote by t .

Differential operators, odd elements of the corresponding Grassmann algebra, have the form

$$\frac{\partial}{\partial \bar{\theta}}, \frac{\partial}{\partial \theta}, \bar{\theta} \frac{\partial}{\partial t}, \theta \frac{\partial}{\partial t}.$$

If we impose the condition (A26.34), we find the two non-trivial combinations

$$Q = \frac{\partial}{\partial \theta} + a\bar{\theta} \frac{\partial}{\partial t}, \quad \bar{Q} = \frac{\partial}{\partial \bar{\theta}} + \bar{a}\theta \frac{\partial}{\partial t}, \quad (A26.36)$$

in which a and \bar{a} are two arbitrary complex constants.

It follows that

$$Q\bar{Q} + \bar{Q}Q = (\bar{a} + a) \frac{\partial}{\partial t}. \quad (A26.37)$$

For $a + \bar{a} \neq 0$, the Lie subgroup is the group of translations on the variable t . This is a realization of quantum mechanics (QM) supersymmetry, as discussed in Section 35.4.

This structure can immediately be generalized to d -dimensional space. We define

$$Q_\alpha = \frac{\partial}{\partial \theta_\alpha} + a_{\alpha\beta}^\mu \bar{\theta}_\beta \frac{\partial}{\partial x_\mu}, \quad \bar{Q}_\alpha = \frac{\partial}{\partial \bar{\theta}_\alpha} + \theta_\beta \bar{a}_{\beta\alpha}^\mu \frac{\partial}{\partial x_\mu}, \quad (A26.38)$$

in which x_μ corresponds to the d commuting variables, $\bar{\theta}_\alpha$ and θ_α to anticommuting variables, and $a_{\alpha\beta}^\mu$, $\bar{a}_{\alpha\beta}^\mu$ are constants. Then,

$$\begin{cases} Q_\alpha Q_\beta + Q_\beta Q_\alpha = 0, \\ \bar{Q}_\alpha \bar{Q}_\beta + \bar{Q}_\beta \bar{Q}_\alpha = 0, \\ Q_\alpha \bar{Q}_\beta + \bar{Q}_\beta Q_\alpha = (\bar{a}_{\alpha\beta}^\mu + a_{\alpha\beta}^\mu) \frac{\partial}{\partial x_\mu}. \end{cases} \quad (A26.39)$$

Again the Lie subgroup is a product of translation groups. Note that to \bar{Q}, Q correspond two other supersymmetry generators:

$$D_\alpha = \frac{\partial}{\partial \theta_\alpha} - \bar{a}_{\alpha\beta}^\mu \bar{\theta}_\beta \frac{\partial}{\partial x_\mu}, \quad \bar{D}_\alpha = \frac{\partial}{\partial \bar{\theta}_\alpha} - \bar{\theta}_\beta \bar{a}_{\beta\alpha}^\mu \frac{\partial}{\partial x_\mu}, \quad (A26.40)$$

which anticommute with \bar{Q}, Q , and thus can play the role of covariant derivatives.

27 Supersymmetric quantum field theory (QFT): Introduction

For completeness, in this chapter, we describe a few properties of supersymmetric QFT [295–297]. Supersymmetry has been proposed, in particular, as a principle to solve the so-called *fine-tuning problem* in particle physics by relating the masses of scalar particles (like Higgs fields) to those of fermions, which can be protected against ‘large’ mass renormalization by chiral symmetry (see Section 23.3.1). However, supersymmetry is, at best, an approximate symmetry broken at a scale beyond the reach of the Large Hadron Collider (LHC), because the possible supersymmetric partners of known particles have not been discovered yet (2020) and thus must be much heavier. Therefore, a curing of the fine-tuning problem seems to require, at least, additional ingredients.

Exact supersymmetry would also have implied the vanishing of the vacuum energy and thus, of the cosmological constant (see Section 28.6.2). The discovery of *dark energy* has a natural interpretation as resulting from a very small cosmological constant. However, a naive model based on broken supersymmetry would still predict 60 orders of magnitude too large a value, compared to 120 orders of magnitude otherwise.

Gauging supersymmetry leads naturally to a unification with gravity, because the commutators of supersymmetry currents involve the energy-momentum tensor.

In Section 27.1, we describe examples of supersymmetric theories involving scalar superfields. Such models constitute the simplest generalizations of supersymmetric quantum mechanics (QM), as it naturally appears, for example, in the study of stochastic differential equations of Langevin type (see Sections A26.4 and 34.7.2). The new feature of supersymmetry in higher dimensions is the combination of supersymmetry with spin, since fermions have spins, and this leads to the modifications we describe in this chapter.

First, we consider a supersymmetric model, with one scalar superfield, in three Euclidean dimensions, where fermion spin is associated with Pauli matrices, and then generalize to $O(N)$ -symmetric models.

We then examine the example of the supersymmetric non-linear σ -model in dimension 2, very much as we have done in the non-supersymmetric examples.

In Section 27.3, we briefly review supersymmetry in four dimensions, a structure potentially relevant for four-dimensional particle physics. We introduce scalar chiral superfields. In Section 27.4, we introduce super vector fields, and discuss gauge invariance.

To discuss supersymmetric models, we use the superfield notation introduced in supersymmetric QM, and the formalism of Grassmann coordinates [298] (see Sections 26.5.2 and 34.7.2).

Notation. In this chapter, implicit summation over repeated indices is always assumed.

27.1 Scalar superfields in three dimensions

We first discuss a class of QFTs with scalar superfields, in three dimensions. We exhibit a few properties of the perturbative expansion and perform one-loop calculations. Note that the $O(N)$ -symmetric theory can also be solved in the large N limit [181].

27.1.1 Supersymmetry and Majorana spinors in $d = 3$

In three Euclidean dimensions, the spin group is $SU(2)$ (see Section A12). A spinor transforms like

$$\psi_U = U\psi, \quad U \in SU(2),$$

or, in component form,

$$[\psi_U]_\alpha = U_\alpha^\beta \psi_\beta. \quad (27.1)$$

The role of Dirac γ matrices is played by the Pauli σ matrices, $\gamma_\mu \equiv \sigma_\mu$. Moreover, σ_2 is antisymmetric while $\sigma_2 \sigma_\mu$ is symmetric. This implies (\bullet^T denotes transposition)

$$\sigma_2 \sigma_\mu \sigma_2 = -\sigma_\mu^T \Rightarrow U^* = \sigma_2 U \sigma_2.$$

We will need the identity tensor δ_β^α , which is the Kronecker δ , and the antisymmetric tensor $\epsilon_{\alpha\beta} = \epsilon^{\alpha\beta} = -\epsilon_{\beta\alpha}$, $\epsilon_{12} = 1$.

A Majorana spinor corresponds to a neutral fermion and has only two independent components ψ_1, ψ_2 . It is ‘real’ in the sense of the conjugation defined in Grassmann algebras in Section 12.3.3.

It is convenient to define, for each spinor ψ_α , the conjugated spinor $(\psi^T \sigma_2)$,

$$\psi^\alpha = i\epsilon^{\alpha\beta} \psi_\beta \Rightarrow \psi_\alpha = i\epsilon_{\alpha\beta} \psi^\beta. \quad (27.2)$$

Then, the transformation (27.1) becomes

$$[\psi_U]^\alpha = \psi^\beta (U^\dagger)_\beta^\alpha.$$

Finally, ψ and χ being two spinors, we use the notation

$$\psi \cdot \chi \equiv \psi^\alpha \chi_\alpha = \chi \cdot \psi, \quad (27.3)$$

where $\psi \cdot \chi$ is an $SU(2)$ invariant.

Corresponding to fermion spinors, we need a spinor θ^α of real Grassmann coordinates. Similarly, we define

$$\theta_\alpha = i\epsilon_{\alpha\beta} \theta^\beta \Leftrightarrow \theta^\alpha = i\epsilon^{\alpha\beta} \theta_\beta.$$

Since the only non-vanishing product is $\theta^1 \theta^2$, we find

$$\theta^\alpha \theta_\beta = \frac{1}{2} \delta_\beta^\alpha \theta^\gamma \theta_\gamma \equiv \frac{1}{2} \delta_\beta^\alpha (\theta \cdot \theta) = i\delta_\beta^\alpha \theta^1 \theta^2.$$

Moreover, we define

$$\theta^\alpha \sigma_{\mu\alpha}^\beta \psi_\beta \equiv (\theta \sigma_\mu \psi) \Rightarrow (\theta \sigma_\mu \psi) = -(\psi \sigma_\mu \theta). \quad (27.4)$$

Other useful identities then are

$$(\theta \cdot \psi)^2 = -\frac{1}{2}(\theta \cdot \theta)(\psi \cdot \psi), \quad (\theta \not{\psi})^2 = \frac{1}{2}p^2(\psi \cdot \psi)(\theta \cdot \theta).$$

We integrate over θ_1, θ_2 with the measure

$$d^2\theta \equiv -\frac{1}{2}id\theta^2 d\theta^1,$$

which is invariant under complex conjugation. Then,

$$\int d^2\theta \theta^\alpha \theta_\beta = \frac{1}{2} \delta_\beta^\alpha, \quad \int d^2\theta (\theta \cdot \theta) = 1.$$

With this convention, the δ -function $\delta^2(\theta' - \theta)$ in θ space is

$$\delta^2(\theta' - \theta) = (\theta' - \theta) \cdot (\theta' - \theta). \quad (27.5)$$

27.1.2 Superfields and covariant derivatives

A scalar superfield $\Phi(x, \theta)$ has the expansion

$$\Phi(x, \theta) = \varphi(x) + (\theta \cdot \psi(x)) + \frac{1}{2}(\theta \cdot \theta)F(x). \quad (27.6)$$

Again, although only two θ variables are independent, it is convenient to define the four covariant derivatives

$$D_\alpha \equiv \frac{\partial}{\partial \theta^\alpha} - \not{\partial}_\alpha^\beta \theta_\beta, \quad D^\alpha \equiv \frac{\partial}{\partial \theta_\alpha} - \theta^\beta \not{\partial}_\beta^\alpha. \quad (27.7)$$

Then, the anticommutation relation is

$$\{D_\alpha, D^\beta\} = -2[\not{\partial}]_\alpha^\beta.$$

Also $((\theta \not{\partial})^\alpha \equiv \theta^\beta \not{\partial}_\beta^\alpha)$,

$$D^\alpha D_\alpha = \frac{\partial}{\partial \theta_\alpha} \frac{\partial}{\partial \theta^\alpha} - (\theta \not{\partial})^\alpha \frac{\partial}{\partial \theta^\alpha} - \frac{\partial}{\partial \theta_\alpha} (\not{\partial} \theta)_\alpha + (\theta \cdot \theta) \nabla^2.$$

Since the σ_μ are traceless, using the identity (27.4), one verifies that $(D \cdot D)$ can also be written as

$$D^\alpha D_\alpha = \frac{\partial}{\partial \theta_\alpha} \frac{\partial}{\partial \theta^\alpha} - 2(\theta \not{\partial})^\alpha \frac{\partial}{\partial \theta^\alpha} + (\theta \cdot \theta) \nabla^2$$

and, therefore, in component form

$$D^\alpha D_\alpha \Phi(x, \theta) = 2F(x) - 2(\theta \not{\partial} \psi(x)) + (\theta \cdot \theta) \nabla^2 \varphi(x).$$

27.1.3 Supersymmetry generators and Ward–Takahashi (WT) identities

Supersymmetry is generated by the operators Q^α (or Q_α), which anticommute with D_α (and thus D^α):

$$Q_\alpha \equiv \frac{\partial}{\partial \theta^\alpha} + (\not{\partial} \theta)_\alpha, \quad Q^\alpha \equiv \frac{\partial}{\partial \theta_\alpha} + (\theta \not{\partial})^\alpha. \quad (27.8)$$

Then,

$$\{Q_\alpha, Q^\beta\} = 2[\not{\partial}]_\alpha^\beta.$$

The anticommutator of Q_α and Q^β is the generator of space translations: *supersymmetry implies space translation invariance*.

Supersymmetry also implies WT identities for correlation functions. The n -point function $\widetilde{W}^{(n)}(p, \theta)$, in Fourier representation, satisfies (here, θ_α^k means the k th argument θ_α)

$$Q^\alpha \widetilde{W}^{(n)}(p_1, \theta^1, \dots, p_n, \theta^n) \equiv \left[\sum_{k=1}^n \frac{\partial}{\partial \theta_\alpha^k} + i(\theta^k \not{p}_k)^\alpha \right] \widetilde{W}^{(n)}(p_1, \theta^1, \dots, p_n, \theta^n) = 0.$$

To solve this equation, we set

$$\widetilde{W}^{(n)}(p_1, \theta^1, \dots, p_n, \theta^n) = F^{(n)}(p_1, \theta^1, \dots, p_n, \theta^n) \exp \left[\frac{i}{2n} \sum_{j,k} (\theta^j (\not{p}_j - \not{p}_k) \theta^k) \right], \quad (27.9)$$

where $F^{(n)}$ is a symmetric function in the exchange $\{p_i, \theta_i\} \leftrightarrow \{p_j, \theta_j\}$. It then satisfies

$$\sum_k \frac{\partial}{\partial \theta_\alpha^k} F^{(n)}(p_1, \theta^1, \dots, p_n, \theta^n) = 0,$$

that is, it is translation-invariant in θ space.

In the case of the two-point function, this leads to the general form

$$\begin{aligned} \widetilde{W}^{(2)}(p, \theta', \theta) &= A(p^2) [1 + C(p^2) \delta^2(\theta' - \theta)] e^{i(\theta \not{p} \theta')} \\ &= A(p^2) [1 + C(p^2)((\theta' - \theta) \cdot (\theta' - \theta)) + i(\theta \not{p} \theta') - \frac{1}{4} p^2 (\theta \cdot \theta) (\theta' \cdot \theta')]. \end{aligned} \quad (27.10)$$

The same identities apply to the vertex (1PI) functions $\tilde{\Gamma}^{(n)}$.

27.1.4 General $O(N)$ -symmetric action

We now consider a general $O(N)$ -invariant supersymmetric action,

$$\mathcal{S}(\Phi) = \int d^3x d^2\theta \left[\frac{1}{2} D^\alpha \Phi(x, \theta) \cdot D_\alpha \Phi(x, \theta) + \mathcal{U}(\Phi^2(x, \theta)) \right], \quad (27.11)$$

where \mathcal{U} is a polynomial, and the superscalar field $\Phi \equiv \{\Phi_i\}$ an N -component vector, which can be parametrized as,

$$\Phi(x, \theta) = \varphi(x) + (\theta \cdot \psi(x)) + \frac{1}{2}(\theta \cdot \theta)\mathbf{F}(x). \quad (27.12)$$

In this parametrization,

$$\int d^2\theta D^\alpha \Phi \cdot D_\alpha \Phi = - \int d^2\theta \Phi D^\alpha D_\alpha \Phi = -(\psi \bar{\partial} \psi) + (\nabla_x \varphi)^2 - \mathbf{F}^2. \quad (27.13)$$

Then, since

$$\Phi^2(x, \theta) = \varphi^2(x) + 2\varphi(x)(\theta \cdot \psi(x)) - \frac{1}{2}(\theta \cdot \theta)(\psi(x) \cdot \psi(x)) + \mathbf{F}(x)\varphi(x)(\theta \cdot \theta), \quad (27.14)$$

the potential term expansion reads,

$$\int d^2\theta \mathcal{U}(\Phi^2) = \mathcal{U}'(\varphi^2) \left(-\frac{1}{2}\psi \cdot \psi + \mathbf{F}\varphi \right) - \mathcal{U}''(\varphi^2)(\varphi\psi \cdot \varphi\psi).$$

In the case of a free theory, with $\mathcal{U}(\rho) \equiv \mu\rho$ (μ constant), the action in component form reduces to

$$\begin{aligned} \mathcal{S}(\varphi, \psi, \mathbf{F}) &= \int d^3x \left\{ -\frac{1}{2}\psi(x)\bar{\partial}\psi(x) + \frac{1}{2}(\nabla_x \varphi(x))^2 - \frac{1}{2}\mathbf{F}^2(x) \right. \\ &\quad \left. + \mu \left[-\frac{1}{2}\psi(x) \cdot \psi(x) + \mathbf{F}(x)\varphi(x) \right] \right\}. \end{aligned}$$

After integration over the auxiliary field \mathbf{F} , the action becomes

$$\mathcal{S}(\varphi, \psi) = \int d^3x \left[-\frac{1}{2}\psi(x)(\bar{\partial} + \mu)\psi(x) + \frac{1}{2}(\nabla_x \varphi(x))^2 + \frac{1}{2}\mu^2\varphi^2(x) \right]. \quad (27.15)$$

For a generic super-potential $\mathcal{U}(\rho)$, one finds

$$\begin{aligned} \mathcal{S}(\varphi, \psi, \mathbf{F}) &= \int d^3x \left\{ -\frac{1}{2}\psi(x)\bar{\partial}\psi(x) + \frac{1}{2}(\nabla_x \varphi(x))^2 - \frac{1}{2}\mathbf{F}^2(x) \right. \\ &\quad + \mathcal{U}'(\varphi^2(x)) \left(-\frac{1}{2}\psi(x) \cdot \psi(x) + \mathbf{F}(x)\varphi(x) \right) \\ &\quad \left. - \mathcal{U}''(\varphi^2(x)) [(\psi(x)\varphi(x)) \cdot (\varphi(x)\psi(x))] \right\}, \end{aligned} \quad (27.16)$$

and thus again, after integration over the auxiliary field \mathbf{F} ,

$$\begin{aligned} \mathcal{S}(\varphi, \psi) &= \int d^3x \left[-\frac{1}{2}\psi(x)\bar{\partial}\psi(x) + \frac{1}{2}(\nabla_x \varphi(x))^2 - \frac{1}{2}\mathcal{U}'(\varphi^2(x))\psi(x) \cdot \psi(x) \right. \\ &\quad \left. - \mathcal{U}''(\varphi^2(x))(\psi(x)\varphi(x)) \cdot (\varphi(x)\psi(x)) + \frac{1}{2}\varphi^2(x)\mathcal{U}'^2(\varphi^2(x)) \right]. \end{aligned} \quad (27.17)$$

The theory violates fermion number conservation (Majorana spinors) and parity symmetry. Actually, a space reflection (see definition in Section A12.1.6) is equivalent to the change $\mathcal{U} \mapsto -\mathcal{U}$. Therefore, theories with $\pm\mathcal{U}$ have the same physics properties.

Moreover, under the form (27.16), one notes that the structure is similar to the dynamic action associated to a Langevin equation. It can be shown that the determinant, resulting from the integration over the fermion fields, is the Jacobian of a non-local, non-linear transformation for the scalar field which transforms the action into a free action. This transformation implies that the partition function is 1 at zero temperature.

27.1.5 Perturbative expansion

Power counting. In three dimensions, the superfield has canonical dimension 1/2, and θ dimension $-1/2$

$$[\theta] = -\frac{1}{2}, \quad [\Phi] = \frac{1}{2}.$$

Here, we do not consider the Φ^3 theory for a one-component field, which is super-renormalizable (and has a renormalizable generalization in four dimensions). The $(\Phi^2)^2$ field theory is the only theory renormalizable in three dimensions. Therefore, we now discuss the potential

$$\mathcal{U}(\rho) = \mu\rho + \frac{1}{2}g\rho^2, \quad (27.18)$$

where g is the coupling constant. Prior to a more refined analysis, one expects coupling constant and field renormalizations (with logarithmic divergences), and a mass renormalization with linear divergences. In terms of the parametrization (27.10) for the two-point vertex function $\tilde{\Gamma}^{(2)}$, one infers that the coefficient $A(p^2)$ has at most a logarithmic divergence, which corresponds to the field renormalization, while the coefficient $C(p^2)$ can have a linear divergence, which corresponds to a mass renormalization.

The super-propagator. To generate perturbation theory, we need the field propagator. In Fourier space and in terms of the field components,

$$\begin{aligned} & (-D^\alpha D_\alpha + 2\mu) \delta^2(\theta' - \theta) \\ &= 4 \left[-1 - i(\theta k \theta') + \frac{1}{2}\mu(\theta' - \theta)(\theta' - \theta) + \frac{1}{4}k^2(\theta\theta)(\theta'\theta') \right]. \end{aligned} \quad (27.19)$$

The super-propagator Δ of the Φ field, in Fourier representation, solution of

$$(-D^\alpha D_\alpha + 2\mu) \tilde{\Delta}(k, \theta, \theta') = \delta^2(\theta' - \theta),$$

can be obtained by solving the equation

$$(-D^\alpha D_\alpha + 2\mu)\Phi = J,$$

or, in component form ($\tilde{\varphi}, \tilde{\psi}, \tilde{F}$ are the Fourier transforms of φ, ψ, F)

$$2\mu\tilde{\varphi}(k) - 2\tilde{F}(k) + 2(\theta(i\mathbf{k} + \mu)\tilde{\psi}(k)) + (\theta\theta)(k^2\tilde{\varphi}(k) + 2\mu\tilde{F}(k)) = J(\theta, k).$$

It is given by

$$\begin{aligned} \tilde{\Delta}(k, \theta, \theta') &= \frac{1 + \frac{1}{2}\mu\delta^2(\theta - \theta')}{k^2 + \mu^2} e^{i(\theta\mathbf{k}\theta')} \\ &= \frac{1}{k^2 + \mu^2} \left[1 + \frac{1}{2}\mu((\theta\theta) + (\theta'\theta')) + \theta(i\mathbf{k} - \mu)\theta' - \frac{1}{4}k^2(\theta\theta)(\theta'\theta') \right]. \end{aligned} \quad (27.20)$$

Expression (27.20) exhibits the $\langle \tilde{\varphi}(k)\tilde{\varphi}(-k) \rangle$ propagator $(k^2 + \mu^2)^{-1}$ and the $\langle \tilde{\psi}(k)\tilde{\psi}(-k) \rangle$ propagator $(i\mathbf{k} - \mu)/(k^2 + \mu^2)$. The coefficients of $(\theta\theta) + (\theta'\theta')$ and of $(\theta\theta)(\theta'\theta')$ are the $\langle \tilde{\varphi}(k)\tilde{F}(-k) \rangle$ and $\langle \tilde{F}(k)\tilde{F}(-k) \rangle$ propagators, respectively.

One-loop divergences. The one-loop diagram contributing to the two-point function has the simple form

$$2\delta^2(\theta - \theta')(N + 2)g \text{tr } \tilde{\Delta}(k, \theta, \theta') = 2\delta^2(\theta - \theta') \frac{N + 2}{(2\pi)^3} g \int \frac{d^3k}{k^2 + \mu^2}.$$

It exhibits the expected linear mass divergence, and requires a mass renormalization $\delta\mu$, for example,

$$\delta\mu = -(N+2)g \frac{1}{(2\pi)^3} \int \frac{d^3k}{k^2}.$$

(We do not discuss here the problem of supersymmetric regularization.)

The contribution to the four-point function is proportional to the usual bubble diagram,

$$g^2 \int \frac{d^3k}{(2\pi)^3} \tilde{\Delta}(k, \theta, \theta') \tilde{\Delta}(p - k, \theta, \theta').$$

Then,

$$\tilde{\Delta}(k, \theta, \theta') \tilde{\Delta}(p - k, \theta, \theta') = \frac{[1 + \mu\delta^2(\theta' - \theta)] e^{i\theta p \theta'}}{(k^2 + \mu^2)[(p+k)^2 + \mu^2]}.$$

Note the cancellation of the factor $e^{i\theta p \theta'}$, which renders the integral more convergent than one would naively expect. The integral over k then yields the finite three-dimensional scalar bubble diagram (see figure 18.2)

$$B(p) = \frac{1}{(2\pi)^3} \int \frac{d^3k}{(k^2 + \mu^2)[(p+k)^2 + \mu^2]} = \frac{1}{4\pi p} \text{Arctan}(p/2|\mu|),$$

and no coupling constant renormalization is required, as in the non-supersymmetric, super-renormalizable $(\phi^2)^2$ field theory.

27.2 The $O(N)$ supersymmetric non-linear σ model

The action for the supersymmetric non-linear σ -model, in dimensions $2 \leq d \leq 3$ (a model that can also be solved in the large N limit [181]), reads [299]

$$\mathcal{S}(\Phi) = \frac{1}{2\kappa} \int d^d x d^2\theta D\Phi(x, \theta) \cdot D\Phi(x, \theta), \quad (27.21)$$

where κ is the coupling constant and Φ a scalar superfield and an N -component vector, which satisfies

$$\Phi(x, \theta) \cdot \Phi(x, \theta) = 1. \quad (27.22)$$

In terms of the parametrization (27.12), the constraint takes the form (equation (27.14)),

$$\varphi^2(x) = 1, \quad \varphi(x) \cdot \psi(x) = 0, \quad F(x) \cdot \varphi(x) = \frac{1}{2}\psi(x) \cdot \psi(x).$$

Using the expansion (27.13), and integrating over the auxiliary field F , one obtains the action in the explicit form

$$\mathcal{S} = \frac{1}{2\kappa} \int d^d x \left[(\nabla \varphi(x))^2 - \bar{\psi}(x) \not{\partial} \psi(x) - \frac{1}{4}(\psi(x) \psi(x))^2 \right].$$

The constraint (27.22) can also be implemented by introducing the scalar superfield,

$$\mathbf{L}(x, \theta) = L(x) + \theta \cdot \ell(x) + \frac{1}{2}(\theta \cdot \theta)\lambda(x), \quad (27.23)$$

adding to the action

$$\mathcal{S}_L = \frac{1}{\kappa} \int d^d x d^2\theta \mathbf{L}(x, \theta) [\Phi^2(x, \theta) - 1], \quad (27.24)$$

and integrating over \mathbf{L} , or equivalently, over $L(x)$, $\lambda(x)$, and $\ell(x)$.

The partition function is then given by ($\mathcal{S}(\Phi, \mathbf{L}) = \mathcal{S} + \mathcal{S}_L$)

$$\mathcal{Z} = \int [d\Phi][dL] e^{-\mathcal{S}(\Phi, \mathbf{L})}.$$

In terms of the field parametrizations (27.12) and (27.23), after integration over the auxiliary field F , the total action reads

$$\begin{aligned} \mathcal{S} = & \frac{1}{2\kappa} \int d^d x [\varphi(x)(-\nabla^2 + L^2(x))\varphi(x) - \bar{\psi}(x)(\not{D} + L(x))\psi(x) \\ & + \lambda(x)(\varphi^2(x) - 1) - 2\ell(x)(\psi(x) \cdot \varphi(x))] . \end{aligned} \quad (27.25)$$

In the superfield formalism, at one-loop, the calculation of the renormalization group (RG) β -function is analogous to the calculation in the non-supersymmetric model (equations (19.96, 19.97), and (19.99)); one simply changes the propagator. Actually, the β and η functions have been calculated up to four loops [300]. In the minimal subtraction (MS) scheme,

$$\beta(\tilde{\kappa}) = (d-2)\tilde{\kappa} - (N-2)\tilde{\kappa}^2 - \frac{3}{2}\zeta(3)(N-2)(N-3)\tilde{\kappa}^5 + O(\tilde{\kappa}^6), \quad (27.26)$$

$$\eta(\tilde{\kappa}) = 2-d+\tilde{\kappa}(N-1)-\frac{1}{4}(N-1)(N-2)(N-3)\tilde{\kappa}^4 + O(\tilde{\kappa}^5), \quad (27.27)$$

with $\tilde{\kappa} = N_d \kappa$ (N_d is the loop factor (10.12)), and $\zeta(s)$ is Riemann's function. For a class of generalized non-linear σ -models, the RG β -functions are related to instanton calculus [301].

The form (27.25) of the action is convenient for solving the model in the large N limit [181]. For example, the critical exponents η and ν , which are known up to order $1/N^2$ and $1/N^3$ [302], at order $1/N$, are

$$\eta = \frac{4\Gamma(d-2)}{\Gamma(d/2)\Gamma^2(d/2-1)\Gamma(2-d/2)} \frac{1}{N} + O(N^{-2}), \quad \frac{1}{\nu} = (d-2) + O(N^{-2}).$$

Dimension 2. Power counting shows that the model is renormalizable in two dimensions, and exhibits at leading order, massless boson and fermion fields, consequence of supersymmetry and the spontaneous breaking of the $O(N)$ symmetry. To avoid infrared (IR) divergences, it is also necessary here to add to the action a mass term that breaks the $O(N)$ symmetry explicitly, for example, by adding a term linear in Φ .

The supersymmetric model, like the non-supersymmetric model (Section 19.14), is asymptotically free (equation (27.26)). The supersymmetry is unbroken, and the particles are massive. Much is known about the model in two dimensions, including the exact S -matrix. Some of these properties are shared by generalized non-linear σ -models [303].

27.3 Supersymmetry in four dimensions

Even though no particle predicted by supersymmetric extensions of the Standard Model of particle physics has been discovered yet, in this section we present, for completeness, a very brief introduction to supersymmetry in four dimensions [304]. This requires the addition of gauge superfields, something we have not done yet.

The algebra of generators of supersymmetry is briefly described in Section A26.4. We again use the superfield formalism of Section 27.1 and, for instance, combine scalar boson and spin 1/2 fermion fields into scalar superfields.

In four dimensions, the Euclidean spinor group $\text{Spin}(4)$ is identical to $SU(2) \times SU(2)$. The spinorial representation can be reduced into left-handed and right-handed components (Weyl spinors). Moreover, the representations of $SU(2)$ are self-conjugated. It is thus sufficient to associate Grassmann coordinates with the two Weyl spinors. We need four Grassmann coordinates $\theta_\alpha, \bar{\theta}_\alpha$.

Renormalizable field theories are constrained by power counting. In four dimensions, these restrictions somewhat destroy the beautiful simplicity of the lower-dimensional examples discussed so far (an incomplete effective theory?).

In any space dimension, the dimension $[\theta]$ of the Grassmann coordinates θ is $[\theta] = [\text{scalar}] - [\text{spinor}] = -1/2$.

In this form, the Grassmann volume element has only dimension 2, and an action bilinear in the scalar superfield can be found: such a contribution to an action is called a D-term. However, this leads only to a free-field action. As we will explain, to construct interaction terms, one has to consider cubic polynomials of only one kind of chiral superfield (left- or right-handed). Such a polynomial only depends on two variables θ or $\bar{\theta}$, and thus can be integrated with the volume elements $d\theta_1 d\theta_2$ ($d\bar{\theta}_1 d\bar{\theta}_2$ respectively), which have dimension 1. The corresponding contribution is called an F-term.

27.3.1 Grassmann coordinates and supersymmetry

To define the generators of supersymmetry (see Section 27.1.1), it is convenient to introduce a notation for the two-component spinors transforming under the representation and conjugated representation of $SU(2)$:

$$\theta^\alpha = -(\sigma_2)^{\alpha\beta} \theta_\beta \equiv i\epsilon^{\alpha\beta} \theta_\beta, \quad \bar{\theta}_\alpha = (\sigma_2)_{\alpha\beta} \theta_\beta \equiv -i\epsilon_{\alpha\beta} \bar{\theta}^\beta. \quad (27.28)$$

We also define

$$\int \prod_\alpha (d\bar{\theta}_\alpha d\theta_\alpha) \bar{\theta}^\alpha \bar{\theta}_\beta \theta^\gamma \theta_\delta = -\delta_\beta^\alpha \delta_\delta^\gamma. \quad (27.29)$$

We now introduce the four 2×2 matrices σ_μ , where for $\mu = 1, 2, 3$, σ_μ are simply the standard Pauli matrices and $\sigma_4 = -i\mathbf{1}$ and define $x^\mu = x_\mu \sigma_\mu$.

Two relations are useful in what follows:

$$\sigma_2 \sigma_\mu \sigma_2 = -\sigma_\mu^*, \quad \sigma_\mu^\dagger \sigma_\nu + \sigma_\nu^\dagger \sigma_\mu = 2\delta_{\mu\nu} \Rightarrow x^\mu x^\dagger = x^2.$$

Under a $SU(2) \times SU(2)$ transformation corresponding to the $SU(2)$ matrices $U = i\psi$ with $u^2 = 1$, $V = i\psi'$ with $v^2 = 1$, the coordinates transform like

$$\theta'_\alpha = U_\alpha^\beta \theta_\beta, \quad \bar{\theta}^{\alpha'} = V_\beta^{*\alpha} \bar{\theta}^\beta, \quad x' = V x U.$$

$SU(2) \times SU(2)$ -invariant quantities are then, for example, $(\theta \cdot \theta)$, $(\bar{\theta} \cdot \bar{\theta})$ or $(\bar{\theta} \not{x} \theta)$. The supersymmetry generators

$$\bar{Q}_\alpha = \frac{\partial}{\partial \bar{\theta}^\alpha} + \frac{1}{2} \sigma_{\mu\alpha}^\beta \theta_\beta \frac{\partial}{\partial x_\mu}, \quad (27.30a)$$

$$Q^\alpha = \frac{\partial}{\partial \theta_\alpha} + \frac{1}{2} \bar{\theta}^\beta \sigma_{\mu\beta}^\alpha \frac{\partial}{\partial x_\mu}, \quad (27.30b)$$

belong to the representations $(0, 1/2)$ and $(1/2, 0)$ of $SU(2) \times SU(2)$, respectively. Their anticommutation relations are

$$\{\bar{Q}_\alpha, Q^\beta\} = \not{\partial}_\alpha^\beta, \quad \{Q, Q\} = \{\bar{Q}, \bar{Q}\} = 0. \quad (27.31)$$

To Q, \bar{Q} correspond two operators D, \bar{D} , which also satisfy the anticommutation relations of supersymmetry generators, anticommute with them, and thus play the role of covariant derivatives:

$$\bar{D}_\alpha = \frac{\partial}{\partial \theta^\alpha} - \frac{1}{2} \sigma_{\mu\alpha}^\beta \theta_\beta \frac{\partial}{\partial x_\mu}, \quad D^\alpha = \frac{\partial}{\partial \theta_\alpha} - \frac{1}{2} \bar{\theta}^\beta \sigma_{\mu\beta}^\alpha \frac{\partial}{\partial x_\mu}, \quad (27.32)$$

$$\{\bar{D}_\alpha, D^\beta\} = -\not{\partial}_\alpha^\beta, \quad \{D, D\} = \{\bar{D}, \bar{D}\} = 0. \quad (27.33)$$

Remark. The supersymmetry current has spin 3/2. In particular, this implies that the spontaneous breaking of supersymmetry generates spin 1/2 Goldstone fermions. This is an unwanted feature from the point of view of particle physics phenomenology. Even the neutrinos, with their small mass, do not have the properties of approximate Goldstone fermions.

Since the commutator of supersymmetries is the generator of translations, an attempt to gauge supersymmetry (which could provide a solution to the unwanted Goldstone fermion) necessarily leads to theories which contain gravitation. Therefore, the problem of spontaneous supersymmetry breaking does not seem to have a solution outside supergravity, and thus within the framework of renormalizable field theories.

27.3.2 Scalar chiral superfields

Because \bar{D} anticommutes with \bar{Q}, Q , the fields that satisfy

$$\bar{D}_\alpha \phi(x, \theta, \bar{\theta}) = 0, \quad (27.34)$$

form a space of representation for the supersymmetry generators. The general solution of equation (27.34) can be written as

$$\phi(x, \theta, \bar{\theta}) = \phi(y, \theta),$$

where the translated space coordinate (which has the correct transformation properties under $SU(2) \times SU(2)$)

$$y_\mu = x_\mu + \frac{1}{2} \bar{\theta}^\alpha \sigma_{\mu\alpha}^\beta \theta_\beta, \quad (27.35)$$

also satisfies the equation

$$Q^\alpha y_\mu(x, \theta, \bar{\theta}) = 0. \quad (27.36)$$

Note that the variable y plays a role similar to time in the one-dimensional example (Section 35.4). A scalar right-handed superfield can be expanded on the θ_α basis. Since α takes only two values, the most general expression has the form

$$\phi(y, \theta) = \varphi(y) + \psi^\alpha(y) \theta_\alpha + \frac{1}{2} E(y) \theta^\alpha \theta_\alpha. \quad (27.37)$$

where φ and E are two complex scalar fields; φ as well as ϕ itself have dimension 1 and thus E has dimension 2. A renormalizable action can be at most quadratic in E and, as in the one-dimensional example (Section 35.4), E does not propagate and can be eliminated from the action by using the corresponding equation of motion.

Acting on functions of y, θ , the generators Q, \bar{Q} take the form

$$\bar{Q}_\alpha = \frac{\partial}{\partial \theta^\alpha} + \sigma_{\mu\alpha}^\beta \theta_\beta \frac{\partial}{\partial y_\mu}, \quad Q^\alpha = \frac{\partial}{\partial \theta_\alpha}.$$

The action of the supersymmetry generator $\bar{\eta}^\alpha \bar{Q}_\alpha - \eta_\alpha Q^\alpha$ in component form is then

$$\delta\varphi = \eta_\alpha \psi^\alpha, \quad \delta\psi^\beta = \bar{\eta}^\alpha \partial_\alpha^\beta \varphi - \eta^\beta E, \quad \delta E = -\bar{\eta}^\alpha \partial_\alpha^\beta \psi_\beta. \quad (27.38)$$

For later purpose, it is useful to also expand ϕ at fixed space coordinate x :

$$\begin{aligned} \phi(x, \theta, \bar{\theta}) &= \varphi(x) + \psi^\alpha(x)\theta_\alpha + \frac{1}{2}E(x)\theta^\alpha\theta_\alpha + \frac{1}{2}\partial_\alpha^\beta\varphi(x)\bar{\theta}^\alpha\theta_\beta \\ &\quad + \frac{1}{2}\theta_\alpha\bar{\theta}^\beta\partial_\beta^\gamma\psi^\alpha(x)\theta_\gamma + \frac{1}{16}\nabla_x^2\varphi(x)\theta^\alpha\theta_\alpha\bar{\theta}^\beta\bar{\theta}_\beta. \end{aligned} \quad (27.39)$$

Similarly, left-handed chiral superfields can be defined, which satisfy $D^\alpha \bar{\phi} = 0$, and thus depend on $\bar{\theta}$ and a space variable \bar{y} ,

$$\bar{y}_\mu = x_\mu - \frac{1}{2}\bar{\theta}^\alpha\sigma_{\mu\alpha}^\beta\theta_\beta, \quad \bar{Q}_\alpha\bar{y}_\mu = 0. \quad (27.40)$$

The corresponding chiral field can be written as $(\bar{\theta}_\alpha\bar{\psi}^\alpha = \bar{\psi}_\alpha\bar{\theta}^\alpha)$

$$\bar{\phi}(\bar{y}, \bar{\theta}) = \varphi^*(\bar{y}) + \bar{\theta}_\alpha\bar{\psi}^\alpha(\bar{y}) - \frac{1}{2}E^*(\bar{y})\bar{\theta}^\alpha\bar{\theta}_\alpha. \quad (27.41)$$

Acting on functions of $\bar{y}, \bar{\theta}$, the generators Q, \bar{Q} take the form

$$\bar{Q}_\alpha = \frac{\partial}{\partial\bar{\theta}^\alpha}, \quad Q^\alpha = \frac{\partial}{\partial\theta_\alpha} + \bar{\theta}^\beta\sigma_{\mu\beta}^\alpha\frac{\partial}{\partial x_\mu}.$$

The action of the supersymmetry generator $\bar{\eta}^\alpha \bar{Q}_\alpha - \eta_\alpha Q^\alpha$ in component form is then

$$\delta\varphi^* = \bar{\psi}_\alpha\bar{\eta}^\alpha, \quad \delta\bar{\psi}_\alpha = -\eta_\beta\partial_\alpha^\beta\varphi^* - \bar{\eta}_\alpha E^*, \quad \delta E^* = -\eta_\alpha\partial_\beta^\alpha\bar{\psi}^\beta. \quad (27.42)$$

The expansion of $\bar{\phi}$, at fixed space variable x , reads

$$\begin{aligned} \bar{\phi}(x, \theta, \bar{\theta}) &= \varphi^*(x) + \bar{\psi}^\alpha(x)\bar{\theta}_\alpha - \frac{1}{2}\partial_\alpha^\beta\varphi^*(x)\bar{\theta}^\alpha\theta_\beta - \frac{1}{2}E^*(x)\bar{\theta}^\alpha\bar{\theta}_\alpha \\ &\quad + \frac{1}{4}\partial_\alpha^\beta\bar{\psi}^\alpha(x)\theta_\beta\bar{\theta}^\gamma\bar{\theta}_\gamma + \frac{1}{16}\nabla_x^2\varphi^*(x)\theta^\alpha\theta_\alpha\bar{\theta}^\beta\bar{\theta}_\beta. \end{aligned} \quad (27.43)$$

The free action. The definition (27.29) implies

$$\int d^4\theta (\theta \cdot \theta)(\bar{\theta} \cdot \bar{\theta}) = -4.$$

The Lagrangian

$$\mathcal{L}_D = \bar{\phi}(\bar{y}, \bar{\theta})\phi(y, \theta), \quad (27.44)$$

then corresponds to the free action

$$\int d^4x d^4\theta \mathcal{L}_D = \int d^4x [\nabla_x \varphi^*(x)\nabla_x \varphi(x) - \bar{\psi}^\alpha(x)\partial_\alpha^\beta\psi_\beta(x) + E^*(x)E(x)]. \quad (27.45)$$

In the case of several superfields, one simply adds the corresponding contributions.

Interaction terms. Here we write the F-term contributions for several superfields ϕ_i (we need at least two to construct a Dirac fermion). The most general renormalizable Lagrangian density has the form,

$$\mathcal{L}_F(\phi) = c_i\phi_i + \frac{1}{2}M_{ij}\phi_i\phi_j + \frac{1}{3!}g_{ijk}\phi_i\phi_j\phi_k. \quad (27.46)$$

Integrating over θ_1, θ_2 , one obtains

$$\int i d\theta_2 d\theta_1 \mathcal{L}_F = c_i E_i + M_{ij} (\varphi_i E_j - \frac{1}{2} \psi_i^\alpha \psi_{\alpha j}) + \frac{1}{2} g_{ijk} (\varphi_i \varphi_j E_k - \psi_i^\alpha \psi_{\alpha j} \varphi_k). \quad (27.47)$$

Adding the kinetic term (27.45, 27.47) and its conjugated left-handed contribution,

$$\begin{aligned} \bar{\mathcal{L}}_F(\bar{\phi}) &= c_i^* \bar{\phi}_i + \frac{1}{2} M_{ij}^* \bar{\phi}_i \bar{\phi}_j + \frac{1}{3!} g_{ijk}^* \bar{\phi}_i \bar{\phi}_j \bar{\phi}_k, \\ \int i d\bar{\theta}_2 d\bar{\theta}_1 \bar{\mathcal{L}}_F &= c_i^* E_i^* + M_{ij}^* (\varphi_i^* E_j^* - \frac{1}{2} \bar{\psi}_i^\alpha \bar{\psi}_{\alpha j}) \\ &\quad + \frac{1}{2} g_{ijk}^* (\varphi_i^* \varphi_j^* E_k^* - \bar{\psi}_i^\alpha \bar{\psi}_{\alpha j} \varphi_k^*), \end{aligned} \quad (27.48)$$

one obtains a physically acceptable supersymmetric Lagrangian. A useful exercise is to verify, using the explicit expressions (27.38) and (27.42), that the resulting Lagrangian \mathcal{L} varies in a supersymmetry transformation by a total derivative, and the action thus is invariant.

We can now integrate over E, E^* , since the integral is Gaussian. This is equivalent to using the corresponding field equations:

$$\frac{\partial \mathcal{L}}{\partial E_i} = E_i^* + c_i + M_{ij} \varphi_j + \frac{1}{2} g_{ijk} \varphi_j \varphi_k = 0, \quad (27.49a)$$

$$\frac{\partial \mathcal{L}}{\partial E_i^*} = E_i + c_i^* + M_{ij}^* \varphi_j^* + \frac{1}{2} g_{ijk} \varphi_j^* \varphi_k^* = 0. \quad (27.49b)$$

The Lagrangian can then be cast into the form

$$\begin{aligned} \mathcal{L} &= \nabla_x \varphi^* \nabla_x \varphi - \bar{\psi}^\alpha \sigma_{\mu\alpha}^\beta \partial_\mu \psi_\beta + \frac{\partial \mathcal{L}_F(\varphi)}{\partial \phi_i} \frac{\partial \bar{\mathcal{L}}_F(\varphi^*)}{\partial \phi_i} \\ &\quad - \frac{1}{2} \frac{\partial^2 \mathcal{L}_F(\varphi)}{\partial \phi_i \partial \phi_j} \psi_i^\alpha \psi_{\alpha j} - \frac{1}{2} \frac{\partial^2 \bar{\mathcal{L}}_F(\varphi^*)}{\partial \bar{\phi}_i \partial \bar{\phi}_j} \bar{\psi}_i^\alpha \bar{\psi}_{\alpha j}. \end{aligned}$$

27.4 Vector superfields and gauge invariance

So far, we have constructed supersymmetric actions containing only scalar and spin 1/2 fermion fields. For a realistic theory of particles, vector fields are also required. We now introduce a real *vector superfield*. We can parametrize it in the form

$$\begin{aligned} V(x, \theta, \bar{\theta}) &= B(x) + (\chi(x) \cdot \theta) + (\bar{\chi}(x) \cdot \bar{\theta}) + \frac{1}{2} C(x)(\theta \cdot \theta) + \frac{1}{2} C^*(x)(\bar{\theta} \cdot \bar{\theta}) \\ &\quad + i 2^{-1/2} (\bar{\theta} V(x) \theta) + \frac{1}{2} (\bar{\theta} (\bar{\lambda}(x) - \frac{1}{2} \partial \chi(x))) (\theta \cdot \theta) \\ &\quad + \frac{1}{2} ((\lambda(x) + \frac{1}{2} \partial \bar{\chi}(x)) \cdot \theta) (\bar{\theta} \cdot \bar{\theta}) + \frac{1}{4} (K(x) + \frac{1}{4} \nabla_x^2 B(x)) (\theta \cdot \theta) (\bar{\theta} \cdot \bar{\theta}), \end{aligned} \quad (27.50)$$

in which B, K, V_μ are real fields.

The reason for such a parametrization will become clearer in the following examples. In this form, the vector superfield contains four real scalar fields, four Weyl spinors and one vector field.

27.4.1 Abelian gauge invariance and supersymmetry

Supersymmetry naturally leads to a generalized form of gauge symmetry. A charged scalar superfield transforms under the global $U(1)$ group as $\phi \mapsto e^{-\Lambda} \phi$. However, if one wants to define space-dependent $U(1)$ transformations consistent with supersymmetry, one cannot simply replace Λ by $\Lambda(x)$. One has to replace the constant Λ by parameters of the form of a scalar chiral superfield ((27.37, 27.40)):

$$\phi(y, \theta) \mapsto e^{-\Lambda(y, \theta)} \phi(y, \theta), \quad \text{with } \bar{D}_\alpha \Lambda = 0, \quad (27.51a)$$

$$\bar{\phi}(\bar{y}, \bar{\theta}) \mapsto e^{-\bar{\Lambda}(\bar{y}, \bar{\theta})} \bar{\phi}(\bar{y}, \bar{\theta}), \quad \text{with } D^\alpha \bar{\Lambda} = 0. \quad (27.51b)$$

One immediately notes that, if the charges and couplings are such that the interaction term (27.46) is invariant under global $U(1)$ transformations, it is also invariant under the transformations (27.51). However, the free term (27.44) is not invariant, since $\bar{\phi}\phi \mapsto \bar{\phi}\phi e^{-(\Lambda+\bar{\Lambda})}$. To render the kinetic term gauge invariant, we replace it by

$$\mathcal{L}_D = \bar{\phi} e^V \phi, \quad (27.52)$$

where the vector superfield V must transform like

$$V \mapsto V + (\Lambda + \bar{\Lambda}). \quad (27.53)$$

With the parametrization of Λ as

$$\Lambda(y, \theta) = \eta(y) + \zeta^\alpha(y)\theta_\alpha + \frac{1}{2}D(y)\theta^\alpha\theta_\alpha,$$

and, similarly, for $\bar{\Lambda}$, the gauge transformation in component form reads

$$\begin{aligned} B &\mapsto B + \eta + \eta^*, & \chi &\mapsto \chi + \zeta, \\ C &\mapsto C + D, & V_\mu &\mapsto V_\mu - i 2^{-1/2} \partial_\mu (\eta - \eta^*), \\ \lambda &\mapsto \lambda, & K &\mapsto K. \end{aligned}$$

With our parametrization, λ and K are gauge invariant, while V_μ transforms like a usual Abelian gauge field.

Wess–Zumino gauge. At first sight, it would seem that such a theory cannot be renormalizable, but at least in the Wess–Zumino gauge, the expansion of the exponential reduces to the three first terms (property (27.55)) and contains no term of dimension larger than 4.

Indeed, in a theory that has this kind of gauge invariance, the vector superfield can be simplified, the fields B, C, χ being eliminated by a gauge transformation. The vector field then reduces to

$$V = i 2^{-1/2} \bar{\theta}^\alpha \mathcal{V}_\alpha^\beta \theta_\beta + \frac{1}{2} (\bar{\theta} \cdot \bar{\lambda})(\theta \cdot \theta) + \frac{1}{2} (\lambda \cdot \theta)(\bar{\theta} \cdot \bar{\theta}) + \frac{1}{4} K(\theta \cdot \theta)(\bar{\theta} \cdot \bar{\theta}). \quad (27.54)$$

However, the *Wess–Zumino gauge* breaks the explicit supersymmetry.

The physical degrees of freedom reduce to a massless vector field V_μ and a massless spin 1/2 fermion λ , since the field K does not propagate.

In this gauge, powers of the vector field have the property

$$V^2 = -\frac{1}{4} V_\mu V_\mu (\theta \cdot \theta)(\bar{\theta} \cdot \bar{\theta}), \quad V^n = 0 \text{ for } n \geq 3. \quad (27.55)$$

27.4.2 Supersymmetric curvature tensor

The vector superfield is not chiral and, therefore, as in the one-dimensional example, the operators \bar{D}, D can be used to construct other superfields. Note that quantities of the form $\bar{D}^\beta \bar{D}_\beta W$ and $D^\beta D_\beta W$, where W is an arbitrary superfield, are right- and left-handed chiral fields, respectively, because the product of three operators D or \bar{D} vanishes:

$$\bar{D}_\alpha \bar{D}^\beta \bar{D}_\beta W = D^\alpha D^\beta D_\beta W = 0, \quad \forall W.$$

In particular, the quantities

$$F^\alpha = \frac{1}{2} \bar{D}^\beta \bar{D}_\beta D^\alpha V, \quad \bar{F}_\alpha = \frac{1}{2} D^\beta D_\beta \bar{D}_\alpha V, \quad (27.56)$$

are chiral and gauge invariant (in the sense of the transformation (27.53)), since, for example,

$$\bar{D}^\beta \bar{D}_\beta D^\alpha (\phi + \bar{\phi}) = \bar{D}^\beta \bar{D}_\beta D^\alpha \phi = -\sigma_{\mu\beta}^\alpha \partial_\mu \bar{D}^\beta \phi = 0,$$

(we have used the anticommutation relation (27.33) and the chirality conditions), and thus generalize the gauge field curvature.

To calculate them, it is convenient to express V in terms of the variables y and \bar{y} :

$$\begin{aligned} V &= i2^{-1/2} V_\mu(y) (\bar{\theta} \sigma_\mu \theta) + \frac{1}{2} (\bar{\theta} \lambda(y)) (\theta \cdot \theta) + \frac{1}{2} (\lambda(y) \cdot \theta) (\bar{\theta} \cdot \bar{\theta}) \\ &\quad + \frac{1}{4} [K(y) - i\partial_\mu V_\mu(y)] (\theta \cdot \theta) (\bar{\theta} \cdot \bar{\theta}), \\ &= i2^{-1/2} V_\mu(\bar{y}) (\bar{\theta} \sigma_\mu \theta) + \frac{1}{2} (\bar{\theta} \cdot \bar{\lambda}(\bar{y})) (\theta \cdot \theta) + \frac{1}{2} (\lambda(\bar{y}) \cdot \theta) (\bar{\theta} \cdot \bar{\theta}) \\ &\quad + \frac{1}{4} [K(\bar{y}) + i\partial_\mu V_\mu(\bar{y})] (\theta \cdot \theta) (\bar{\theta} \cdot \bar{\theta}). \end{aligned}$$

One then finds

$$\begin{aligned} F^\alpha &= -\lambda^\alpha(y) - i2^{-3/2} \theta^\beta (\sigma_\nu^\dagger \sigma_\mu)_\beta^\alpha F_{\mu\nu}(y) - K(y) \theta^\alpha + \frac{1}{2} \partial_\mu \bar{\lambda}^\beta(y) \sigma_{\mu\beta}^\alpha (\theta \cdot \theta), \\ \bar{F}_\alpha &= \bar{\lambda}_\alpha(\bar{y}) - i2^{-3/2} (\sigma_\mu \sigma_\nu^\dagger)_\alpha^\beta F_{\mu\nu}(\bar{y}) \bar{\theta}_\beta + K(\bar{y}) \bar{\theta}_\alpha + \frac{1}{2} \sigma_{\mu\alpha}^\beta \partial_\mu \lambda_\beta(\bar{y}) (\bar{\theta} \cdot \bar{\theta}), \end{aligned}$$

with the usual notation $F_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$. We notice that indeed F and \bar{F} depend only on the gauge-invariant combinations $F_{\mu\nu}, \lambda, K$.

A verification of these expressions is provided by the simple relation (derived from the definitions (27.56) after a few commutations)

$$D_\alpha F^\alpha + \bar{D}^\alpha \bar{F}_\alpha = 0.$$

The field F^α is chiral and has dimension 3/2. Therefore, $\mathcal{L}_F = -\frac{1}{4} F^\alpha F_\alpha$ is candidate to contribute to the kinetic term. We then find ($\epsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor with $\epsilon_{1234} = 1$),

$$\int d\theta_1 d\theta^1 F^\alpha F_\alpha = 2\bar{\lambda}^\alpha \sigma_{\mu\alpha}^\beta \partial_\mu \lambda_\beta + K^2 - \frac{1}{2} F_{\mu\nu} F_{\mu\nu} + \frac{1}{4} \epsilon_{\lambda\mu\nu\rho} F_{\lambda\mu} F_{\nu\rho}, \quad (27.57)$$

(the last term is a total derivative and thus is topological, see Section 23.6), where we have used the identity

$$\frac{1}{2} \text{tr} \sigma_\lambda^\dagger \sigma_\mu \sigma_\nu^\dagger \sigma_\rho = \epsilon_{\lambda\mu\nu\rho} + (\delta_{\lambda\mu} \delta_{\nu\rho} + \delta_{\mu\nu} \delta_{\rho\lambda} - \delta_{\lambda\nu} \delta_{\mu\rho}).$$

If we add the conjugated contribution coming from $-\frac{1}{4}\bar{F}^\alpha\bar{F}_\alpha$, we obtain the supersymmetric free gauge action \mathcal{S} , which can be written as

$$\mathcal{S} = \int d^4x \left[\frac{1}{4}F_{\mu\nu}^2(x) - \bar{\lambda}^\alpha(x)\sigma_{\mu\alpha}^\beta\partial_\mu\lambda_\beta(x) - K^2(x) \right]. \quad (27.58)$$

This action could also have been obtained by noting that the vector superfield has dimension 0. A term of the form $V\bar{D}\bar{D}DV$ has dimension 2 and is thus candidate to be a D term.

Finally, note that it is also possible to give a mass to a vector superfield by adding the D contribution of V^2 .

We have thus described all the ingredients required to construct a supersymmetric version of quantum electrodynamics (QED).

27.4.3 Non-Abelian gauge theories

The extension to non-Abelian gauge theories is simple. We denote by t_a the matrices generators of the gauge group G . The gauge transformation is now parametrized by a set of scalar superfields Λ_a , and can be written as

$$\phi \mapsto e^{-\Lambda} \phi, \quad \Lambda \equiv \Lambda_a t_a, \quad \bar{\phi} \mapsto \bar{\phi} e^{-\bar{\Lambda}}, \quad \bar{\Lambda} \equiv \bar{\Lambda}_a t_a.$$

A gauge invariant Lagrangian is

$$\mathcal{L}_D = \bar{\phi} e^V \phi, \quad \text{with } V = V_a t_a,$$

where the gauge field transforms like

$$e^V \mapsto e^{\bar{\Lambda}} e^V e^\Lambda.$$

The generalized form of the curvature F^α is

$$F^\alpha = \bar{D}^\beta \bar{D}_\beta e^{-V} D^\alpha e^V, \quad (27.59)$$

which is chiral, and indeed transforms like

$$F^\alpha \mapsto e^\Lambda F^\alpha e^{-\Lambda}. \quad (27.60)$$

Finally, the contribution to the action takes the form $\text{tr } F^\alpha F_\alpha$.

28 Elements of classical and quantum gravity

This chapter has two purposes: to describe a few elements of differential geometry that appear in different places in this work [305], and to provide, for completeness, a short introduction to a field theory of gravity known under the name of General Relativity (GR), and the problem of its quantization. GR generalizes Relativity, solves one problem of Newton mechanics: the puzzle of the action at a distance of Newton's potential, and explains the cancellation of mass between acceleration term and gravitational force.

We first briefly recall a few concepts related to reparametrization (more accurately, local diffeomorphism) of Riemannian manifolds. We introduce the notions of parallel transport, affine connection, curvature, in analogy with gauge theories as discussed in Chapters 22–26. To define fermions on Riemannian manifolds, additional mathematical objects are required, the vielbein and the spin connection. We then construct Einstein–Hilbert's action for classical gravity (the theory of General Relativity (GR)) and derive the equation of motion (see, *e.g.* Ref. [306]). Finally, in the last section, we describe some formal aspects of the quantization of GR [232, 307–309], following the lines of the quantization of non-Abelian gauge theories of Chapter 22.

Because GR is not renormalizable in four dimensions (even in its extended forms, like supersymmetric gravity), at the present time, a reasonable assumption is that GR is the low-energy, large-distance, remnant (an effective quantum field theory (QFT)) of a more complete theory that probably no longer has the form of a QFT (strings, non-commutative geometry?). In the terminology of critical phenomena, GR belongs to the class of irrelevant interactions: due to the presence of the massless graviton, GR can be compared with an interacting theory of Goldstone modes at low temperature, in the ordered phase. The scale of this new physics seems to be of the order of $\sqrt{\hbar c/G_N} \sim 10^{19}$ GeV (Planck's mass), where G_N is Newton's gravitational constant. Still, because the equations of GR follow from varying Einstein–Hilbert action S_{EH} , some regularized form of $\int e^{iS_{\text{EH}}/\hbar}$ is expected to be relevant to quantum gravity.

In the framework of GR, the presence of a *cosmological constant*, generated by the quantum vacuum-energy density, is expected. The observed acceleration of the expansion of the Universe, attributed to some *dark energy* can be interpreted as resulting from a cosmological constant (also consistent with the measured equation of state), but its value is about 10^{-120} smaller than one would (very naively) guess, pointing, by contrast, towards a natural energy scale of 10^{-3} eV.

The discovery of large scale gravitational phenomena, generally attributed to some missing *dark matter*, has motivated some authors to look, as an alternative, for large-scale modifications of GR, but no compelling model has emerged yet.

In this chapter, we do not focus on mathematical rigour, and our notation is old-fashioned. For instance, we write most expressions in terms of *local coordinates*, ignoring topological aspects and, because it is not essential for our purpose, that several sets of overlapping coordinates (charts) with transition functions are in general required to fully describe a manifold. The framework of fibre bundles is avoided. The reader interested in more details is referred to the literature.

Notation. The convention of *summation over repeated lower and upper indices will always be used, except when the metric is explicitly Euclidean.*

28.1 Manifolds. Change of coordinates. Tensors

Let x^i , $i = 1, \dots, N$ be a set of local coordinates, which parametrize an N -dimensional (real or complex) manifold \mathfrak{M} . A (locally) non-singular change of coordinates or mapping $x \mapsto x'$ is defined by a set of infinitely differentiable functions $x^i(x')$,

$$x^i = x^i(x'), \quad (28.1)$$

such that the mapping $x \mapsto x'$ is locally invertible (the proper extension to the complete manifold is a diffeomorphism).

We define

$$dx^i = T_j^i(x') dx'^j, \text{ where } T_j^i(x') = \frac{\partial x^i}{\partial x'^j}. \quad (28.2)$$

Then, to the mappings $x \mapsto x' \mapsto x''$ corresponds the product $T_j^i(x') T_k^j(x'')$. Since T_j^i is invertible, the matrix T_j^i is an element of the defining representation of the general linear group $GL(N, R \text{ or } C)$ (general invertible matrices).

28.1.1 Fields on manifolds: Classification

We now consider fields defined on \mathfrak{M} , and classify them according to their transformation properties in the change of local coordinates (28.1).

Scalar fields. Scalar fields $S(x)$ transform by the simple substitution,

$$S'(x') = S(x(x')). \quad (28.3)$$

Vector fields. Differentiating equation (28.3), one obtains the transformation of partial derivatives of a scalar field,

$$\frac{\partial S'}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial S}{\partial x^j} = T_j^i(x') \frac{\partial S}{\partial x^j}. \quad (28.4)$$

The partial derivatives transform like vectors (also called *covariant vectors*):

$$V'_i(x') = T_i^j(x') V_j(x). \quad (28.5)$$

Dual vector fields. The transformation of vector fields $W^i(x)$ of the dual space (also called *contravariant vectors*), which belong to the space tangent to the manifold, can be defined by the property that the scalar product $V_i(x) W^i(x)$ is a scalar. This implies the transformation

$$W'^i(x') = [T^{-1}]_j^i(x') W^j(x), \quad (28.6)$$

where $[T^{-1}]_j^i$ is the inverse of the transposed of the matrix T_j^i (and, therefore, belongs to a different representation of $GL(N)$).

Transformations of vectors are thus defined in terms of the substitution $x \mapsto x'$ and a linear transformation.

Tensor fields. More generally, one can classify all fields with respect to their transformation properties under the linear group. General n -tensors $V_{i_{p+1} \dots i_n}^{i_1 \dots i_p}$ transform like the tensor product of p vectors and $(n - p)$ dual vectors:

$$V'^{i_1 \dots i_p}_{i_{p+1} \dots i_n}(x') = [T^{-1}]_{j_1}^{i_1} \dots [T^{-1}]_{j_p}^{i_p} T_{i_{p+1}}^{j_{p+1}} \dots T_{i_n}^{j_n} V^{j_1 \dots j_p}_{j_{p+1} \dots j_n}(x), \quad (28.7)$$

transformation which can be rewritten symbolically as

$$\mathbf{V}'(x') = \mathbf{T}(x') \mathbf{V}(x). \quad (28.8)$$

Remarks

- (i) Summing a n -tensor over a pair of upper and lower indices yields a $(n - 2)$ -tensor. This operation can be called taking a *covariant trace*.
- (ii) The Kronecker δ_i^j is a coordinate-invariant tensor.
- (iii) The representation (28.7) can in general be reduced according to the irreducible representations of the permutation group acting on lower or upper indices.

28.1.2 Infinitesimal change of coordinates

An infinitesimal change of coordinates can be written as

$$x^i = x'^i + \varepsilon^i(x'), \quad (28.9)$$

and all quantities can then be expanded in powers of ε . For example, $T_j^i(x') - \delta_j^i \sim \partial_j \varepsilon^i(x')$ ($\partial_i \equiv \partial/\partial x_i$). We denote by $\delta_\varepsilon V$, the variation at first order in ε of a tensor. One finds

$$\begin{aligned} \delta_\varepsilon V_{i_{p+1} \dots i_n}^{i_1 \dots i_p}(x) &= \varepsilon^j \partial_j V_{i_{p+1} \dots i_n}^{i_1 \dots i_p}(x) - \sum_{\ell=1}^p \partial_j \varepsilon^{i_\ell} V_{i_{p+1} \dots i_n}^{i_1 \dots j \dots i_p}(x) \\ &\quad + \sum_{\ell=p+1}^n \partial_{i_\ell} \varepsilon^j V_{i_{p+1} \dots j \dots i_n}^{i_1 \dots i_p}(x). \end{aligned} \quad (28.10)$$

By construction, $\delta_\varepsilon V$, being the difference between two tensors, is a tensor. Applied to a tensor product, δ_ε satisfies,

$$\delta_\varepsilon(V \otimes W) = (\delta_\varepsilon V) \otimes W + V \otimes (\delta_\varepsilon W).$$

Finally, note the commutation relation

$$[\delta_{\varepsilon_1}, \delta_{\varepsilon_2}] = \delta_{\varepsilon_3}, \quad \text{with} \quad \varepsilon_3^i = \varepsilon_1^j \partial_j \varepsilon_2^i - \varepsilon_2^j \partial_j \varepsilon_1^i. \quad (28.11)$$

28.1.3 Differential forms

Antisymmetric tensors transform under an irreducible representation of $GL(N)$. The simplest example is the completely antisymmetric tensor $\epsilon_{12\dots N}$, which transforms into $\epsilon_{12\dots N} \det(T)$ and, thus, is invariant for unimodular transformations (elements of $SL(N)$).

Antisymmetric tensors are associated with differential forms (see Section 1.5). They can be conveniently represented by contracting indices with the generators θ^i of a Grassmann or exterior algebra (see Section 1.5) *also denoted, when no confusion is possible, as dx (and using then the wedge product \wedge)*:

$$\Omega(\theta, x) = \theta^{i_1} \theta^{i_2} \cdots \theta^{i_n} \Omega_{i_1 i_2 \dots i_n}(x).$$

The form differentiation defined by

$$D_x \equiv \theta^j \frac{\partial}{\partial x_j}, \quad (28.12)$$

satisfies the property of a *cohomology* operator, since $D_x^2 = \theta^j \partial_j \theta^k \partial_k = 0$.

Examples are provided by gauge theories where the vector potential A_μ can be considered as a one-form and the curvature tensor $F_{\mu\nu}$ as a two-form (see equations (21.72) and (22.44)). With the notation (28.12), the two-form \mathbf{F} associated with the curvature tensor $F_{\mu\nu}$ can be related to the gauge field one-form \mathbf{A} by (equation (22.18))

$$\mathbf{F} = 2(D\mathbf{A} + \mathbf{A}^2).$$

The transformation (28.7) acting on Ω can then be written as

$$\Omega'(\theta, x') = \Omega_{i_1 i_2 \dots i_n}(x) \prod_{\ell=1}^n D_{x'} x'^{i_\ell}.$$

The operator D_x has an important property: acting on an n -form, it generates an $(n+1)$ -form, that is, an antisymmetric tensor with $(n+1)$ indices. Indeed,

$$D_x \Omega(\theta, x) = \theta^j \partial_j (\theta^{i_1} \dots \theta^{i_n} \Omega_{i_1 \dots i_n}) = \frac{1}{n+1} \theta^{i_1} \dots \theta^{i_{n+1}} [\partial_{i_{n+1}} \Omega_{i_1 \dots i_n}]_{\text{antisym}}. \quad (28.13)$$

Let us verify that the antisymmetric quantity in the right-hand side is indeed a tensor. Using $D_x^2 = 0$, one finds

$$\begin{aligned} D_{x'} \Omega'(x') &= D_{x'} (D_{x'} x'^{i_1} D_{x'} x'^{i_2} \dots D_{x'} x'^{i_n} \Omega_{i_1 i_2 \dots i_n}(x)) , \\ &= D_{x'} x'^{i_1} D_{x'} x'^{i_2} \dots D_{x'} x'^{i_n} D_{x'} x'^j \partial_j \Omega_{i_1 i_2 \dots i_n}(x), \end{aligned}$$

which proves the tensor property.

A form Ω that satisfies $D\Omega = 0$ is called a *closed form*. A form Ω that can be written as $\Omega = D\Omega'$ is *exact*. Since $D^2 = 0$, any exact form is closed.

Conversely, a standard problem in differential forms is to determine whether a closed form is exact. The corresponding partial differential equations can always be integrated locally but, depending on the topology of the manifold, sometimes no global solution can be found, a situation encountered in the quantization of spin variables (Section A3.3.1).

28.2 Parallel transport: Connection, covariant derivative

We have shown that the gradient $\partial_i S$ of a scalar is a tensor. However, the property that, by differentiation, one generates new tensors is not general: the gradient of a vector is not a tensor, as one easily verifies. What is required is a modified derivative, called *covariant derivative*. To introduce it, we first construct covariant quantities depending on products of vector fields at different points. This requires the concept of *parallel transport*. Covariant derivatives then appear in the limit of infinitesimally close points. An analogous construction is encountered in the context of gauge theories in Chapters 21–26.

28.2.1 Parallel transport

We consider an oriented, continuous, piecewise differentiable curve C on the manifold joining the origin \mathbf{x}_1 to the end-point \mathbf{x}_2 . To this curve, we associate a linear mapping $\mathbf{U}(C)$ from the space tangent to \mathfrak{M} , at point \mathbf{x}_1 to the space tangent at point \mathbf{x}_2 , represented by a matrix with elements $U_j^i(C)$. We call this mapping *parallel transport* from \mathbf{x}_1 to \mathbf{x}_2 . Let $V^i(\mathbf{x}_1)$ be a vector belonging to the space tangent at point \mathbf{x}_1 . The vector

$$V_U^i = U_j^i(C) V^j(\mathbf{x}_1) \quad (28.14)$$

transforms like a vector belonging to the space tangent at the point \mathbf{x}_2 .

We impose to the mapping $\mathbf{U}(C)$ to satisfy several conditions:

If a curve C_{21} goes from \mathbf{x}_1 to \mathbf{x}_2 and a curve C_{32} from \mathbf{x}_2 to \mathbf{x}_3 then,

$$\mathbf{U}(C_{32} \cup C_{21}) = \mathbf{U}(C_{32}) \mathbf{U}(C_{21}). \quad (28.15)$$

The definition (28.15) implies that a curve reduced to a point corresponds to the group identity:

$$\mathbf{U}(C \equiv 1 \text{ point}) = \mathbf{1} \Leftrightarrow U_i^j(C \equiv 1 \text{ point}) = \delta_i^j. \quad (28.16)$$

If C_{21} is a curve going from \mathbf{x}_1 to \mathbf{x}_2 , we denote by C_{21}^{-1} the same curve but oriented from \mathbf{x}_2 to \mathbf{x}_1 . Then, we impose

$$\mathbf{U}(C^{-1}) = \mathbf{U}^{-1}(C). \quad (28.17)$$

The transformation properties of other tensors follow from the following rules: a scalar is invariant in a parallel transport. The condition that the scalar product should be invariant then determines the form of parallel transport for vectors:

$$[V_U]_i = [U^{-1}]_i^j(C) V_j(\mathbf{x}_1). \quad (28.18)$$

We now denote by $\mathbf{U}(C)$ the extension of the linear transformation to all tensors. The tensor product $V \otimes W$ then transforms like

$$\mathbf{U}(C)(V \otimes W) = \mathbf{U}(C)V \otimes \mathbf{U}(C)W. \quad (28.19)$$

Therefore, we can define parallel transport for all tensors, by imposing that a tensor transforms as a tensor product of vectors.

Diffeomorphisms. The condition that parallel transport is compatible with a change of coordinates (equation (28.6)) is

$$T_j^i(\mathbf{x}'_2)[V']_U^j = U_j^i(C) T_k^j(\mathbf{x}'_1) V'^k,$$

and, therefore,

$$U_j^i(C) = T_k^i(\mathbf{x}'_2) U_l^{ik}(C') [T^{-1}]_j^l(\mathbf{x}'_1), \quad (28.20)$$

in which we have assumed that the curve C (resp. C') goes from the point \mathbf{x}_1 to the point \mathbf{x}_2 (resp. \mathbf{x}'_1 to the point \mathbf{x}'_2). Equation (28.20) shows that $U_j^i(C)$ is a (non-local) tensor.

More generally, using the notation (28.8), we can write the effect of the change of coordinates on the parallel transporter $U(C)$ as,

$$\mathbf{U}(C) = \mathbf{T}(\mathbf{x}'_2) \mathbf{U}'(C') \mathbf{T}^{-1}(\mathbf{x}'_1).$$

28.2.2 Affine connection

The composition law (28.15), makes it possible to define parallel transport by parallel transport for infinitesimal curves. We thus consider an infinitesimal differentiable curve, that is, a straight line C connecting two close points \mathbf{x} and $\mathbf{x} + \boldsymbol{\varepsilon}$, and we assume that $U(C)$ can be expanded at first order in $\boldsymbol{\varepsilon}$:

$$\mathbf{U}(C) = \mathbf{1} - \boldsymbol{\Gamma}_k \boldsymbol{\varepsilon}^k + o(\|\boldsymbol{\varepsilon}\|), \quad (28.21)$$

where $\boldsymbol{\Gamma}_k$ is called the affine *connection* in the manifold. The connection is entirely characterized by its action on vectors. We denote by Γ_{ik}^j , the Christoffel symbol, the matrix elements of the connection. Then,

$$U_j^i(C) = \delta_j^i - \Gamma_{jk}^i(x) \boldsymbol{\varepsilon}^k + o(\|\boldsymbol{\varepsilon}\|). \quad (28.22)$$

Conversely, the connection field completely characterizes parallel transport. Indeed, once a curve C is parametrized in terms of a parameter s , the corresponding parallel-transporter $U(C)$ and $\Gamma_{jk}^i(x)dx^k/ds$ play the roles respectively of the evolution operator and the Hamiltonian of quantum mechanics (QM). Therefore, as shown in Section A5, $U(C)$ can be written as a path-ordered integral along the curve C :

$$\mathbf{U} = \text{P exp} \left[- \oint_C \Gamma_k dx^k \right], \quad (28.23)$$

where P stands for path-ordered.

Using equation (28.20), it is possible to determine how the affine connection transforms under reparametrization. It is convenient to express it in terms of the one-form $\boldsymbol{\Gamma} = \Gamma_k dx^k$ associated with the matrix Γ_k . One then finds,

$$\boldsymbol{\Gamma} = \mathbf{T}\boldsymbol{\Gamma}'\mathbf{T}^{-1} - d\mathbf{T}\mathbf{T}^{-1} = \mathbf{T}\boldsymbol{\Gamma}'\mathbf{T}^{-1} + \mathbf{T}d\mathbf{T}^{-1}, \quad (28.24)$$

where $d\mathbf{T}$ is the one-form $dx'^i \partial_i \mathbf{T}$. In component form, in the defining representation, the equation becomes

$$\Gamma'_{jk}^i = (T^{-1})_l^i \partial_k T_j^l + (T^{-1})_l^i \Gamma_{mn}^l T_k^n T_j^m. \quad (28.25)$$

Because the transformation (28.25) is affine, the connection does not transform like a tensor. However, we can decompose the tensor Γ_{jk}^i into symmetric and antisymmetric parts, \mathcal{G}_{jk}^i and \mathcal{T}_{jk}^i , respectively, in the exchange $j \leftrightarrow k$:

$$\Gamma_{jk}^i = \mathcal{G}_{jk}^i + \mathcal{T}_{jk}^i, \quad \text{with} \quad \mathcal{G}_{jk}^i = \mathcal{G}_{kj}^i, \quad \mathcal{T}_{jk}^i = -\mathcal{T}_{kj}^i. \quad (28.26)$$

Because T_j^i has the form (28.2), the inhomogeneous term in (28.25) is symmetric in $j \leftrightarrow k$. Hence, both quantities transform independently under the transformation (28.25). Therefore, the antisymmetric part \mathcal{T}_{jk}^i is a tensor. Moreover, the restriction to connections Γ_{jk}^i symmetric in $j \leftrightarrow k$ (*i.e.* $\mathcal{T}_{jk}^i = 0$) is consistent with the transformation (28.25) and characterizes the special class of *parallel transports without torsion*.

28.2.3 The covariant derivative

It is now possible to compare the vector field at point $\mathbf{x} + \boldsymbol{\varepsilon}$ with the same vector field parallel transported, along a curve C , from the point \mathbf{x} to the point $\mathbf{x} + \boldsymbol{\varepsilon}$. Using the definition (28.21) and equation (28.14), one obtains

$$[\mathbf{U}(C)V]^i(x) - V^i(x + \boldsymbol{\varepsilon}) = - [\partial_k V^i(x) + \Gamma_{jk}^i(x) V^j(x)] \boldsymbol{\varepsilon}^k + o(\|\boldsymbol{\varepsilon}\|). \quad (28.27)$$

The difference in the left-hand side is a vector belonging to the tangent space at $\mathbf{x} + \boldsymbol{\varepsilon}$. In terms of the *covariant derivative* ∇_k , whose action on vectors is

$$\nabla_k V^i = \partial_k V^i + \Gamma_{jk}^i V^j, \quad (28.28)$$

the equation can be written at leading order in $\boldsymbol{\varepsilon}$ as

$$[\mathbf{U}(C)V]^i(x) - V^i(x + \boldsymbol{\varepsilon}) = -\boldsymbol{\varepsilon}^k \nabla_k V^i(x) + o(\|\boldsymbol{\varepsilon}\|). \quad (28.29)$$

The covariant derivative characterizes infinitesimal parallel transport and makes it possible to derive new local tensors from derivatives of tensors.

Parallel transport of general tensors yields the form of the corresponding covariant derivative, which we denote as

$$\nabla_i = \partial_i \mathbf{1} + \boldsymbol{\Gamma}_i, \quad (28.30)$$

and to which one can associate the covariant form differentiation

$$\nabla = d + \boldsymbol{\Gamma} \equiv dx^i (\mathbf{1} \partial_i + \boldsymbol{\Gamma}_i). \quad (28.31)$$

The explicit representation of the covariant derivative depends on the nature of the tensor it is acting on. On scalars, for example, $\nabla_i = \partial_i$. The form of the covariant derivative when acting on a vector is

$$\nabla_i V_j = \partial_i V_j - \Gamma_{ji}^k V_k. \quad (28.32)$$

From equation (28.19), one infers that ∇_i obeys the usual rules of differential operators. It is linear and, moreover, if V and W are two tensors, then for the tensor product $V \otimes W$ one finds

$$\nabla(V \otimes W) = \nabla(V) \otimes W + V \otimes \nabla(W).$$

The general form of the covariant derivative of a tensor with n indices follows:

$$[\nabla_i V]_{j_{p+1} \dots j_n}^{j_1 \dots j_p} = \partial_i V_{j_{p+1} \dots j_n}^{j_1 \dots j_p} + \sum_{\ell=1}^p \Gamma_{k_\ell i}^{j_\ell} V_{j_{p+1} \dots j_n}^{j_1 \dots k_\ell \dots j_p} - \sum_{\ell=p+1}^n \Gamma_{j_\ell i}^{k_\ell} V_{j_{p+1} \dots k_\ell \dots j_n}^{j_1 \dots j_p}, \quad (28.33)$$

which is a tensor with $(n+1)$ indices.

Remarks

(i) For the class of symmetric Christoffel connections Γ_{jk}^i (torsion-free transport), the derivative of a form and its covariant derivative are identical. In the notation (28.13),

$$D \equiv \theta^i \partial_i = \theta^i \nabla_i,$$

as one verifies. In particular, the ordinary curl of a vector is a two-form and thus the covariant curl and the ordinary curl coincide:

$$\partial_i V_j - \partial_j V_i = \nabla_i V_j - \nabla_j V_i. \quad (28.34)$$

(ii) The infinitesimal form of the transformation (28.25) of the Christoffel symbol corresponding to equation (28.9), which is not homogeneous, can be written as

$$\delta_\varepsilon \Gamma_{jk}^i = \partial_j \partial_k \varepsilon^i + \varepsilon^l \partial_l \Gamma_{jk}^i - \partial_l \varepsilon^i \Gamma_{jk}^l + \partial_k \varepsilon^l \Gamma_{jl}^i + \partial_j \varepsilon^l \Gamma_{lk}^i. \quad (28.35)$$

Although this may not be obvious from this expression, $\delta_\varepsilon \Gamma_{jk}^i$ is a tensor.

28.3 Riemannian manifold. The metric tensor

We now restrict \mathfrak{M} to be a real Riemannian manifold. In a Riemannian manifold, which is smooth, one defines the notion of distance. It is characterized by the line element ds , distance from the point x^i to the infinitesimal close point $x^i + dx^i$, such that

$$(ds)^2 = g_{ij}(x)dx^i dx^j,$$

where g_{ij} is the (symmetric) *metric tensor*, and the quadratic form is non-degenerate: $\det \mathbf{g} \neq 0$.

Distances do not depend on the choice of coordinates, and this determines the transformation of the metric tensor $g_{ij}(x)$ in a reparametrization:

$$g_{ij}(x) = (T^{-1})_i^k (T^{-1})_j^l g'_{kl}(x'). \quad (28.36)$$

It is consistent with the notation (28.7) to write the inverse (in the sense of matrices) of the metric tensor as $g^{ij}(x)$:

$$g_{ik}g^{kj} = \delta_i^j. \quad (28.37)$$

The tensors g_{ij} and g^{ij} can be used to lower or raise indices: the metric tensor establishes an isomorphism between the tangent vector space and its dual. It is thus a standard notation to use the same symbol for tensors which are deduced from one of them by lowering or raising indices with the metric tensor, for example,

$$V_i = g_{ij}V^j. \quad (28.38)$$

Compatibility of parallel transport with the metric. Parallel transport is compatible with the metric if it leaves invariant the scalar product of two vectors:

$$V^i(x)g_{ij}(x)W^j(x) = \tilde{V}^i(\tilde{x})g_{ij}(\tilde{x})\tilde{W}^j(\tilde{x}), \quad (28.39)$$

in which \tilde{V}^i and \tilde{W}^i are the parallel-transported of V^i and W^i from \mathbf{x} to $\tilde{\mathbf{x}}$ along a curve C . Using the definition (28.14), one can rewrite the equation as

$$g_{ij}(x) = U_i^k(C)g_{kl}(\tilde{x})U_j^l(C). \quad (28.40)$$

Taking then the limit of an infinitesimal curve, one concludes

$$\nabla_k g_{ij} = 0. \quad (28.41)$$

Determination of the connection in torsion-free parallel transport. We have shown that the conservation of the scalar product of vectors under parallel transport implies the compatibility condition (28.41). Using the explicit form (28.33), it can be written as

$$\partial_k g_{ij} - g_{il}\Gamma_{jk}^l - \Gamma_{ik}^l g_{lj} = 0. \quad (28.42)$$

If, we consider only torsion-free parallel transport and, thus, assume Γ_{jk}^i symmetric, we can solve the compatibility condition and express Γ_{jk}^i explicitly in terms of the metric tensor. We find

$$\Gamma_{jk}^i = \frac{1}{2}g^{il}(\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk}). \quad (28.43)$$

Compatibility with the metric *uniquely* determines torsion-free parallel transport.

28.3.1 Covariant volume element

To construct, for example, a classical action, we need a volume element invariant under reparametrization. The Euclidean measure transforms like

$$\prod_i dx^i = \det T_l^k \prod_i dx'^i. \quad (28.44)$$

We denote by g the determinant of the metric g_{ij} :

$$g = \det \mathbf{g}. \quad (28.45)$$

It transforms like

$$g' = \det \mathbf{g}' = (\det T_j^i)^2 g. \quad (28.46)$$

Therefore, the volume element

$$dv(x) = \sqrt{g} \prod_{i=1}^N dx^i, \quad (28.47)$$

is invariant. More generally, the volume element $dv(x)$ multiplied by a scalar (like a function of the scalar curvature) is also invariant.

Remark. Differentiating the identity

$$\ln g \equiv \ln \det \mathbf{g} = \text{tr} \ln \mathbf{g}, \quad (28.48)$$

one finds

$$g^{-1} \partial_k g = g^{ij} \partial_k g_{ji}. \quad (28.49)$$

Comparing the right-hand side with expression (28.43) for the connection, one infers the simple relation

$$\Gamma_{ki}^k = \frac{1}{\sqrt{g}} \partial_i \sqrt{g}. \quad (28.50)$$

In particular, it implies for the covariant divergence of a vector,

$$\nabla_i V^i = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} V^i) = \partial_i V^i + (\partial_i \ln \sqrt{g}) V^i, \quad (28.51)$$

and, therefore,

$$\int \left(\prod_i dx^i \right) \sqrt{g} \nabla_j V^j = \int \left(\prod_i dx^i \right) \partial_j (\sqrt{g} V^j). \quad (28.52)$$

28.4 The curvature (Riemann) tensor

We consider an infinitesimal parallelogram C in the manifold, joining the points: $x, x + \varepsilon_1, x + \varepsilon_1 + \varepsilon_2, x + \varepsilon_2$, and back to x (see Fig. 28.1), and calculate the corresponding parallel transporter:

$$\begin{aligned} \mathbf{U}(C) &= \mathbf{U}^{-1}(x + \varepsilon_2, x) \mathbf{U}^{-1}(x + \varepsilon_1 + \varepsilon_2, x + \varepsilon_2) \\ &\times \mathbf{U}(x + \varepsilon_1 + \varepsilon_2, x + \varepsilon_1) \mathbf{U}(x + \varepsilon_1, x). \end{aligned} \quad (28.53)$$

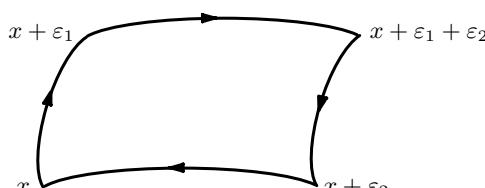


Fig. 28.1 The loop C

We parametrize the expansion of $\ln \mathbf{U}(C)$ up to second order in ε_1 and ε_2 as

$$\ln \mathbf{U}(x + \varepsilon, x) = -\boldsymbol{\Gamma}_i \varepsilon^i + \frac{1}{2} \boldsymbol{\Gamma}_{ij} \varepsilon^i \varepsilon^j + O(\varepsilon^2), \quad (28.54)$$

in which $\boldsymbol{\Gamma}_i$ is the connection.

The right-hand side of equation (28.53) can be evaluated by using Baker–Hausdorff's formula repeatedly (the calculation follows the lines of the derivation of the commutation relations of generators in a Lie algebra, see also Section 25.2.2):

$$e^A e^B = e^{(A+B+[A,B]/2+\dots)}.$$

The first order in ε in the exponential of the right-hand side vanishes as a consequence of equation (28.17). At second order, one finds

$$\mathbf{U}(C) = \mathbf{1} - \varepsilon_1^i \varepsilon_2^j \mathbf{R}_{ij} + O(\varepsilon^3),$$

where the antisymmetric tensor \mathbf{R}_{ij} is given by

$$\mathbf{R}_{ij} = [\nabla_i, \nabla_j] = \partial_i \boldsymbol{\Gamma}_j - \partial_j \boldsymbol{\Gamma}_i + [\boldsymbol{\Gamma}_i, \boldsymbol{\Gamma}_j]. \quad (28.55)$$

We have obtained the expression of the *curvature tensor* \mathbf{R}_{ij} (or Riemann tensor) in terms of the connection. The curvature tensor characterizes the variation of tensors in a transport along infinitesimal closed curves.

Note that we could have used the expression (28.23) to perform the calculation for an arbitrary closed curve. We fix one point x on the curve and write a generic point $x + \varepsilon$. Expanding the path-ordered integral up to second order, we find

$$\mathbf{U}(C) = \mathbf{1} - \oint \boldsymbol{\Gamma}_k(x + \varepsilon) d\varepsilon^k + \frac{1}{2} \oint \oint d\varepsilon^k d\varepsilon^l P[\boldsymbol{\Gamma}_k(x + \varepsilon) \boldsymbol{\Gamma}_l(x + \varepsilon)] + O(\varepsilon^3).$$

We then expand $\boldsymbol{\Gamma}_k$ for ε small. In the first integral, the term proportional to $\boldsymbol{\Gamma}_k(x)$ vanishes because the curve is closed. In the second integral, we can neglect the dependence in ε at this order. A short calculation then yields

$$\mathbf{U}(C) - \mathbf{1} \sim -\frac{1}{2} \int_D \mathbf{R},$$

where \mathbf{R} is the two-form $\mathbf{R}_{ij} dx^i \wedge dx^j$ (see equation (28.31)):

$$\mathbf{R} = 2\nabla^2 = 2(d\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^2),$$

and D , the domain of integration, a surface which has the curve C as boundary: $\partial D = C$.

Finally, the curvature tensor \mathbf{R}_{ij} is characterized by its matrix elements, when acting on vectors,

$$[\mathbf{R}_{ij} V]^k \equiv R_{lij}^k V^l.$$

Equation (28.55) in component form reads

$$R_{lij}^k = \partial_i \Gamma_{lj}^k - \partial_j \Gamma_{li}^k + \Gamma_{mi}^k \Gamma_{lj}^m - \Gamma_{mj}^k \Gamma_{li}^m. \quad (28.56)$$

A general tensor transforms like

$$[\mathbf{R}_{ij} V]_{l_1 l_2 \dots l_q}^{k_1 k_2 \dots k_p} = \sum_{r=1}^p R_{mij}^r V_{l_1 l_2 \dots l_q}^{k_1 k_2 \dots m \dots k_p} - \sum_{r=1}^q R_{l_r ij}^m V_{l_1 l_2 \dots m \dots l_q}^{k_1 k_2 \dots k_p}. \quad (28.57)$$

28.4.1 Curvature tensor and metric

When parallel transport is torsion-free and compatible with the metric, the curvature tensor is determined by the metric tensor. A short calculation yields for the tensor with only lower indices, $R_{klkj} = g_{km}R_{lij}^m$, the expression

$$\begin{aligned} R_{klkj} &= \frac{1}{2}(\partial_k\partial_l g_{kj} - \partial_i\partial_k g_{jl}) - \frac{1}{4}(\partial_i g_{mk} + \partial_k g_{mi} - \partial_m g_{ik}) \\ &\quad \times g^{mn}(\partial_j g_{nl} + \partial_l g_{nj} - \partial_n g_{jl}) - (i \leftrightarrow j). \end{aligned} \quad (28.58)$$

A few properties. Being defined as a commutator, R_{lij}^k satisfies

$$R_{lij}^k = -R_{lji}^k. \quad (28.59)$$

For the same reason, the curvature tensor satisfies the consequence of the Jacobi identity,

$$[\nabla_i, [\nabla_j, \nabla_k]] + \text{cyclic permutations } (i, j, k) = 0,$$

which, in terms of R_{lij}^k , reads

$$\nabla_i R_{mjk}^l + \nabla_j R_{mki}^l + \nabla_k R_{mij}^l = 0, \quad (28.60)$$

and, in the framework of Riemannian geometry, is called the *Bianchi identity*.

Note that the compatibility condition (28.41) implies

$$[\nabla_i, \nabla_j]g_{kl} = 0 = R_{kij}^m g_{ml} + R_{lij}^m g_{km}.$$

This equation implies that the tensor R_{klkj} is antisymmetric in $k \leftrightarrow l$:

$$R_{klkj} = -R_{lkij}, \quad (28.61)$$

a property that can be verified on the explicit form (28.58).

We also note that R_{klkj} is antisymmetric in $(k \leftrightarrow l)$, and symmetric in the exchange $(kl) \leftrightarrow (ij)$:

$$R_{klkj} = R_{ijkl}. \quad (28.62)$$

It is simple to verify the cyclic identity:

$$R_{ijkl} + R_{kijl} + R_{jkil} = 0. \quad (28.63)$$

The covariant trace of the curvature tensor,

$$R_{ij} = R_{ijk}^k, \quad R_{ij} = R_{jij}, \quad (28.64)$$

is called the *Ricci tensor*. It is a symmetric tensor and, therefore, has properties analogous to the metric. Taking the covariant trace again, one obtains the *scalar curvature*,

$$R = R_i^i \equiv R^{ij}g_{ji} \equiv R_{kl}^{kl}. \quad (28.65)$$

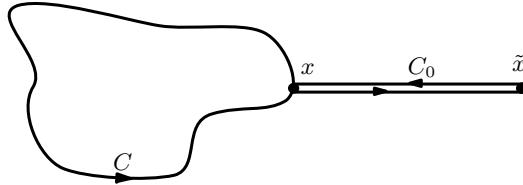


Fig. 28.2 Holonomy group: independence of the initial point

Remarks

(i) An important problem is to classify local tensors which are functionals of the metric tensor. The compatibility condition (28.41) implies that the covariant derivative of g_{ij} is not a new tensor.

However, we note that when the connection is a functional of the metric tensor, this also applies to the curvature tensor. All tensors depending only on the metric can then be obtained from the curvature tensor and its covariant derivatives.

(ii) In the case of symmetric Christoffel connections, the variation (28.35) can be rewritten in a simple way, which exhibits its tensor character, as

$$\delta_\varepsilon \Gamma_{jk}^i = \nabla_k \nabla_j \varepsilon^i - R_{jkl}^i \varepsilon^l. \quad (28.66)$$

The symmetry of the right-hand side in the exchange $j \leftrightarrow k$ relies on the cyclic identity (28.63).

28.4.2 Holonomy variables, holonomy group

The parallel-transporters $\mathbf{U}(C)$ associated with closed curves C are often called *holonomy variables*. In particular, they probe topological properties of a manifold \mathfrak{M} , which we assume to have dimension N . This is especially clear in the case of parallel transport associated with a vanishing curvature tensor. In the latter case, the contour C can be continuously deformed without changing the corresponding holonomy variable.

In this context, it can be useful to consider the group formed by holonomies $\mathbf{U}(x; C)$ originating from a common fixed point x , called the *holonomy group*.

Independence of the holonomy group of the initial point. The holonomy groups associated with different initial points are isomorphic. Indeed, if x and \tilde{x} are two different initial points, we introduce a fixed curve C_0 joining them. To each curve C passing through x , we can now associate a curve \tilde{C} passing through \tilde{x} (see Fig. 28.2):

$$\tilde{C} = C_0^{-1} \cup C \cup C_0. \quad (28.67)$$

In this way, we have constructed a one-to-one mapping between curves passing through x and curves passing through \tilde{x} . The corresponding relation between parallel transporters is

$$\mathbf{U}(\tilde{C}) = \mathbf{U}(C_0)\mathbf{U}(C)\mathbf{U}^{-1}(C_0). \quad (28.68)$$

It establishes an isomorphism between the holonomy groups associated with the points x and \tilde{x} . Therefore, the abstract holonomy group is independent of the initial point, and intrinsic to the set of curves in the manifold equipped with the equivalence relation (28.67).

Holonomy and metric. For a general parallel transport, the compatibility condition (28.40) for a closed curve reads

$$g_{ij}(x) = U_i^k(C)g_{kl}(x)U_l^j(C). \quad (28.69)$$

If the metric is positive, equation (28.69) shows that the matrices $\mathbf{U}(x; C)$, corresponding to closed curves passing through a point x , belong to a subgroup of the $O(N)$ orthogonal group. Let us now consider the curve C formed by the infinitesimal parallelogram of Fig. 28.1. If we expand equation (28.69) up to second order in ε , we obtain a condition which, in terms of the curvature tensor, can be written as

$$\varepsilon_1^k \varepsilon_2^l (R_{ijkl} + R_{jikl}) = 0. \quad (28.70)$$

This result provides a geometric interpretation for the antisymmetry (28.61) of the curvature tensor: it is the antisymmetry of the generators of the orthogonal group.

28.5 Fermions, vielbein, spin connection

Notation. In what follows, we use the first letters of the alphabet a, b, c, d, \dots to represent indices corresponding to the Euclidean metric, and the letters i, j, k, \dots to represent tensor indices. Again, we assume that the dimension of \mathfrak{M} is N .

We now briefly explain how one can define spin 1/2 fermions living on Riemannian manifolds, because the construction is not trivial.

28.5.1 Local frame: Vielbein

It is first necessary to introduce a local frame in the space tangent to the manifold (this can always be done locally but may lead to topological obstructions). The set of vectors $e^{ai}(x)$, which form a local basis, is called the *vielbein*. We assume that the metric is positive. We can then choose vectors orthogonal with respect to the metric $g_{ij}(x)$, of length 1:

$$e^{ai}(x) g_{ij}(x) e^{bj}(x) = \delta_{ab}. \quad (28.71)$$

We do not distinguish upper and lower internal indices a, b because, as equation (28.71) shows, the internal metric is Euclidean. Introducing the vectors e_i^a , obtained as usual from the e^{ai} by lowering the index with the metric tensor, we can rewrite equation (28.71) as

$$e^{ai} e_i^b = e_i^a e^{bi} = \delta_{ab}, \quad (28.72)$$

which shows that the matrix e^{ai} is the inverse of the matrix e_i^b . Finally, combining equations (28.72) and (28.71), one finds

$$g_{ij}(x) = e_i^a(x) e_j^a(x), \quad (28.73)$$

an equation that expresses the metric tensor in terms of the vielbein.

The metric is invariant under orthogonal transformations acting on the local frame:

$$e_i^a(x) = O_{ab}(x) (e')_i^b(x), \quad O_{ab}(x) O_{cb}(x) = \delta_{ac}. \quad (28.74)$$

Spinors then transform under the spin group $\text{Spin}(N)$ associated with the local $O(N)$ group. As described in Section 12.3.2, this implies that the spinors ψ and $\bar{\psi}$ transform like

$$\begin{aligned} (\psi)_\alpha(x) &= \Lambda_{\alpha\beta}^{-1} \psi'_\beta(x), \\ (\bar{\psi})_\alpha(x) &= \bar{\psi}'_\beta(x) \Lambda_{\beta\alpha}, \end{aligned} \quad (28.75)$$

in which the matrix O_{ab} is given in terms of Λ by (γ_a is a Dirac matrix)

$$\Lambda \gamma_b O_{ab} \Lambda^{-1} = \gamma_a. \quad (28.76)$$

With these transformation properties, the quantity $\bar{\psi}\gamma^a e^{ai}\psi$ is a tensor under reparametrization, independent of the local frame.

We recall that if Λ is written as

$$\Lambda = \exp\left(\frac{i}{4}\theta_{ab}\sigma_{ab}\right), \quad (28.77)$$

in which θ_{ab} is an antisymmetric real matrix, then,

$$O_{ab} = (e^\theta)_{ab}. \quad (28.78)$$

28.5.2 Gauge invariance and spin connection

Since the choice of the local frame is arbitrary, one requires invariance of all physical quantities under the transformations of the local orthogonal group (28.74). To enforce this condition, one introduces a new parallel transporter, which takes the form of an orthogonal curve-dependent matrix $\mathbf{O}(C)$. In a change of local frame, it transforms like

$$\mathbf{O}'(C) = \mathbf{O}(\tilde{x})\mathbf{O}(C)\mathbf{O}^{-1}(x), \quad (28.79)$$

in which the curve C goes from x to \tilde{x} . Such transformations are called *gauge transformations* (for details see Chapter 22). For infinitesimal differentiable curves, $\mathbf{O}(C)$ can be expressed in terms of the *spin connection* ω_i^{ab} , which has the form of a gauge field. The connection is a vector on the manifold, and a matrix belonging to the Lie algebra of $O(N)$:

$$\mathbf{O}(C) = \mathbf{1} - \boldsymbol{\omega}_i(x)dx^i + o(\|dx^i\|). \quad (28.80)$$

In the case of the gauge theories discussed in Chapter 22, the gauge field is a new independent mathematical quantity. However, here we note that if we consider the matrix (which is a scalar under reparametrization)

$$U^{ab}(C) = e_i^a(\tilde{x})U_j^i(C)e^{jb}(x), \quad (28.81)$$

then, as a consequence of the compatibility condition (28.40) and equation (28.73), it satisfies

$$(U^T(C))^{ac}U^{cb}(C) = \delta_{ab}, \quad (28.82)$$

and is, therefore, orthogonal.

Equation (28.74) implies that it also satisfies equation (28.79). Therefore, the matrix $U^{ab}(C)$ has the properties required from $O^{ab}(C)$, and one identifies $O^{ab}(C) \equiv U^{ab}(C)$. Equation (28.81), in the limit of an infinitesimal curve, yields the relation between the spin connection ω_i^{ab} and the connection,

$$\Gamma_{ik}^j = e^{aj}\partial_k e_i^a + e^{aj}e_i^b\omega_k^{ab}. \quad (28.83)$$

Equation (28.43) expresses Γ_{ik}^j in terms of the metric tensor. It follows that the spin connection ω_i^{ab} can be expressed in terms of the vielbein, which replaces the metric tensor as the basic geometric quantity in theories with fermions and thus spinors:

$$\omega_i^{ab} = \frac{1}{2}e^{aj}(\partial_i e_j^b - \partial_j e_i^b) + \frac{1}{4}e^{aj}e^{bk}(\partial_k e_j^c - \partial_j e_k^c)e_i^c - (a \leftrightarrow b). \quad (28.84)$$

By expanding all tensors on the basis formed by the vielbein, one can replace the condition of covariance under reparametrization by the condition of independence of the local reference frame (gauge invariance).

A covariant derivative ∇_i is associated to the connection ω_i^{ab} , which, on a vector of components $V^a = V^i e_i^a$, acts like

$$\nabla_i V^a = \partial_i V^a + \omega_i^{ab}V^b. \quad (28.85)$$

Finally, one defines a general covariant derivative, whose action on Euclidean indices is given by equation (28.85), and its action on tensor indices by equation (28.33). Let us then calculate, for example,

$$\nabla_i e_j^a = \partial_i e_j^a + \omega_i^{ab} e_j^b - \Gamma_{ij}^k e_k^a. \quad (28.86)$$

With this definition, equation (28.83) takes the simple form

$$\nabla_i e_j^a = 0, \quad (28.87)$$

and, therefore,

$$\nabla_i V^a = e_j^a \nabla_i V^j. \quad (28.88)$$

This last equation directly follows from the definition of the covariant derivative and equation (28.81).

A last remark: when C is a closed curve, equation (28.81) becomes a similarity relation between matrices. In particular, the curvature tensor R_{kl}^{ab} , associated with the connection ω_i^{ab} , is simply related to the Riemann tensor by the equation

$$R_{kl}^{ab} = R_{ijkl} e_i^a e_j^b = -R_{kl}^{ba}. \quad (28.89)$$

Covariant derivative and fermions. In the case of fermions, the covariant derivative ∇ takes a form that follows from the considerations of Chapter 22, the transformation properties of spinors and the expression of matrices of the spinor representation (28.77):

$$\nabla = e^{ai} \gamma^a (\partial_i + \frac{i}{4} \sigma_{bc} \omega_i^{bc}). \quad (28.90)$$

28.6 Classical GR. Equations of motion

In the theory of GR, the metric tensor field $g_{ij}(x)$ becomes a dynamical variable. Using the previous framework, one can construct a classical action for a metric tensor coupled to matter, which is relativistic, *local* (unlike Newton's theory) and purely geometrical, that is, independent of the system of coordinates in the manifold.

28.6.1 The classical action

The pure gravity action $S_{\text{GR}}(\mathbf{g})$ must have the form of a scalar integrated with an invariant volume element. If we look for a local action with only two derivatives, we see that the scalar curvature is the only candidate. We then obtain Einstein's (or Einstein–Hilbert) action,

$$S_{\text{GR}}(\mathbf{g}) = \int d\nu(x) R(\mathbf{g}(x)), \quad (28.91)$$

where $d\nu(x)$ is defined in equation (28.47).

To Einstein's gravity action, one can add the derivative-free term, called *cosmological term*,

$$S_{\text{cosmo.}}(\mathbf{g}) = \Lambda \int d\nu(x). \quad (28.92)$$

More generally, covariance and locality alone allow any function of the curvature tensor and the scalars obtained by contracting covariant derivatives of the curvature tensor.

Scalar fields. The simplest covariant action for a scalar matter field $\phi(x)$ coupled to gravity takes the form

$$\mathcal{S}_{\text{scalar}}(\phi, \mathbf{g}) = \int dv(x) \left[\frac{1}{2} g^{ij}(x) \partial_i \phi(x) \partial_j \phi(x) + \mathcal{V}(\phi(x)) \right]. \quad (28.93)$$

Gauge fields. In the case of vector field \mathbf{A}_i associated to an external gauge symmetry of the type discussed in Chapter 22, we note that the gauge field curvature (equation (22.16))

$$\mathbf{F}_{ij} = \partial_i \mathbf{A}_j - \partial_j \mathbf{A}_i + [\mathbf{A}_i, \mathbf{A}_j],$$

is a tensor because it is a two-form (see equation (28.34)). Therefore, a possible action is

$$\mathcal{S}_{\text{gauge}}(\mathbf{A}, \mathbf{g}) = \frac{1}{4} \int dv(x) \text{tr } \mathbf{F}_{ij}(x) \mathbf{F}^{ij}(x). \quad (28.94)$$

Fermion matter coupled to gravity. The considerations of Section 28.5.1, in particular, the definitions of the vielbein and the covariant derivative (28.90), make it possible to define the action for Dirac fermion matter coupled to gravity

$$\mathcal{S}_{\text{fermion}}(\psi, \bar{\psi}, \mathbf{e}) = - \int d^N x \det \mathbf{e} \bar{\psi}(x) (\not{\nabla} + M) \psi(x). \quad (28.95)$$

28.6.2 Classical equation of motion

Let us derive the equation of motion for pure gravity, obtained from the variation of the action (28.91):

$$\mathcal{S}_{\text{GR}}(\mathbf{g}) = \int dv(x) R(\mathbf{g}).$$

Note the identity (be careful about the signs),

$$R_{lij}^k = \nabla_i \Gamma_{lj}^k - \nabla_j \Gamma_{li}^k - \Gamma_{mi}^k \Gamma_{lj}^m + \Gamma_{mj}^k \Gamma_{li}^m. \quad (28.96)$$

Using this identity and integrating by parts, one can also write the action as

$$\mathcal{S}_{\text{GR}}(\mathbf{g}) = \int dv(x) g^{il} \left(-\Gamma_{mi}^j \Gamma_{lj}^m + \Gamma_{mj}^j \Gamma_{li}^m \right).$$

If one adds a cosmological term (equation (28.92)), the action becomes

$$\mathcal{S}(\mathbf{g}) = \mathcal{S}_{\text{GR}}(\mathbf{g}) + \Lambda \int dv(x). \quad (28.97)$$

Equation of motion. We have to calculate the variation of various quantities when the metric g_{ij} varies:

$$\delta_h g_{ij} = h_{ij}. \quad (28.98)$$

Note that the variation of the connection is a tensor (the inhomogeneous part of the transformation cancels in the variation). Indeed, after a short calculation one finds

$$\delta_h \Gamma_{jk}^i = \frac{1}{2} g^{il} (\nabla_j h_{lk} + \nabla_k h_{lj} - \nabla_l h_{jk}). \quad (28.99)$$

In the same way, one can calculate the variation of the curvature in terms of $\delta\Gamma$. Since $\delta\Gamma$ is a tensor, one finds, as expected,

$$\delta_h R_{lij}^k = \nabla_i \delta\Gamma_{jl}^k - \nabla_j \delta\Gamma_{il}^k.$$

Substituting, one obtains,

$$\delta_h R_{lij}^k = \frac{1}{2} g^{km} ([\nabla_i, \nabla_j] h_{ml} + \nabla_i \nabla_l h_{mj} - \nabla_j \nabla_l h_{mi} - \nabla_i \nabla_m h_{lj} + \nabla_j \nabla_m h_{il}). \quad (28.100)$$

Since

$$[\nabla_i, \nabla_j] h_{ml} = -R_{mij}^n h_{nl} - R_{lij}^n h_{nm},$$

one finds,

$$\begin{aligned} \delta_h R &= \delta \left(g^{li} R_{lij}^j \right) = -R^{ij} h_{ij} + g^{li} \delta R_{lij}^j, \\ &= -R^{ij} h_{ij} + \frac{1}{2} (g^{il} g^{jm} - g^{im} g^{jl}) \nabla_i \nabla_l h_{mj}. \end{aligned}$$

Moreover,

$$\begin{aligned} \delta_h (\sqrt{g} R) &= \frac{1}{2\sqrt{g}} g^{ij} h_{ij} R + \sqrt{g} \delta_h R, \\ &= \frac{1}{2\sqrt{g}} g^{ij} h_{ij} R - \sqrt{g} R^{ij} h_{ij} + \frac{1}{2} (g^{il} g^{jm} - g^{im} g^{jl}) \nabla_i \nabla_l h_{mj}. \end{aligned}$$

Finally, integrating by parts in the action and using $\nabla_i g_{jk} = 0$, one obtains

$$\delta_h \mathcal{S} = \int dv(x) h_{ij} \left(\frac{1}{2} g^{ij} R - R^{ij} \right),$$

and, thus, the equation of motion

$$R g^{ij} - 2R^{ij} = 0. \quad (28.101)$$

Note that in two dimensions, the curvature tensor has only one component and thus $R^{ij} \propto g^{ij}$. Equation (28.101) reduces to only one equation. Then, taking the covariant trace, one finds (in N dimensions)

$$(N-2)R = 0,$$

an equation that is identically satisfied in two dimensions. We recover Gauss–Bonnet’s theorem, which implies that *Einstein’s action in two dimensions is topological*.

The cosmological constant. The action (28.97) with a cosmological term leads to the equation of motion

$$(R + \Lambda)g^{ij} - 2R^{ij} = 0. \quad (28.102)$$

Taking the covariant trace, one finds

$$(N-2)R + N\Lambda = 0,$$

which shows that the cosmological constant induces a non-vanishing curvature of space, even in the absence of matter.

Matter coupled to gravity. When matter is coupled to gravity, for example, the action terms (28.93) and (28.95) are added to the gravity action, the equation of motion becomes

$$(R + \Lambda)g^{ij} - 2R^{ij} + \mathcal{T}^{ij} = 0, \quad (28.103)$$

where the quantity

$$\mathcal{T}^{ij} = 2 \frac{\delta \mathcal{S}_{\text{matter}}}{\delta g_{ij}},$$

in the limit $g_{ij} \rightarrow \delta_{ij}$, coincides with the energy–momentum tensor defined in Section A13.2.

28.7 Quantization in the temporal gauge: Pure gravity

It is simple to verify that, due to the covariance of the equations of motion under local change of coordinates, it is impossible to quantize the theory in the standard way, because the time components of the metric tensor have no conjugate momenta and thus generate only constraints [307, 232, 308]. This is a problem one also encounters in gauge theories, and the same strategy can be used to solve it. Here, we choose to quantize the theory in the temporal gauge, and we consider, for simplicity, only pure gravity.

By a change of coordinates, one can reduce the metric to the form $g_{00} = 1$ and $g_{0i} = 0$ for $i \neq 0$. If one then specializes the action to such metrics one obtains, as equation of motions, the space components of equations (28.101). The remaining equations have to be imposed as additional constraints. As a notation, *we below use the letters a, b, c, and d for space indices.*

28.7.1 The action in the temporal or Weyl gauge

We first calculate the components of the curvature tensor:

$$\begin{aligned} R_{a00d} &= R_{0ad0} = \frac{1}{2}\partial_0^2 g_{ad} - \frac{1}{4}\partial_0 g_{am} g^{mn} \partial_0 g_{nd}, \\ R_{abcd} &= R_{abcd}^{\text{sp.}} - \frac{1}{4}\partial_0 g_{ac} \partial_0 g_{bd} + \frac{1}{4}\partial_0 g_{bc} \partial_0 g_{ad}, \end{aligned}$$

where $R_{abcd}^{\text{sp.}}$ is the curvature tensor in the $(N - 1)$ dimensional space. It follows that

$$\begin{aligned} R_{00} &= \frac{1}{2}g^{ad}\partial_0^2 g_{ad} + \frac{1}{4}\partial_0 g^{ab}\partial_0 g_{ab}, \\ R &= R_{abcd}g^{ad}g^{bc} + 2R_{00}, \\ &= R^{\text{sp.}} + \frac{3}{4}\partial_0 g^{ab}\partial_0 g_{ab} + \frac{1}{4}(g^{ab}\partial_0 g_{ab})^2 + g^{ab}\partial_0^2 g_{ab}. \end{aligned}$$

We have to integrate by parts to eliminate the second derivatives:

$$\sqrt{g}g^{ab}\partial_0^2 g_{ab} = -\sqrt{g}\left[\partial_0 g_{ab}\partial_0 g^{ab} + \frac{1}{2}(g^{ab}\partial_0 g_{ab})^2\right] + \text{total derivatives}.$$

Finally, we obtain

$$\mathcal{S}_{\text{temp}} = \int d^N x \mathcal{L}, \quad \text{with } \mathcal{L} = \sqrt{g}\left[R^{\text{sp.}} - \frac{1}{4}\partial_0 g_{ab}\partial_0 g^{ab} - \frac{1}{4}(g^{ab}\partial_0 g_{ab})^2\right]. \quad (28.104)$$

The corresponding conjugate momenta are

$$\Pi^{ab} = -\frac{1}{2}\sqrt{g}(\partial_0 g^{ab} - g^{ab}g_{cd}\partial_0 g^{cd}) = -\frac{1}{2\sqrt{g}}\partial_0(gg^{ab}), \quad (28.105)$$

and, conversely,

$$\partial_0 g^{ab} = -\frac{2}{\sqrt{g}}\left(\Pi^{ab} - \frac{g^{ab}}{N-2}\Pi_c^c\right). \quad (28.106)$$

The Hamiltonian density follows,

$$\mathcal{H} = \Pi^{ab}\partial_0 g_{ab} - \mathcal{L} = \frac{1}{\sqrt{g}}\left[\Pi_{ab}\Pi^{ab} - \frac{1}{N-2}(\Pi_c^c)^2\right] - \sqrt{g}R^{\text{sp.}}. \quad (28.107)$$

We note that covariance implies that the quadratic form in the momenta is a homogeneous function of the metric tensor. Therefore, ordering problems cannot be avoided, and this reflects in a functional measure which is not flat in metric space, but multiplied by a power of g :

$$\mathcal{Z} = \int [dg_{ij}g^{N(N-5)/8}] \prod_{i=0,N-1} \delta(g_{0i}) \exp[-\mathcal{S}(\mathbf{g})].$$

It is then possible to use the same functional techniques as in gauge theories, and introduce covariant gauges, which lead to a BRST symmetric quantized action.

Constraints. We must still study the constraints which are

$$R^{0c} = 0, \quad R - 2R^{00} = 0. \quad (28.108)$$

The first set is a simple generalization of Gauss's law in gauge theories:

$$R^{0c} = 0 = g^{ad}g^{bc}R_{ab0d} = -\frac{1}{\sqrt{g}}\nabla_a^{\text{sp.}}\Pi^{ac}, \quad (28.109)$$

where $\nabla_a^{\text{sp.}}$ is the covariant derivative in $(N-1)$ -dimensional space.

Note that, to prove this result, one can use

$$\nabla_c^{\text{sp.}}g_{ab} = 0. \quad (28.110)$$

This constraint implies that the wave functional is invariant in a change of space coordinates. Indeed, in an infinitesimal coordinate transformation, the variation of the metric tensor is (equation (28.10))

$$x^i \mapsto x^i + \varepsilon^i(x) \Rightarrow h_{ij} = g_{ik}\partial_j\varepsilon^k + g_{jk}\partial_i\varepsilon^k + \partial_k g_{ij}\varepsilon^k = \nabla_i\varepsilon_j + \nabla_j\varepsilon_i. \quad (28.111)$$

Thus, the invariance of the wave functional is expressed by

$$\nabla_b^{\text{sp.}}\frac{\delta\Psi}{\delta g_{ab}} = 0, \quad (28.112)$$

which is exactly the constraint. This constraint commutes with the Hamiltonian, since time-independent change of coordinates remains a symmetry in the temporal gauge.

The last constraint is related to time-independent time reparametrization, and thus has no equivalent in gauge theories:

$$R - 2R^{00} = R^{\text{sp.}} + \frac{1}{4}\partial_0 g^{ab}\partial_0 g_{ab} + \frac{1}{4}(g^{ab}\partial_0 g_{ab})^2 = 0. \quad (28.113)$$

A short calculation shows that

$$R - 2R^{00} = -\frac{1}{\sqrt{g}}\mathcal{H} \Rightarrow \mathcal{H}\Psi = 0. \quad (28.114)$$

Therefore, the constraint, known as *Wheeler–DeWitt equation*, implies that the physical states correspond to wave functionals that are eigenfunctions with eigenvalue 0 of the Hamiltonian. In the temporal gauge, there is no time evolution in the space of physical states. Dynamics is entirely encoded in the very definition of physical states.

28.7.2 Quantization: A few remarks

Beyond these algebraic considerations, quantization of gravity remains an open topic. The quantization method presented here has been criticized, because it relies on the possibility of defining a space-like surface, a notion ill-defined in the presence of strong metric fluctuations. As for gauge fields, it is indeed completely justified only in perturbation theory, when the background (or classical) metric is static, and typical deviations of the metric from the background (or classical) metric are small.

The definition of quantum gravity, beyond the formal level, leads to a number of unsolved problems, of which we list only a few.

While, in general, the simplest way to define a QFT is first to construct the Euclidean theory, and then proceed by analytic continuation to real time, in quantum gravity the Euclidean action is not bounded from below. Moreover, the connection between Euclidean and Minkowskian gravity is much less obvious than for non-gravitational theories because a change in the signature of the metric is involved.

Lattice regularization by simplicial gravity can be easily achieved in two dimensions, and is under investigation in four dimensions [310].

In the real-time formulation, perturbation theory is expanded around a fixed background metric η_{ij} which, in the case of asymptotic flat space, is simply the Minkowski metric $\{+1, -1, -1, -1, \dots\}$. How nature chooses this particular signature is a non-trivial interesting problem. In perturbation theory, pure quantum gravity then describes the self-interaction of a hypothetical spin-two massless particle, called the *graviton*. From the equations of motions, one infers that it has $\frac{1}{2}N(N+1)/2 - 2N = \frac{1}{2}N(N-3)$ dynamical degrees of freedom. In particular, for gravity, the dimension 3 is the analogous of dimension 2 for gauge theories, due to reparametrization invariance and constraints, the metric has no dynamical degrees of freedom.

Pure quantum gravity is non-renormalizable in four dimensions. The best we can expect in a covariant gauge is that the metric has canonical dimension $\frac{1}{2}(N-2)$. However, the action is non-polynomial. As in the case of the non-linear σ model, we thus expect pure gravity to be renormalizable only in dimension 2, but then the action is topological, or more generally in dimension $(2+\varepsilon)$. Attempts have been made to follow the methods successfully employed in the case of the non-linear σ model, and to look for a non-trivial ultraviolet (UV) fixed point in $(2+\varepsilon)$ dimension. However, the analysis of singularities of perturbation theory when the dimension approaches 2 is complicated, because the theory does not exist in two dimensions and, therefore, remains inconclusive (but the inclusion of a scalar field called dilaton seems to improve the situation).

In fact, it is widely believed that the theory, even in its supersymmetric form remains non-renormalizable, a property that could indicate the breakdown of local QFT close to Planck's scale. The gravity action can only be considered as the first term in a local expansion, in the sense of an *effective field theory*. The property that gravity is extremely weak at short distance is consistent with its non-renormalizability.

Finally, it is hard to figure out how direct effects due to quantum gravity could be observed. In particular, no experimental device has been proposed to detect gravitons, the massless spin 2 particles associated with quantum gravity. The most discussed issues in the past have been black hole evaporation due to semi-classical effects and the Beckenstein–Hawking entropy,

$$\Sigma = \frac{Ac^3}{4G_N\hbar},$$

where A is the event horizon area and G_N Newton's constant.

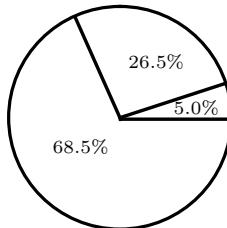


Fig. 28.3 Energy content of the Universe (2015): 68.5% dark energy, 26.5% dark matter, 5.0% ordinary matter

28.8 Observational cosmology: A few comments

In the last decades, a combination of satellite and ground observations had confirmed the existence of black holes (a prediction of GR) of different types, super-massive black holes at the centre of galaxies as well as black holes of several solar masses, of dark matter and dark energy. Moreover, starting from year 2015, classical GR has found new confirmations with the direct detection of gravitational waves (spin-two waves), by the LIGO–VIRGO collaboration (indirect evidence came earlier from the variation of the rate of emissions of a binary pulsar), the signals being consistent with the collapse of black holes and neutron stars [311]. Additional results have been, for example, evidence of about 30 to more than 100 solar mass black holes, and confirmation that gravitational waves propagate, as predicted, at the speed of light.

Deviations, at the 10^{-5} level, from a simple relic 2.7% K cosmic microwave background (CMB) by the WMAP (Wilkinson microwave anisotropy probe) and, later, the Planck satellites have been measured with increasing precision.

All these observations have strongly modified our understanding of cosmology and the evolution of the Universe, since the presumably Big Bang some 13.8 billion years ago.

28.8.1 An emerging model of the Universe and its evolution

A simple model, Λ CDM (Lambda Cold Dark Matter), can account for the measurement of the deviations, at the 10^{-5} level, from a simple 2.7% K CMB and number of other observations. It is then found that the Universe at large scale is flat. It is filled with 15% of normal matter and 85% of *dark matter*, a hypothetical form of matter, without electromagnetic interactions, needed, in particular, to explain the rotation curves of stars in galaxies and the cohesion of clusters of galaxies.

Alternative to the hypothesis of a new form of matter, models based on a modification of the GR theory at the larger scales have been proposed, like the so-called $f(R)$ gravity, when R is the scalar curvature. But no compelling model has emerged.

Moreover, it has been observed (supernovae 1A distribution, CMB) that, since about five billion years, the expansion of the Universe is accelerating while gravitational pull leads to a slowing down.

This effect, which corresponds to a negative pressure, when parametrized in terms of an energy content (7×10^{-30} g/cm³), dubbed *dark energy*, accounts for about 68% of the total energy content of the Universe (Fig. 28.3).

From the observational determination of the equation of state, one verifies that this dark energy is compatible with the effect of a *cosmological constant*, an effect that is expected to be generated by the vacuum energy density of quantum matter, since the vacuum energy becomes an observable in a GR background. However, this cosmological constant, expressed in the natural scale of Planck's mass, is exceedingly small.

It had been noticed that a supersymmetric theory has an exact vanishing cosmological constant, because the vacuum energy vanishes, but, since experimentally supersymmetry, even if it is a symmetry of nature, is broken at presently accessible energies, this hypothesis does not provide a solution to the problem: it reduces 120 to 60 orders of magnitude, again assuming that the effect has a scale provided by Planck's mass.

Formulated differently, the scale associated to the dark energy density does not seem to be related to Planck's mass (about 10^{19} GeV), but rather to a very small mass of the order of 10^{-3} eV (of the order of neutrino masses? or related to dark matter?).

29 Generalized non-linear σ -models in two dimensions

In this chapter, we describe the formal properties, and then discuss the renormalization, of a class of quantum field theories (QFT) based *on homogeneous spaces*, that is, coset spaces of the form G/H , where G is a compact Lie group and H a Lie subgroup. In physics, they appear naturally in the case of *spontaneous symmetry breaking*, and describe the interactions between *Goldstone modes* (see Section 13.4).

Homogeneous spaces are associated with non-linear realizations of group representations [312]. They can be studied in different parametrizations corresponding to different choices of coordinates when these spaces are considered as Riemannian manifolds. However, in contrast with arbitrary manifolds, there exist natural ways to embed these manifolds in flat Euclidean spaces, spaces in which the symmetry group acts linearly. Following a method explained in the example of the non-linear σ -model in Chapter 19, we first use such an embedding, because the renormalization properties are simpler, and the physical interpretation of correlation functions more direct. We then also examine some properties of the models in a generic parametrization. The renormalization problem is then solved by the introduction of a Becchi–Rouet–Stora–Tyutin (BRST)-like symmetry with anticommuting (Grassmann) parameters (see Chapter 26 for details).

Then, we study the more specific properties of models corresponding to a special class of homogeneous spaces: symmetric spaces. The non-linear σ -model (Chapter 19) provides a simple example. These models are characterized by the uniqueness of the metric and thus of the classical action, up to a multiplicative constant. In two dimensions, from the classical field equations an infinite number of non-local conservation laws can be derived.

The quantum models depend on only one coupling constant. We calculate the field and coupling renormalization group (RG) functions at one-loop order, in two dimensions, and find that the models all exhibit the property of ultraviolet (UV) asymptotic freedom (see also Chapter 23).

The chapter ends with comments about more general models based on non-compact groups and arbitrary Riemannian manifolds. The Appendix contains a few additional remarks about metric and curvature in homogeneous spaces. We also briefly describe classical families of symmetric spaces.

In the description of these models, two different formalisms and sets of notation can be employed, depending on whether one wants to emphasize the group structure, or the Riemannian manifold point of view.

29.1 Homogeneous spaces and Goldstone modes

The study of QFT based on homogeneous spaces can be motivated by some physics arguments. We consider an N -component field $\phi(x)$ transforming under a linear matrix representation $\mathcal{R}(G)$ of a compact Lie group G . We denote by \mathbf{t}^α are the generators of the Lie algebra $\mathcal{L}(G)$ corresponding to the representation $\mathcal{R}(G)$. We use the notation of Chapter 13 for the infinitesimal group transformation.

In an infinitesimal group transformation with parameters ω_α , the field transforms as,

$$\delta_\omega \phi(x) = \sum_\alpha \omega_\alpha \mathbf{t}^\alpha \phi(x). \quad (29.1)$$

Since we only consider compact groups, we can assume that the representation is orthogonal, and that the generators of $\mathcal{L}(G)$ are $N \times N$ antisymmetric matrices.

29.1.1 Spontaneous symmetry breaking, Goldstone modes, and homogeneous spaces

As in Chapter 13, our starting point is a G -invariant action $\mathcal{S}(\phi)$ of the form

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + V(\phi(x)) \right]. \quad (29.2)$$

We assume that, in the classical (or tree) approximation, the model exhibits the phenomenon of spontaneous symmetry breaking, that is, $V(\phi)$ has non- G -invariant, continuously connected, global minima. We distinguish one of them, ϕ^c , around which we expand perturbation theory, and call H the subgroup that leaves ϕ^c invariant (the isotropy or little group, or stabilizer of ϕ^c). We have argued in Section 13.4.2 that, in this situation, a number of field components of $\phi(x)$, which correspond to the generators of the Lie algebra of G that do not belong to the Lie algebra of H , are massless Goldstone modes.

If we are only interested in the long-distance behaviour (infrared (IR) limit) of correlation functions, or equivalently, if the mass of the massive fields is sent to infinity (the low-temperature limit of the corresponding statistical model), the fluctuations of the massive fields in the field integral can be neglected. In this limit, the field $\phi(x)$ is reduced to its massless components, and can be entirely parametrized in terms of a matrix $\mathbf{R}(g(x))$ belonging to the representation $\mathcal{R}(G)$, acting on the vector ϕ_c :

$$\phi(x) = \mathbf{R}(g(x))\phi_c, \quad g(x) \in G. \quad (29.3)$$

Note that if one multiplies $g(x)$ on the right by an element $h(x)$ of H , since ϕ_c is left invariant by the group H , $\phi(x)$ is not modified:

$$\phi(x) = \mathbf{R}(g(x)h(x))\phi_c = \mathbf{R}(g(x))\mathbf{R}(h(x))\phi_c = \mathbf{R}(g(x))\phi_c.$$

This shows that $\phi(x)$ is a function only of the elements of the coset (homogeneous) space G/H . We then divide the set of generators of the Lie algebra $\mathcal{L}(G)$ into the set of generators belonging to the Lie algebra $\mathcal{L}(H)$, $\{\mathbf{t}^\alpha\}$, $\alpha > \ell$, and the complementary set, which we denote by $\mathcal{L}(G/H)$ of generators $\{\mathbf{t}^\alpha\}$, $1 \leq \alpha \leq \ell$, which is such that

$$\sum_{\alpha=1}^{\ell} c_\alpha \mathbf{t}^\alpha \phi_c = 0 \Rightarrow c_\alpha = 0.$$

When acting on ϕ^c , the matrix \mathbf{R} can be canonically parametrized in terms of fields $\xi_a(x)$ as

$$\mathbf{R}(\xi_a(x)) = \exp \left(\sum_{a=1}^{\ell} \xi_a(x) \mathbf{t}^a \right). \quad (29.4)$$

29.1.2 Goldstone mode effective action

Since $V(\phi)$ is derivative-free and group-invariant, it is independent of $g(x)$ and thus yields an irrelevant constant contribution to the action, which can be omitted. In terms of the representation (29.3), the action $\mathcal{S}(\phi)$ then reduces to

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x \phi_c \nabla \mathbf{R}^{-1}(g(x)) \cdot \nabla \mathbf{R}(g(x)) \phi_c. \quad (29.5)$$

By expanding the action (29.5) for $g(x)$ close to the identity, it is simple to verify that all remaining fields are massless.

Notation. It can be convenient to use the bra–ket notation of quantum mechanics and to indicate vectors, denoting by $|0\rangle$ the vector ϕ^c . Equation (29.3) then takes the form

$$|\phi(x)\rangle = \mathbf{R}(\xi(x))|0\rangle, \quad (29.6)$$

in which $|\phi(x)\rangle$ is a notation for the field $\phi(x)$. With this notation, the classical action $\mathcal{S}(\phi)$ can be rewritten in terms of $\mathbf{R}^{-1}\nabla\mathbf{R}$ as

$$\mathcal{S}(\phi) = -\frac{1}{2} \int d^d x \langle 0 | (\mathbf{R}^{-1}(x) \nabla \mathbf{R}(x))^2 | 0 \rangle. \quad (29.7)$$

Any matrix of the form $\mathbf{R}^{-1}\partial\mathbf{R}$ (a pure gauge) belongs to the Lie algebra of G .

29.1.3 Metric and action in general coordinates

In this section, we use a notation and conventions adapted to Riemannian geometry (see Chapter 28 for details). In particular, we now use the convention of *implicit summation over successive upper and lower indices*.

We parametrize, locally, the manifold (coset space) G/H in terms of a set of coordinates φ^i . The action (29.5) can then be written as

$$\mathcal{S}(\varphi) = \frac{1}{2} \int d^d x g_{ij}(\varphi(x)) \nabla \varphi^i(x) \nabla \varphi^j(x), \quad (29.8)$$

where g_{ij} is a metric on G/H considered as a Riemannian manifold. In what follows ∂_i with Roman indices means derivative with respect to φ^i .

We have shown in Section 29.1.1 that the matrix \mathbf{R} , when acting on the vector $|0\rangle$, can be parametrized in the form (29.4). The variables ξ^a in equation (29.4) thus provide an example of a set of coordinates on the coset space G/H .

Comparing the expressions (29.8) and (29.7), we can relate the metric to other geometric objects. The current $\mathbf{R}^{-1}\nabla\mathbf{R}$ can be written as

$$\mathbf{R}^{-1}\nabla\mathbf{R} = \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi^i} \nabla \varphi^i(x). \quad (29.9)$$

We can expand $\mathbf{R}^{-1}\nabla\mathbf{R}$ on a basis of generators of the Lie algebra of G (more details can be found in Section A29.1):

$$\mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi^i} \equiv \mathbf{R}^{-1} \partial_i \mathbf{R} = L_{\alpha,i}(\varphi) \mathbf{t}^\alpha, \quad (29.10)$$

and, therefore,

$$\mathbf{R}^{-1}\nabla\mathbf{R} = L_{\alpha,i}(\varphi) \mathbf{t}^\alpha \nabla \varphi^i(x). \quad (29.11)$$

An expression of the metric tensor in terms of $L_{\alpha,i}(\varphi)$ follows:

$$g_{ij}(\varphi) = L_{a,i}(\varphi) L_{b,j}(\varphi) \mu^{ab}, \quad (29.12)$$

with the definition

$$\mu^{ab} = -\langle 0 | \mathbf{t}^a \mathbf{t}^b | 0 \rangle. \quad (29.13)$$

The Roman letters (a, b) indicate that the indices only run over values corresponding to generators belonging to $\mathcal{L}(G/H)$.

One can verify that the quantities $L_{\alpha,i}$ are independent of the representation (Section A29.1). Therefore, the dependence on the choices of the classical vector $|0\rangle$ and the particular representation of the group G is entirely contained in the positive matrix μ^{ab} . The form of the matrix is only restricted by the symmetry under the subgroup H :

$$\langle 0 | \mathbf{t}^a \boldsymbol{\tau}^\gamma \mathbf{t}^b | 0 \rangle = \langle 0 | \mathbf{t}^a [\boldsymbol{\tau}^\gamma, \mathbf{t}^b] | 0 \rangle = \langle 0 | [\mathbf{t}^a, \boldsymbol{\tau}^\gamma] \mathbf{t}^b | 0 \rangle,$$

for $\mathbf{t}^a, \mathbf{t}^b \in \mathcal{L}(G/H)$, $\boldsymbol{\tau}^\gamma \in \mathcal{L}(H)$ and, therefore,

$$\mu^{ac} f_c^{\gamma b} - f_c^{a\gamma} \mu^{cb} = 0, \quad (29.14)$$

where $f_\beta^{\alpha\gamma}$ are the structure constants of $\mathcal{L}(G)$. The interpretation of the equation is simple: the vectors $\mathbf{t}^a |0\rangle$ transform under a (in general reducible) representation of H , which has the matrices $f_b^{a\gamma}$ as generators.

The symmetric matrix $\boldsymbol{\mu}$ commutes with all generators \mathbf{f}^γ ($[\mathbf{f}^\gamma]_a^b \equiv f_b^{a\gamma}$) of $\mathcal{L}(H)$:

$$[\boldsymbol{\mu}, \mathbf{f}^\gamma] = 0.$$

As a consequence, the number of parameters on which μ^{ab} depends, is the number of different H -invariant scalar products one can form with the irreducible components of the vector $\mathbf{t}^a |0\rangle$ (Schur's lemma).

This result exhausts all consequences of the G -symmetry.

29.1.4 Quantization and perturbation theory near dimension 2

To quantize the classical theory, we could start from the Euclidean action (29.5), derive a quantum Hamiltonian, and use canonical quantization [193]. It is faster to begin with action (29.2) and freeze the massive degrees of freedom as explained previously in Section 29.1. The result is the same and yields the field integral representation

$$\mathcal{Z} = \int \left[g^{1/2}(\varphi) d(\varphi) \right] \exp \left[-\frac{1}{2} \int d^d x g_{ij}(\varphi(x)) \nabla \varphi^i(x) \nabla \varphi^j(x) \right], \quad (29.15)$$

where $g = \det g_{ij}$. The φ -integration measure is the covariant volume element in the manifold (see Section 28.3.1), the G -invariant measure induced by the flat ϕ measure, and is also the restriction to the coset space G/H of the Haar group measure of G .

From the point of view of power counting, the theory is clearly renormalizable in two dimensions for any parametrization, since the φ -field has dimension $\frac{1}{2}(d-2)$ and the interaction terms contain two derivatives and arbitrary powers of the φ -field.

To prove the structural stability of the action (29.5), we derive in the coming sections a set of Ward–Takahashi (WT) identities, consequences of the non-linearly realized G -symmetry. For the proof, a lattice regularization respecting the G -symmetry is required, because it regularizes the field-integration measure, and it is implicit in the derivation of WT identities. Dimensional regularization, which ignores the integration measure, but generates vertex functions that satisfy the WT identities, can then be used for practical calculations. The choice of a G -symmetric regularization corresponds to a choice of quantization of the classical Hamiltonian consistent with the symmetry.

29.2 WT identities and renormalization in linear coordinates

For renormalization problems, it is convenient to embed the homogeneous space into an Euclidean space, as we have explained in Chapter 19.

Remark. Note that the vector space \mathcal{V} spanned by the family of vectors of the form (29.3) may have a lower dimension than the original space of representation to which ϕ belongs, if this representation is reducible. The space \mathcal{V} is still a space of representation for the group G . In what follows we *restrict* ϕ to its components in \mathcal{V} .

29.2.1 Linear coordinates

We consider the vector space \mathcal{V} spanned by the vectors of form (29.3), as well as the subspace \mathcal{V}' of dimension ℓ spanned by the vectors $\mathbf{t}^\alpha \phi_c$. We choose an orthogonal basis in \mathcal{V} such that the vectors of \mathcal{V}' only have the first ℓ components non-vanishing. We then distinguish the ℓ first components of the vector ϕ of equation (29.3), denoting them by $\pi_a(x)$, and denoting the remaining ones by $\sigma_i(x)$:

$$\phi(x) = \begin{cases} \pi_a(x), & 1 \leq a \leq \ell \\ \sigma_i(x), & \ell < i \end{cases} . \quad (29.16)$$

The fields $\sigma_i(x)$ and $\pi_a(x)$ are functions of the fields $\xi_a(x)$. Expanding equation (29.4) in powers of ξ ,

$$\phi(x) = \phi_c + \sum_{a=1}^{\ell} \xi_a(x) \mathbf{t}^a \phi_c + O(\xi^2) , \quad (29.17)$$

we note, comparing expressions (29.3) and (29.16), that $\pi_a(x)$ and $\xi_a(x)$ are canonically related:

$$\pi(x) = \sum_{b=1}^{\ell} \xi_b(x) \mathbf{t}^b \phi_c + O(\xi^2) , \quad (29.18)$$

that is, the relation can be inverted to express the fields ξ_a as a power series in the fields π_a . On the other hand, the fields $\sigma_i(x)$ are of the form

$$\sigma(x) = \phi_c + O(\xi^2) ,$$

and, therefore, can be calculated in terms of the fields π_a . The fields $\pi_a(x)$ and $\sigma_i(x)$ transform under different linear representations of the group H . However, the fields $\pi_a(x)$ transform under a non-linear representation of the group G , since the generators $\{\mathbf{t}^\alpha\}$ mix the fields $\pi_a(x)$ and the $\sigma_i(x)$, which are functions of π_a . The infinitesimal transformation of parameters ω_α takes the form

$$\delta_\omega \pi_a(x) = [t_{ab}^\alpha \pi_b(x) + t_{aj}^\alpha \sigma_j(\pi(x))] \omega_\alpha . \quad (29.19)$$

Since the σ_i are functions of π_a , the transformation laws

$$\delta_\omega \sigma_i(\pi(x)) = [t_{ib}^\alpha \pi_b(x) + t_{ij}^\alpha \sigma_j(\pi(x))] \omega_\alpha , \quad (29.20)$$

are now consequences of equation (29.19) and the functional form of the $\sigma_i(\pi)$.

29.2.2 Correlation functions, WT identities

We consider the generating functional of correlation functions

$$\mathcal{Z}(J) = \int [g^{1/2}(\pi) d(\pi)] \exp \left[-\mathcal{S}(\phi) + \int d^d x \mathbf{J}(x) \cdot \phi(x) \right] , \quad (29.21)$$

the integrand being expressed in terms of these special coordinates. Sources have been added for all components of ϕ in \mathcal{V} for reasons that have already been explained in Chapter 19: when we try to derive WT identities expressing the consequences of the symmetry for correlation functions, the composite σ_a fields appear in the variation of the π -field. Moreover, since all π -fields are massless, we have to break the G -symmetry explicitly to provide an IR cut-off. This can be achieved by expanding perturbation theory starting from fixed constant non-vanishing values of the σ sources.

WT identities. We now derive the WT identities corresponding to the non-linearly realized symmetry under the group G , and show how they imply the structural stability of the action (29.5). We perform an infinitesimal transformation (29.19) in the field integral (29.21). Since we have introduced sources for all components of ϕ , the corresponding WT identity for $\mathcal{Z}(J)$ and $\mathcal{W}(J)$ is identical to the identity obtained for linearly realized symmetries:

$$\int d^d x J_i(x) t_{ij}^\alpha \frac{\delta \mathcal{Z}(J)}{\delta J_j(x)} = 0 \quad \Rightarrow \quad \int d^d x J_i(x) t_{ij}^\alpha \frac{\delta W}{\delta J_j(x)} = 0. \quad (29.22)$$

Again the difference appears in the Legendre transformation. We denote by $J_a(x)$, $1 \leq a \leq \ell$, the sources for the fields π_a , and H_i , $\ell < i \leq n$, the sources for the composite fields σ_i . The Legendre transform has only to be performed on J_a . The transformation takes the form,

$$\Gamma(\pi, H) + \mathcal{W}(J, H) = \int d^d x \sum_{a=1}^{\ell} \pi_a(x) J_a(x), \quad (29.23)$$

$$\pi_a(x) = \frac{\delta W}{\delta J_a(x)} \Leftrightarrow J_a(x) = \frac{\delta \Gamma}{\delta \pi_a(x)}. \quad (29.24)$$

Since, as a consequence of equation (7.65),

$$\frac{\delta \Gamma}{\delta H_i(x)} = - \frac{\delta \mathcal{W}}{\delta H_i(x)},$$

the identity (29.22) implies for the generating functional Γ of vertex functions,

$$\int d^d x \left\{ \frac{\delta \Gamma}{\delta \pi_a(x)} \left[t_{ab}^\alpha \pi_b(x) - t_{aj}^\alpha \frac{\delta \Gamma}{\delta H_j(x)} \right] + H_i(x) \left[t_{ib}^\alpha \pi_b(x) - t_{ij}^\alpha \frac{\delta \Gamma}{\delta H_j(x)} \right] \right\} = 0. \quad (29.25)$$

29.2.3 Renormalization

As in the example of the σ -model (equation (19.35)), the identity (29.25) has a quadratic form. The arguments of Section 19.6 apply and make it possible to prove the stability of the identity under renormalization. The same identity is thus fulfilled by the renormalized action \mathcal{S}_r :

$$\int d^d x \left\{ \frac{\delta \mathcal{S}_r}{\delta \pi_a(x)} \left[t_{ab}^\alpha \pi_b(x) - t_{aj}^\alpha \frac{\delta \mathcal{S}_r}{\delta H_j(x)} \right] + H_i(x) \left[t_{ib}^\alpha \pi_b(x) - t_{ij}^\alpha \frac{\delta \mathcal{S}_r}{\delta H_j(x)} \right] \right\} = 0. \quad (29.26)$$

Power counting implies that $\mathcal{S}_r(\pi, H)$ and $H_a(x)$ have dimension 2. Therefore, again \mathcal{S}_r is linear in the source $H_a(x)$. We write \mathcal{S}_r as

$$\mathcal{S}_r(\pi, H) = - \int \sigma_{r,i}(\pi) H_i(x) d^d x + \Sigma_r(\pi), \quad (29.27)$$

in which the functions $\sigma_{r,i}(\pi)$ are derivative-free and Σ_r has dimension 2, that is, has at most two derivatives. The term linear in H in equation (29.26) yields

$$\frac{\delta \sigma_{r,i}(\pi)}{\delta \pi_a(x)} \left[t_{ab}^\alpha \pi_b(x) + t_{aj}^\alpha \sigma_{r,j}(\pi(x)) \right] = t_{ib}^\alpha \pi_b(x) + t_{ij}^\alpha \sigma_{r,j}(\pi(x)). \quad (29.28)$$

These partial differential equations for the function $\sigma_{r,i}(\pi)$ imply that if the fields π_a transform as

$$\delta_\alpha \pi_a(x) = t_{ab}^\alpha \pi_b(x) + t_{aj}^\alpha \sigma_{r,j}(\pi(x)), \quad (29.29)$$

then, as a consequence,

$$\delta_\alpha \sigma_i(\pi(x)) = t_{ib}^\alpha \pi_b(x) + t_{ij}^\alpha \sigma_{r,j}(\pi(x)). \quad (29.30)$$

Thus, the field ϕ with component $(\pi_a, \sigma_{r,i})$ transforms under a linear representation of the group G . We now consider the equation obtained by setting $H = 0$:

$$\int d^d x \frac{\delta \Sigma_r}{\delta \pi_a(x)} [t_{ab}^\alpha \pi_b(x) + t_{aj}^\alpha \sigma_{r,j}(x)] = 0. \quad (29.31)$$

The equation implies that $\Sigma_r(\pi)$ is the most general functional of $\pi(x)$ with two derivatives invariant under the group G .

The action can thus be constructed starting from the most general G -invariant action with two derivatives expressed in terms of the field $\phi(x)$, and then eliminating the fields $\sigma_{ri}(x)$. Also, we have shown in Section 29.1.3 that the action, expressed in terms of the metric tensor, can be parametrized in terms of the matrix μ (equation (29.13)). Here, we find that, even if in the tree approximation we begin with a special choice of the matrix μ solution of equation (29.14), after renormalization we obtain the most general solution of this equation.

29.2.4 Field renormalizations

Here, we briefly comment about the solutions of equation (29.28). For renormalization purpose, we are looking for *generic* solutions $\sigma_i(\pi)$ expandable in powers of π :

$$\sigma_i = S_i + S_i^{a_1} \pi_{a_1} + \frac{1}{2} S_i^{a_1 a_2} \pi_{a_1} \pi_{a_2} + \dots, \quad (29.32)$$

in which $S_i, S_i^{a_1} \dots$ are constants.

The equation of order 0 in π is

$$S_i^a t_{aj}^\alpha S_j = t_{ij}^\alpha S_j. \quad (29.33)$$

If t^α belongs to the Lie algebra $\mathcal{L}(H)$, t_{aj}^α vanishes, since π and σ belong to different representations and, therefore,

$$t_{ij}^\alpha S_j = 0 \quad \forall t^\alpha \in \mathcal{L}(H). \quad (29.34)$$

As a consequence, the generic vector S_j is the most general vector having H as an isotropy group. Note that, since S_i differs from ϕ_c only at one-loop order, the isotropy group (stabilizer) of S_i cannot be larger than H .

For the same reason, the generators of $\mathcal{L}(G/H)$ are such that $t_{aj}^\alpha S_j$ spans the π subspace. Therefore, equation (29.33) implies

$$S_i^a t_{aj}^\alpha S_j = t_{ij}^\alpha S_j, \quad 1 \leq \alpha \leq \ell. \quad (29.35)$$

The scalar product of the vectors S_i with all vectors of a complete basis is known, therefore, all S_i are determined. In particular, if the π subspace contains no vector invariant under the action of the group H , the right-hand side of equation (29.35) and, therefore, also all vectors S_i vanish. This property holds for symmetric spaces.

Higher-order equations determine the coefficients of monomials of increasing degree in π . If we assume that we know $\sigma_i(\pi)$ at order π^k , then the equation for the coefficient of π^{k+1} takes the form

$$S_i^{ab_1 \dots b_k} \dots t_{aj}^\alpha S_j = \text{known quantities}.$$

The coefficients $S_i^{ab_1 \dots b_k}$ are completely determined, since, considered as vectors $\mathbf{S}_i^{b_1 \dots b_k}$, their scalar products with all vectors of a complete basis are given.

The conclusion is that the functions $\sigma_i(\pi)$ depend on as many parameters as the number of independent vectors S_i which have H as an isotropy group. The renormalized action is then the most general ‘free’ massless action in the linear ϕ variables. Therefore, we have enumerated all the renormalization constants of the model.

29.3 Renormalization in general coordinates: BRST symmetry

In Section 29.2, we have shown how a special choice of parametrization leads to a simple discussion of the symmetry properties and renormalization of correlation functions.

In Chapter 19, in the case of the non-linear σ -model, we have indicated that another choice would have required the introduction of a number, generically infinite, of additional renormalization constants, corresponding to a renormalization of the parametrization of the manifold. The field is no longer multiplicatively renormalized, the bare field becoming a general function of the renormalized field.

However, as we have indicated in Section A7.2, in such situations, only parametrization-independent quantities (related to geometric properties of the manifold or to the S -matrix) are physical. Therefore, it is useful to also investigate models on homogeneous spaces in an arbitrary parametrization to more clearly exhibit the parametrization dependence. Moreover, the example of the calculation of the one-loop β -function will show that some parametrizations are more convenient for practical calculations.

In this section, we derive WT identities expressing the group symmetry in an arbitrary parametrization, and show that they are stable under renormalization. The general strategy, that is, to add to the action sources for a set of composite operators that is closed under infinitesimal group transformations on the field is only suitable if the minimum number of different operators is finite. Then, the renormalization of the theory follows from a rather straightforward generalization of the arguments given in the first part of this chapter. However, for a generic parametrization an infinite number of composite fields is required, and this strategy is no longer useful. Therefore, we introduce a new method which, in some generalized form is also relevant to the renormalization of gauge theories, and which is based on infinitesimal group transformations with anticommuting parameters (see Chapter 26).

29.3.1 Infinitesimal group transformations

We consider a non-linear realization of the representation of a group G acting on a field $\varphi^i(x)$. We express the infinitesimal group transformations corresponding to parameters ω^α in the form

$$\delta_\omega \varphi^i(x) = D_\alpha^i(\varphi(x)) \omega^\alpha. \quad (29.36)$$

We assume that the functions $D_\alpha^i(\varphi)$ are smooth, that is, infinitely differentiable, at $\varphi = 0$. We write the action in the form (29.8) as

$$\mathcal{S}(\varphi) = \frac{1}{2} \int d^d x g_{ij}(\varphi(x)) \nabla \varphi^i(x) \nabla \varphi^j(x). \quad (29.37)$$

The invariance of the action $\mathcal{S}(\varphi)$ under the transformations (29.36) implies

$$\int d^d x D_\alpha^i(\varphi(x)) \frac{\delta \mathcal{S}(\varphi)}{\delta \varphi^i(x)} = 0, \quad (29.38)$$

and, therefore, for the metric tensor $g_{ij}(\varphi)$,

$$D_\alpha^i \frac{\partial g_{jk}}{\partial \varphi^i} + g_{ik} \frac{\partial D_\alpha^i}{\partial \varphi^j} + g_{ij} \frac{\partial D_\alpha^i}{\partial \varphi^k} = 0. \quad (29.39)$$

Solving the equation is equivalent to finding all possible metric tensors on a given homogeneous space, compatible with the group structure. Section A29.2 contains some details about the nature and structure of the equation.

The functions $D_\alpha^i(\varphi)$ satisfy identities that can be obtained either by direct calculation, or by noting that the differential operators

$$\mathcal{D}_\alpha = D_\alpha^i(\varphi) \frac{\partial}{\partial \varphi^i}, \quad (29.40)$$

acting on functions of φ , form themselves a representation of the Lie algebra (see also Section 13.1). Therefore, they have commutation relations of the form

$$[\mathcal{D}_\alpha, \mathcal{D}_\beta] = f_{\alpha\beta}^\gamma \mathcal{D}_\gamma, \quad (29.41)$$

where $f_{\alpha\beta}^\gamma$ are the structure constants of the Lie algebra. Note that these commutation relations are also compatibility conditions for equation (29.38), considered as a set of linear differential equations for $\mathcal{S}(\varphi)$. Calculating explicitly the commutator in terms of the functions $D_\alpha^i(\varphi)$, one finds,

$$D_\alpha^j(\varphi) \frac{\partial D_\beta^i(\varphi)}{\partial \varphi^j} - D_\beta^j(\varphi) \frac{\partial D_\alpha^i(\varphi)}{\partial \varphi^j} = f_{\alpha\beta}^\gamma D_\gamma^i(\varphi). \quad (29.42)$$

This is the form of the commutation relations of the Lie algebra that is useful in the discussion of non-linear representations. Moreover, in Sections 26.3.2 (equation (26.39)) and 26.6 (equation (26.78)), we encounter equations that are formally identical, and play, in particular, an essential role in the discussion of the renormalization of gauge theories.

29.3.2 WT identities

We assume that we have introduced a group-invariant regularization, and consider the generating functional

$$\mathcal{Z}(J, K) = \int [d\rho(\varphi)] \exp \left\{ -\mathcal{S}(\varphi) + \int d^d x [K(x)A(\varphi(x)) + J_i(x)\varphi^i(x)] \right\}. \quad (29.43)$$

The measure $[d\rho(\varphi)]$ is the group-invariant measure. We have added sources for the field $\varphi^i(x)$ and for a local derivative-free function $A(\varphi(x))$ of $\varphi(x)$ to the action. The function $A(\varphi)$ has to only satisfy one condition: it has to begin with a term of order φ^2 such that, for $K(x)$ constant, masses are generated for all components of φ , which can serve as IR regulators.

To derive consequences of the non-linear symmetry (29.38), we introduce a set of anticommuting constants \bar{C}^α , and ε , elements of a *Grassmann algebra*, and perform the shift $\varphi \mapsto \varphi + \delta_{\text{BRST}}\varphi$, with

$$\delta_{\text{BRST}}\varphi^i(x) = \varepsilon D_\alpha^i(\varphi(x))\bar{C}^\alpha. \quad (29.44)$$

Together, with the transformation

$$\delta_{\text{BRST}}\bar{C}^\alpha = -\frac{1}{2}\varepsilon f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma, \quad (29.45)$$

they provide an example of BRST transformations, whose geometric origin is explained in Section 26.3.1, where we discuss *BRST symmetry* for group manifolds (see also Section 26.7.1). We recall that BRST transformations are such that $\delta_{\text{BRST}}^2 = 0$. In particular, in the combined transformations (29.44) and (29.45), one finds

$$\delta_{\text{BRST}}[D_\alpha^i(\varphi)\bar{C}^\alpha] = 0.$$

In expression (29.43), $\mathcal{S}(\varphi)$ and $[\text{d}\rho(\varphi)]$ are invariant. For the source terms,

$$\begin{aligned} \delta_{\text{BRST}} \int d^d x [J_i(x)\varphi^i(x) + K(x)A(\varphi(x))] \\ = \varepsilon \int d^d x \left[J_i(x) + K(x) \frac{\delta A}{\delta \varphi^i(x)} \right] D_\alpha^i(\varphi(x))\bar{C}^\alpha. \end{aligned} \quad (29.46)$$

Because the transformation (29.44) is not linear, it generates two composite operators: $D_\alpha^i(\varphi(x))\bar{C}^\alpha$, and $(\delta A/\delta \varphi^i)D_\alpha^i(\varphi)\bar{C}^\alpha$. In accordance with our general strategy, we introduce two anticommuting sources, $\Lambda_i(x)$ and $L(x)$, for them. Note that if we assign a fermion number +1 to \bar{C}^α and -1 to $\Lambda_i(x)$ and $L(x)$, this fermion number is conserved.

Since the operator $(\delta A/\delta \varphi^i)D_\alpha^i(\varphi)\bar{C}^\alpha$ is a BRST variation (BRST exact), it is also BRST invariant (closed).

We now consider the generating functional,

$$\begin{aligned} \mathcal{Z}(J, K, \bar{C}, \Lambda, L) = \int [\text{d}\rho(\varphi)] \exp \left[-\mathcal{S}(\varphi) + \int d^d x (J_i(x)\varphi^i(x) + K(x)A(\varphi(x))) \right. \\ \left. + \int d^d x \left(\Lambda_i(x) + L(x) \frac{\delta A}{\delta \varphi^i(x)} \right) D_\alpha^i(\varphi(x))\bar{C}^\alpha \right]. \end{aligned} \quad (29.47)$$

According to the preceding analysis, its variation has the form

$$\begin{aligned} \mathcal{Z}(J, K, \bar{C} + \delta_{\text{BRST}}\bar{C}, \Lambda, L) - \mathcal{Z}(J, K, \bar{C}, \Lambda, L) \\ = \varepsilon \int d^d x \left(J_i(x) \frac{\delta}{\delta \Lambda_i(x)} + K(x) \frac{\delta}{\delta L(x)} \right) \mathcal{Z}(J, K, \bar{C}, \Lambda, L), \end{aligned} \quad (29.48)$$

and, therefore,

$$\left[\int d^d x \left(J_i(x) \frac{\delta}{\delta \Lambda_i(x)} + K(x) \frac{\delta}{\delta L(x)} \right) + \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma \frac{\partial}{\partial \bar{C}^\alpha} \right] \mathcal{Z} = 0. \quad (29.49)$$

A similar identity for $\mathcal{W} = \ln \mathcal{Z}$ follows. After Legendre transformation with respect to J_i , one obtains the WT identity

$$\int d^d x \left(\frac{\delta \Gamma}{\delta \varphi^i(x)} \frac{\delta \Gamma}{\delta \Lambda_i(x)} + K(x) \frac{\delta \Gamma}{\delta L(x)} \right) + \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma \frac{\partial \Gamma}{\partial \bar{C}^\alpha} = 0. \quad (29.50)$$

29.3.3 The renormalized action

As explained in Sections 19.6 and 26.9, an equation of the form (29.50) is stable under renormalization, and is also satisfied by the renormalized action \mathcal{S}_r :

$$\int d^d x \left(\frac{\delta \mathcal{S}_r}{\delta \varphi^i(x)} \frac{\delta \mathcal{S}_r}{\delta \Lambda_i(x)} + K(x) \frac{\delta \mathcal{S}_r}{\delta L(x)} \right) + \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma \frac{\partial \mathcal{S}_r}{\partial \bar{C}^\alpha} = 0. \quad (29.51)$$

To solve the equation, we note that, as a consequence of fermion number conservation, Λ , L and \bar{C} can only appear in the combinations $(\Lambda \bar{C})$ and $(L \bar{C})$. In two dimensions, the sources K , $(\Lambda \bar{C})$ and $(L \bar{C})$ have dimension 2, and, therefore, \mathcal{S}_r is a linear function of these sources with derivative-free coefficients:

$$\begin{aligned} \mathcal{S}_r(\varphi, K, \bar{C}, \Lambda, L) &= \mathcal{S}_r(\varphi) \\ &- \int d^d x [K(x) A_r(\varphi(x)) + \Lambda_i(x) D_{r,\alpha}^i(\varphi(x)) \bar{C}^\alpha + L(x) M_\alpha(\varphi(x)) \bar{C}^\alpha]. \end{aligned} \quad (29.52)$$

Identifying the coefficient of $\Lambda_i(x)$ in equation (29.51), one obtains,

$$\frac{\partial D_{r,\alpha}^i(\varphi)}{\partial \varphi^j} D_{r,\beta}^j(\varphi) \bar{C}^\alpha \bar{C}^\beta - \frac{1}{2} f_{\beta\gamma}^\alpha \bar{C}^\beta \bar{C}^\gamma D_{r,\alpha}^i(\varphi) = 0. \quad (29.53)$$

The antisymmetric coefficient of $\bar{C}^\alpha \bar{C}^\beta$ must vanish and, thus,

$$\frac{\partial D_{r,\alpha}^i}{\partial \varphi^j} D_{r,\beta}^j - \frac{\partial D_{r,\beta}^i}{\partial \varphi^j} D_{r,\alpha}^j = f_{\alpha\beta}^\gamma D_{r,\gamma}^i. \quad (29.54)$$

The functions $D_{r,\alpha}^i(\varphi)$ are associated with a non-linear representation of the group G .

We then identify the coefficient of K ,

$$\frac{\delta A_r}{\delta \varphi^i} D_{r,\alpha}^i(\varphi) \bar{C}^\alpha - M_\alpha(\varphi) \bar{C}^\alpha = 0 \Rightarrow M_\alpha(\varphi) = \frac{\delta A_r}{\delta \varphi^i} D_{r,\alpha}^i(\varphi), \quad (29.55)$$

and the coefficient of L :

$$\frac{\delta M_\alpha}{\delta \varphi^i} D_{r,\beta}^i - \frac{\delta M_\beta}{\delta \varphi^i} D_{r,\alpha}^i = f_{\alpha\beta}^\gamma M_\gamma. \quad (29.56)$$

The latter equation is already implied by the two equations (29.54) and (29.55). We conclude that $A_r(\varphi)$ is, in general, an arbitrary function of φ .

Finally, the last equation, independent of the different sources, is

$$\frac{\delta \mathcal{S}_r}{\delta \varphi^i}(\varphi) D_{r,\alpha}^i(\varphi) = 0. \quad (29.57)$$

The renormalized action is invariant under the non-linear transformations of the group G generated by $D_{r,\alpha}^i(\varphi)$.

We do not need to discuss again thoroughly the solutions of the WT identities, which we have reduced to equations (29.54) and (29.57). The latter equation implies a renormalized form of equation (29.39), which is an equation for the renormalized metric tensor.

Remarks. In general, one chooses a H -symmetric parametrization for homogeneous spaces G/H . This imposes additional restrictions upon the renormalized form of $D_{r,\alpha}^i(\varphi)$, and implies that $\mathcal{S}_r(\varphi)$ is H symmetric.

The results obtained by the method of this section are less detailed than those obtained in the case of the linear parametrization in Section 29.2. The latter method should be used when applicable.

29.4 Symmetric spaces: Definition

Symmetric spaces are special homogeneous spaces such that the symmetry group G possesses an involutive automorphism, and the subgroup H is the subgroup of invariant elements. Considering the case in which G is compact, we show in Section A29.4.2 that H is then a maximal proper subgroup. Equivalently, a parity can be assigned to the generators of the Lie algebra, and the Lie algebra $\mathcal{L}(H)$ is the algebra of even elements. More details can be found in Section A29.4.

Field theory models in two dimensions, constructed on symmetric spaces, have special properties both on the classical level and after quantization, which we will examine. The non-linear σ -model in Chapter 19 provides one of the simplest examples. In particular, once the parametrization of the manifold is chosen, the Euclidean action is unique up to a constant multiplicative factor. This reflects the uniqueness of the metric in the manifold compatible with the group structure. The parity of generators of the Lie algebra leads to a parity assignment for the fields, +1 for the composite σ -field, -1 for the π -field.

The coset space G/H , when it is symmetric, can be constructed from a group manifold in the following way: we consider the elements \mathbf{g} of a group \mathfrak{G} of the form (see Section A29.4 for details)

$$\mathbf{g} = \mathbf{g}_2^{-1} \mathbf{g}_0 \mathbf{g}_1,$$

where $\mathbf{g}_0, \mathbf{g}_1, \mathbf{g}_2$ are elements of \mathfrak{G} , which satisfy conditions that we describe. We can then distinguish two cases:

(i) $\mathbf{g}_0 \equiv \mathbf{1}$ and $\mathbf{g}_1, \mathbf{g}_2$ are two independent elements of \mathfrak{G} . The automorphism is $\mathbf{g}_1 \leftrightarrow \mathbf{g}_2$. We recognize the coset space $\mathfrak{G} \times \mathfrak{G}/\mathfrak{G}$, the automorphism exchanging the two components of $\mathfrak{G} \times \mathfrak{G}$. As manifolds, they are identical to the manifold of the group \mathfrak{G} . The corresponding field theories are called *principal chiral models*. In the example of $G \equiv SU(2)$, they are related to the chiral models studied in Section 13.6.

(ii) The element \mathbf{g}_0 is a fixed element different from the identity, which satisfies

$$\mathbf{g}_0^* \mathbf{g}_0 = \epsilon \mathbf{1}, \quad \epsilon = \pm 1,$$

in which the $*$ operation is an involutive automorphism of the group \mathfrak{G} , which may be trivial (for unitary groups, it can also be the complex conjugation). The elements \mathbf{g}_1 and \mathbf{g}_2 are related by

$$\mathbf{g}_2 = (\mathbf{g}_1^{-1})^*,$$

which implies that the elements \mathbf{g} have the form

$$\mathbf{g} = (\mathbf{f}^{-1})^* \mathbf{g}_0 \mathbf{f}, \tag{29.58}$$

where the elements \mathbf{f} are arbitrary elements of the group \mathfrak{G} . As a consequence,

$$\mathbf{g}^* \mathbf{g} = \epsilon \mathbf{1}. \tag{29.59}$$

Then $G \equiv \mathfrak{G}$, the automorphism of G is $\mathbf{g} \mapsto \mathbf{g}_0^{-1} \mathbf{g}^* \mathbf{g}_0$, and H is the subgroup of elements \mathbf{h} that satisfy $\mathbf{h} = \mathbf{g}_0^{-1} \mathbf{h}^* \mathbf{g}_0$.

We will now extend, by continuity, the $*$ operation to the Lie algebra.

29.5 Classical field equations. Conservation laws

We show that two-dimensional models based on symmetric spaces exhibit an infinite number of conservation laws [313].

The action. In the case of symmetric spaces, the classical action \mathcal{S} can always be written in a simple geometric form, since symmetric spaces can be realized in the group manifold itself:

$$\mathcal{S}(\mathbf{g}) = \frac{1}{2} \int d^d x \operatorname{tr} [\nabla \mathbf{g}(x) \nabla \mathbf{g}^{-1}(x)], \quad (29.60)$$

in which $\mathbf{g}(x)$ belongs to some matrix representation of \mathfrak{G} . The different symmetric spaces are characterized by the group \mathfrak{G} , and the constraints imposed on $\mathbf{g}(x)$.

The action can be expressed in terms of the associated current \mathbf{A}_μ , which belongs to the Lie algebra $\mathcal{L}(\mathfrak{G})$:

$$\mathbf{A}_\mu(x) = \mathbf{g}^{-1}(x) \partial_\mu \mathbf{g}(x). \quad (29.61)$$

The field \mathbf{A}_μ can be considered as a connection, or gauge field (these concepts are discussed in Chapters 21, 22, and 28), of a special kind, called a *pure gauge*. As a short calculation shows, it is characterized by the vanishing of the corresponding curvature tensor $\mathbf{F}_{\mu\nu}$:

$$\mathbf{F}_{\mu\nu}(x) \equiv \partial_\mu \mathbf{A}_\nu(x) - \partial_\nu \mathbf{A}_\mu(x) + [\mathbf{A}_\mu(x), \mathbf{A}_\nu(x)] = 0. \quad (29.62)$$

The action can then be written as

$$\mathcal{S}(\mathbf{g}) = -\frac{1}{2} \int d^d x \operatorname{tr} \mathbf{A}_\mu^2(x). \quad (29.63)$$

Classical field equations. A classical field equation corresponds to the action (29.60). In all cases, a general variation of $\mathbf{g}(x)$ can be written as

$$\mathbf{g}(x) \mapsto (1 - \varepsilon^*(x)) \mathbf{g}(x) (1 + \varepsilon(x)), \quad (29.64)$$

in which $\varepsilon(x)$ and $\varepsilon^*(x)$ belong to $\mathcal{L}(\mathfrak{G})$, and are, either independent in the case of chiral models, or otherwise, related by the $*$ automorphism in order to preserve the condition (29.59):

$$\mathbf{g}^* \mathbf{g} = \epsilon \mathbf{1}. \quad (29.65)$$

The corresponding variation of \mathbf{A}_μ is

$$\delta \mathbf{A}_\mu(x) = \mathbf{D}_\mu (\varepsilon(x) - \mathbf{g}^{-1}(x) \varepsilon^*(x) \mathbf{g}(x)),$$

where \mathbf{D}_μ is the *covariant derivative* associated with the current \mathbf{A}_μ . It acts on an element $\omega(x)$ of the Lie algebra as

$$\mathbf{D}_\mu \omega(x) = \partial_\mu \omega(x) + [\mathbf{A}_\mu(x), \omega(x)]. \quad (29.66)$$

The variation of the action then can be written as

$$\delta \mathcal{S} = \int d^d x \operatorname{tr} [\partial_\mu \mathbf{A}_\mu(x) (\varepsilon(x) - \mathbf{g}^{-1}(x) \varepsilon^*(x) \mathbf{g}(x))].$$

In the case of chiral models, we can specialize the equation to $\varepsilon^* = 0$, for example. In the other cases, the relation (29.65) implies

$$\partial_\mu \mathbf{A}_\mu^*(x) = -\mathbf{g}(x) \partial_\mu \mathbf{A}_\mu(x) \mathbf{g}^{-1}(x).$$

In both cases, the classical field equation can be written as

$$\partial_\mu \mathbf{A}_\mu(x) = 0, \quad (29.67)$$

which expresses the conservation of the current associated to the \mathfrak{G} -symmetry.

29.5.1 Non-local conserved currents

In two dimensions, equations (29.61) and (29.67) imply the existence of an infinite number of non-local conserved currents.

We define a covariant derivative \mathbf{D}_μ by

$$\mathbf{D}_\mu = \partial_\mu \mathbf{1} + \mathbf{A}_\mu. \quad (29.68)$$

The explicit form of the covariant derivative is different from the expression (29.66) because, as we have also pointed out in Section 28.2, its form depends on the representation.

We consider the linear partial differential equations for matrices χ :

$$\mathbf{D}_\mu \chi(x, \kappa) = \kappa \epsilon_{\mu\nu} \partial_\nu \chi(x, \kappa), \quad (29.69)$$

in which κ is the spectral parameter and $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$, $\epsilon_{12} = 1$.

Let us show that the linear system (29.69) is compatible. Setting

$$\Delta_\mu = \mathbf{D}_\mu - \mathbf{1} \kappa \epsilon_{\mu\rho} \partial_\rho, \quad (29.70)$$

we have to calculate the commutator

$$[\Delta_\mu, \Delta_\nu] = [\mathbf{D}_\mu, \mathbf{D}_\nu] - \kappa (\epsilon_{\mu\rho} [\mathbf{1}\partial_\rho, \mathbf{D}_\nu] + \epsilon_{\nu\sigma} [\mathbf{D}_\mu, \mathbf{1}\partial_\sigma]). \quad (29.71)$$

In two dimensions, since $\mu \neq \nu$, in the last term of the right-hand side only $\rho = \nu$ and $\sigma = \mu$ give a non-vanishing contribution:

$$[\Delta_\mu, \Delta_\nu] = [\mathbf{D}_\mu, \mathbf{D}_\nu] - \kappa \epsilon_{\mu\nu} [\mathbf{1}\partial_\rho, \mathbf{D}_\rho]. \quad (29.72)$$

The field equation (29.67) implies

$$[\mathbf{1}\partial_\mu, \mathbf{D}_\mu] = \partial_\mu \mathbf{A}_\mu = 0. \quad (29.73)$$

The commutator $[\mathbf{D}_\mu, \mathbf{D}_\nu]$ is the curvature $\mathbf{F}_{\mu\nu}$ of equation (29.62), which vanishes,

$$[\mathbf{D}_\mu, \mathbf{D}_\nu] = \mathbf{F}_{\mu\nu} = 0. \quad (29.74)$$

The conclusion is

$$[\Delta_\mu, \Delta_\nu] = 0, \quad (29.75)$$

and the linear system (29.69) is compatible.

We now define the current

$$\mathbf{J}_\mu(x, \kappa) = \mathbf{D}_\mu \chi(x, \kappa). \quad (29.76)$$

As a consequence of equation (29.69), the current is conserved:

$$\partial_\mu \mathbf{J}_\mu(x, \kappa) = 0. \quad (29.77)$$

The solution of equation (29.69) has an expansion in powers of κ of the form

$$\chi(x, \kappa) = \mathbf{1} + \sum_{n=1}^{\infty} \chi_n(x) \kappa^{-n}. \quad (29.78)$$

Correspondingly, \mathbf{J}_μ has an expansion of form

$$\mathbf{J}_\mu(x, \kappa) = \mathbf{A}_\mu(x) + \sum_{n=1}^{\infty} \mathbf{J}_{\mu,n}(x) \kappa^{-n}. \quad (29.79)$$

The conservation equation (29.77) then generates an infinite number of conserved currents:

$$\partial_\mu \mathbf{J}_{\mu,n}(x) = 0.$$

The interesting question, that we do not investigate here, is then whether these conservation laws survive quantization. Let us simply mention that the corresponding quantum conservation laws lead to the factorization of the S -matrix, which can then often be completely determined. We refer the interested reader to the literature for details [314].

29.6 QFT: Perturbative expansion and RG

We define the QFT by the integral representation of the generating functional

$$\mathcal{Z}(\mathbf{j}) = \int [d\mathbf{f}(x)] \exp \left\{ -\frac{1}{h} \left[\mathcal{S}(\mathbf{g}) - \int d^d x \operatorname{tr} \mathbf{g}(x) \mathbf{j}(x) \right] \right\}, \quad (29.80)$$

in which $\mathbf{f}(x)$ is a group element, either identical to $\mathbf{g}(x)$ for $\mathfrak{G} \times \mathfrak{G}/\mathfrak{G}$, or such that $\mathbf{g}(x)$ is related to $\mathbf{f}(x)$ by equation (29.58):

$$\mathbf{g} = (\mathbf{f}^{-1})^* \mathbf{g}_0 \mathbf{f},$$

and $d\mathbf{f}$ is the de Haar measure for the group \mathfrak{G} . The parameter h is the coupling constant.

We can then parametrize $\mathbf{f}(x)$ in a form analogous to (29.4), in terms of independent field variables $\xi_a(x)$, coefficients of the generators belonging to $\mathcal{L}(G/H)$. We expand perturbation theory around a constant finite value of the source $\mathbf{j}(x)$, to generate a perturbative expansion with an IR cut-off.

Actually, for explicit calculations, it is convenient to return to the field representations (29.6, 29.4), in terms of bra–ket quantum notation,

$$|\phi(x)\rangle = \mathbf{R}(x)|0\rangle \equiv \exp \left[\sum_a t^a \xi_a(x) \right] |0\rangle, \quad (29.81)$$

valid for general homogeneous spaces.

The generating functional of correlation functions can be written as (equation (29.7)),

$$\mathcal{Z}(J) = \int [d\rho(\xi)] \exp \left\{ -\frac{1}{h} \int d^d x \left[\frac{1}{2} \langle \nabla \phi(x) | \nabla \phi(x) \rangle - \langle J(x) | \phi(x) \rangle \right] \right\}, \quad (29.82)$$

where $[d\rho(\xi)]$ is the invariant measure and

$$\langle \nabla \phi(x) | \nabla \phi(x) \rangle = \langle 0 | (\mathbf{R}^{-1}(x) \nabla \mathbf{R}(x))^2 | 0 \rangle. \quad (29.83)$$

To generate the perturbation expansion, we expand the action in powers of $\xi_a(x)$. The geometric part of the Feynman diagram calculation then involves the evaluation of expectation values of the form $\langle 0 | t^{a_1} t^{a_2} \dots t^{a_k} | 0 \rangle$.

29.6.1 RG functions at one-loop order

Notation. In this section, repeated indices mean summation over indices running from 1 to ℓ , values which correspond to the generators of G/H .

To be able to discuss the RG properties of these models, we calculate the renormalization constants and the RG functions at one-loop order. Moreover, these calculations illustrate some of the preceding considerations [315].

Preliminary remarks. We normalize the vector $|0\rangle$ by

$$\langle 0 | 0 \rangle = 1. \quad (29.84)$$

To evaluate the one-loop diagrams, we also need a few tensors like

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b | 0 \rangle = -\delta_{ab}. \quad (29.85)$$

Up to the normalization, the relation follows from the property that, in the case of symmetric spaces, the vectors $\mathbf{t}^a |0\rangle$ form an irreducible representation of the subgroup H .

Because the structure constants are antisymmetric, and the vector $|0\rangle$ is the unique vector having H for stabilizer group, then

$$\mathbf{t}^a \mathbf{t}^a |0\rangle = -\mu |0\rangle. \quad (29.86)$$

Combining equations (29.84–29.86), we obtain the value of μ :

$$\mathbf{t}^a \mathbf{t}^a |0\rangle = -\ell |0\rangle. \quad (29.87)$$

We need two tensors with three indices. Since the generators are represented by anti-symmetric matrices,

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^c | 0 \rangle + \langle 0 | \mathbf{t}^c \mathbf{t}^b \mathbf{t}^a | 0 \rangle = 0.$$

A few commutations and the properties of symmetric spaces lead to

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^c | 0 \rangle = 0. \quad (29.88)$$

A different useful quantity is obtained by replacing \mathbf{t}^b by a generator $\boldsymbol{\tau}^\beta$ of H :

$$\langle 0 | \mathbf{t}^a \boldsymbol{\tau}^\beta \mathbf{t}^c | 0 \rangle = -f_{a\beta c}. \quad (29.89)$$

We still need three tensors with four indices. A combination of some commutation relations and previous relations, yields

$$\langle 0 | \mathbf{t}^a \mathbf{t}^a \mathbf{t}^b \mathbf{t}^b | 0 \rangle = \ell^2, \quad (29.90)$$

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^a \mathbf{t}^b | 0 \rangle = \ell^2 - \sum_{\gamma > l} f_{\gamma ab} f_{\gamma ab}, \quad (29.91)$$

$$\text{and } \langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^b \mathbf{t}^a | 0 \rangle = \langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^a \mathbf{t}^b | 0 \rangle.$$

Using the antisymmetry of the structure constants and the special properties of symmetric spaces, it is simple to verify that

$$\sum_{\alpha > l} \sum_{a=1}^l f_{\alpha ab} f_{\alpha ab'} = \frac{1}{2} \delta_{bb'} C, \quad (29.92)$$

in which C is the Casimir of the group G . Summing over $b = b'$, we obtain an expression for the second term in the right-hand side of equation (29.91). Therefore,

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^a \mathbf{t}^b | 0 \rangle = \ell (\ell - C/2). \quad (29.93)$$

Renormalization constants at one-loop. To obtain the two one-loop renormalization constants, we express that the functional $\mathcal{W}(J)$ is one-loop finite when calculated with the renormalized action

$$\mathcal{S}_r = \frac{1}{h Z_h} \int d^d x \left[\frac{1}{2} \langle \nabla \phi(x) | \nabla \phi(x) \rangle - Z_\phi^{-1/2} Z_h \langle J(x) | \phi(x) \rangle \right], \quad (29.94)$$

for the special constant source

$$|J(x)\rangle = m^2 |0\rangle. \quad (29.95)$$

Dimensional regularization is now used in such a way that the measure term can be omitted. The action has to be expanded up to order $\xi^4(x)$. With the definitions (29.81) and (29.83) using

$$\mathbf{R}^{-1}\nabla\mathbf{R} = \mathbf{t} \cdot \nabla\xi + \frac{1}{2}[\mathbf{t} \cdot \nabla\xi, \mathbf{t} \cdot \xi] + \frac{1}{6}[[\nabla\xi \cdot \mathbf{t}, \xi \cdot \mathbf{t}], \xi \cdot \mathbf{t}] + O(\xi^4), \quad (29.96)$$

which, for *symmetric spaces*, leads to

$$\langle \nabla\phi(x) | \nabla\phi(x) \rangle = (\nabla\xi(x))^2 + \frac{1}{3} \sum_{\gamma} (f_{\gamma ab}\xi_a(x)\partial_{\mu}\xi_b(x))^2 + O(\xi^5), \quad (29.97)$$

we can expand the renormalized action up to order ξ^4 :

$$\begin{aligned} \mathcal{S}_r(\xi) = & \frac{1}{h} \int d^d x \left\{ \frac{1}{Z_h} \left[\frac{1}{2} (\nabla\xi(x))^2 + \frac{1}{6} \sum_{\gamma} (f_{\gamma ab}\xi_a(x)\nabla\xi_b(x))^2 \right] \right. \\ & \left. - Z_{\phi}^{-1/2} m^2 [1 - \frac{1}{2}\xi^2(x) + \frac{1}{24} \langle 0 | t^a \mathbf{t}^b \mathbf{t}^c \mathbf{t}^d | 0 \rangle \xi_a(x)\xi_b(x)\xi_c(x)\xi_d(x)] \right\} + O(\xi^5). \end{aligned}$$

A short calculation yields the free (or vacuum) energy density $W(m^2) = \ln \mathcal{Z}(m^2)/\text{volume}$ up to order h :

$$\begin{aligned} W(m^2) = & Z_{\phi}^{-1/2} \frac{m^2}{h} + \frac{\ell}{2} \int d^d q \ln \left(1 + \frac{m^2}{q^2} Z_{\phi}^{-1/2} Z_h \right) \\ & - \frac{h}{8} m^2 \ell (\ell - C) \left(\int \frac{d^d q}{q^2 + m^2} \right)^2 + O(h^2). \end{aligned} \quad (29.98)$$

Setting $d = 2 + \varepsilon$, we make a Laurent expansion for ε small, and define the renormalization constants by minimal subtraction (MS):

$$Z_h = 1 + \frac{C}{4\pi\varepsilon} h + O(h^2), \quad (29.99)$$

$$Z_{\phi} = 1 + \frac{\ell}{2\pi\varepsilon} h + \frac{\ell(2\ell + C)}{(4\pi\varepsilon)^2} h^2 + O(h^3). \quad (29.100)$$

The coupling constant RG function $\beta(h)$ and the field RG function $\eta(h)$ then are given by

$$\beta(h) = \varepsilon h \left(1 + h \frac{d}{dh} \ln Z_h \right)^{-1} = \varepsilon h - \frac{C}{4\pi} h^2 + O(h^3), \quad (29.101)$$

$$\eta(h) = \beta(h) \frac{d \ln Z_{\phi}}{dh} = \frac{\ell}{2\pi} h + O(h^4). \quad (29.102)$$

In contrast with $\phi_{d=4}^4$ -like QFTs, the sign of the leading term of the β -function, in the dimension in which the theory is just renormalizable ($\varepsilon = 0$), is *negative*. The physical significance of this property (asymptotic freedom) is discussed in Section 19.12.

Symmetric spaces in $2 + \varepsilon$ dimensions. We have shown that, from the point of view of the renormalization group, the properties of the non-linear σ -model generalize to all models defined on symmetric spaces. They are all UV-free in two dimensions and, therefore, have a phase transition at a critical temperature of order ε in $2 + \varepsilon$ dimension. However, the identification of the correlation functions of these models with the correlation functions of a ϕ^4 -type theory at an IR fixed point, in general, is not easy. In particular, the connection through a large N expansion does not exist in general. It is likely that, for some of these models, the transition found from the $2 + \varepsilon$ expansion is actually a first order transition for ε not infinitesimal. Note that the β -function has been calculated up to four loops for a large class of symmetric spaces [316].

29.6.2 One-loop β -function and background field method

If one wants to only calculate the coefficients of the perturbative expansion of the β -function, one can shorten the calculation by using the background field method, whose principles are explained in Section A7.2. As an exercise, and to exhibit some advantages of the method, we again perform the calculation at one-loop order. For this purpose, we evaluate the vacuum amplitude $Z(\theta)$ and the free-energy $\mathcal{W}(\theta)$ in a finite volume, with non-trivial boundary conditions (for details, see Chapter 32). In $(d - 1)$ dimensions, the coordinates x_μ vary in the interval

$$0 \leq x_\mu \leq L_\perp, \quad 1 \leq \mu \leq d - 1,$$

and we impose to the field the periodic boundary conditions

$$\phi(z, x_1, \dots, 0, \dots, x_{d-1}) = \phi(z, x_1, \dots, L_\perp, \dots, x_{d-1}).$$

We have denoted the imaginary time coordinate by z , which varies in the interval $0 \leq z \leq L$, and for which we impose fixed ‘twisted’ boundary conditions,

$$|\phi_\theta(z = 0, \mathbf{x})\rangle = |0\rangle, \quad |\phi_\theta(z = L, \mathbf{x})\rangle = e^\theta |0\rangle, \quad (29.103)$$

in which θ is a linear combination of generators belonging to $\mathcal{L}(G/H)$:

$$\theta = \sum_a t^a \theta_a, \quad t^a \in \mathcal{L}(G/H). \quad (29.104)$$

As a consequence, in the Fourier representation, momenta are quantized:

$$p_\mu = \frac{2\pi}{L_\perp} n_\mu, \quad \text{with } n_\mu \in \mathbb{Z}^{d-1}, \quad p_z = \frac{\pi}{L} m, \quad \text{with } m \in \mathbb{Z}.$$

The large L limit will be taken before the large L_\perp limit.

To deal with the longitudinal boundary conditions, it is convenient to set

$$|\phi_\theta(z, \mathbf{x})\rangle = e^{\theta z/L} |\phi(z, \mathbf{x})\rangle. \quad (29.105)$$

The new field then satisfies

$$|\phi(0, \mathbf{x})\rangle = |\phi(L, \mathbf{x})\rangle = |0\rangle, \quad (29.106)$$

which in the parametrization

$$|\phi(z, \mathbf{x})\rangle = \exp[t^a \xi_a(z, \mathbf{x})] |0\rangle, \quad (29.107)$$

is equivalent to

$$\xi_a(0, \mathbf{x}) = \xi_a(L, \mathbf{x}) = 0. \quad (29.108)$$

The renormalized action then reads ($\partial_z \equiv \partial/\partial z$):

$$\mathcal{S}_r(\xi) = \frac{1}{2hZ_h} \int dz d^{d-1}x \left[\langle \nabla \phi | \nabla \phi \rangle + \frac{2}{L} \langle \partial_z \phi | \theta | \phi \rangle - \frac{1}{L^2} \langle \phi | \theta^2 | \phi \rangle \right]. \quad (29.109)$$

The calculation of the one-loop contribution only involves the expansion of $\mathcal{S}_r(\xi)$ up to order ξ^2 . Using the relations (29.85) and (29.88),

$$\langle 0 | \mathbf{t}^a \mathbf{t}^b | 0 \rangle = -\delta_{ab}, \quad \langle 0 | \mathbf{t}^a \mathbf{t}^b \mathbf{t}^c | 0 \rangle = 0,$$

one finds

$$\langle \nabla \phi | \nabla \phi \rangle = (\nabla \boldsymbol{\xi})^2 + O(\xi^3), \quad (29.110)$$

$$\int dz \langle \partial_z \phi | \boldsymbol{\theta} | \phi \rangle = \int \partial_z \xi_a \theta_a dz + O(\xi^3) = O(\xi^3), \quad (29.111)$$

$$\langle \phi | \boldsymbol{\theta}^2 | \phi \rangle = -\theta_a \theta_a + \frac{1}{2} \xi_a \xi_b \langle 0 | [\mathbf{t}^a, [\mathbf{t}^b, \boldsymbol{\theta}^2]] | 0 \rangle + O(\xi^3). \quad (29.112)$$

To evaluate the last term, we need

$$V_{abcd} = \frac{1}{2} \langle 0 | [\mathbf{t}^a, [\mathbf{t}^b, \mathbf{t}^c \mathbf{t}^d]] | 0 \rangle = \frac{1}{2} \sum_{\varepsilon} (f_{\varepsilon ad} f_{\varepsilon bc} + f_{\varepsilon ac} f_{\varepsilon bd}). \quad (29.113)$$

The action in the Gaussian approximation is

$$\mathcal{S}_r(\xi) = \frac{L_\perp^{d-1} \theta_a \theta_a}{2LhZ_h} + \int d^{d-1}x dz \left[\frac{1}{2} (\nabla \boldsymbol{\xi})^2 - \frac{1}{2L^2} \xi_a \xi_b \theta_c \theta_d V_{abcd} \right] + O(\xi^3). \quad (29.114)$$

At one-loop order, the free energy as a function of θ_a is given by

$$\mathcal{W}(\theta) = -\frac{L_\perp^{d-1} \theta_a \theta_a}{2LhZ_h} - \frac{1}{2} \text{tr} \ln \left\{ -[\partial_z^2 + \nabla_\perp^2] \delta_{ab} - \frac{1}{L^2} V_{abcd} \theta_c \theta_d \right\}, \quad (29.115)$$

where ∇_\perp^2 is the Laplacian in $(d-1)$ dimensions. To determine the renormalization constant Z_h , it is sufficient to expand up to second order in θ :

$$\mathcal{W}(\theta) = -\frac{L_\perp^{d-1} \theta_a \theta_a}{2LhZ_h} + \frac{1}{2L^2} V_{aacd} \theta_c \theta_d \sum_{\substack{p_z=m\pi/L \\ \mathbf{p}=2\pi\mathbf{n}/L_\perp}} \frac{1}{p_z^2 + \mathbf{p}^2}. \quad (29.116)$$

We have shown that in a symmetric space (equation (29.92))

$$V_{aacd} = \sum_{\varepsilon} f_{\varepsilon ad} f_{\varepsilon ac} = \frac{1}{2} C \delta_{cd} > 0,$$

in which C is the Casimir of the group G .

Since we want to evaluate the UV divergences of the one-loop sum for $d = 2 + \varepsilon$, we can replace the sum by an integral:

$$\sum_{p_z, \mathbf{p}} \frac{1}{p_z^2 + \mathbf{p}^2} \sim \frac{LL_\perp}{(2\pi)^2} \int_{|p|>1} \frac{d^d p}{\mathbf{p}^2} \sim -\frac{LL_\perp}{2\pi\varepsilon}. \quad (29.117)$$

It follows that Z_h in the MS scheme is

$$Z_h = 1 + \frac{C}{4\pi\varepsilon} h + O(h^2), \quad (29.118)$$

in agreement with the result (29.99) of Section 29.6.1.

29.7 Generalizations

Non-compact homogeneous spaces. So far, we have restricted the discussion to homogeneous spaces based on compact groups. However, most arguments can be generalized to a class of homogeneous spaces G/H in which G is non-compact. The restrictions are that the metric g_{ij} should be positive, and the group G unimodular, to preserve the G invariance of the functional measure. For example, we can take G semi-simple. The positivity of the metric implies that H is then the maximal compact subgroup of G .

Some care has to be taken to define the field integral, since the volume of the group manifold is in general infinite, so that appropriate boundary conditions are required.

Symmetric spaces corresponding to non-compact groups. Simple examples of models belonging to this class are provided by analytic continuations of compact symmetric spaces, replacing, formally, in the orthogonal representation the generators \mathbf{t}^a in G/H by $i\mathbf{t}^a$. For example, after this transformation, the compact Grassmannian manifold $O(N+M)/O(M) \times O(N)$ becomes the manifold $O(M, N)/O(M) \times O(N)$, in which $O(M, N)$ is the pseudo-orthogonal group leaving the signature of the metric with $M +$ signs and $N -$ signs invariant.

From the point of view of perturbation theory, the only difference is that the sign of the coupling has changed. However, this means that these models are no longer UV-free but IR-free in two dimensions. They have a non-trivial IR fixed point in $d = 2 - \varepsilon$ dimension. The physics is thus completely different: the situation bears some analogies with the behaviour of the ϕ^4 theory in four dimensions. The existence of massless modes below two dimensions is not in contradiction with rigorous theorems, because the symmetry group is not compact. Such models play a role in the theory of localization [317].

Arbitrary Riemannian manifolds. A much wider class of models has been studied [318]. One considers an arbitrary smooth (infinitely differentiable) Riemannian manifold, with a smooth positive definite metric $g_{ij}(\varphi)$, and one takes the classical action

$$\mathcal{S}(\varphi) = \frac{1}{2} \int d^d x \nabla \varphi^i(x) g_{ij}(\varphi(x)) \nabla \varphi^j(x). \quad (29.119)$$

The generating functional of correlation functions $\mathcal{Z}(J)$ is then given by

$$\mathcal{Z}(J) = \int [d\rho(\varphi)] \exp \left[-\mathcal{S}(\varphi) + \int d^d x J_i(x) \varphi^i(x) \right], \quad (29.120)$$

in which $d\rho(\varphi)$ has to be a smooth, strictly positive, covariant measure on the manifold, for example (see Chapter 28),

$$d\rho(\varphi) = \sqrt{g} d\varphi,$$

in which g is the determinant of the metric tensor.

As we discuss in Sections 3.3 and 34.9, in one dimension, the field integral (29.120) is associated with a Hamiltonian of the form of a Laplace operator in the manifold with metric g_{ij} , and is related to Brownian motion or diffusion processes on the manifold.

In two dimensions, the action (29.119) corresponds to theories renormalizable by power counting, which generalizes theories on homogeneous spaces. However, in the absence of symmetries, and, thus, of the Goldstone modes which forces fields to remain massless, important differences appear after renormalization.

- (i) In general, the space of all possible metrics is infinite dimensional.
- (ii) A metric can involve an infinite number of scalars like the scalar curvature (see Chapter 28), and a covariant integration measure multiplied by a scalar is still covariant.

As a consequence, in contrast to homogeneous cases, derivative-free terms will be generated in the renormalization, and it is only possible to maintain the form (29.119) of the action by adjusting an infinite number of parameters (a massive fine-tuning problem).

In addition, the renormalized metric will be generically the most general metric in the manifold. In other words, the renormalized action is the most general action allowed by power counting arguments.

Considerations based on covariance (Chapter 28), nevertheless, are useful, since they simplify perturbative calculations in the general situation by restricting the form of the counter-terms at a given order in the loop expansion.

For example, the equivalent of the coupling constant RG function is a functional of the metric $\beta(g_{ij})$. It has the covariance of the metric tensor. At one-loop order, it can only involve first and second derivatives of the metric and, therefore, R_{ij} the Ricci tensor and Rg_{ij} in which R is the scalar curvature. By inspection, it is possible to eliminate Rg_{ij} and the constant in front of R_{ij} can be obtained from a particular model. We have expressed the coefficients in terms of g^{ij} , the inverse of the metric tensor g_{ij} , because it naturally orders perturbation theory. In $2 + \varepsilon$ dimensions [318],

$$\beta^{ij}(g) = \varepsilon g^{ij} - \frac{1}{2\pi} R^{ij} - \frac{1}{8\pi^2} R^{iklm} R_{klm}^j + O\left((g^{ij})^4\right), \quad (29.121)$$

in which R^{ij} is obtained from the Ricci tensor (equation (28.64)) by raising the indices with g^{ij} and R_{klm}^j is the curvature tensor (equation (28.56), see Chapter 28 for details).

A29 Homogeneous spaces: A few algebraic properties

The Appendix first describes a few additional properties of homogeneous spaces, when considered as Riemannian manifolds. It assumes a minimal familiarity with the elements of differential geometry presented in Chapter 28. The second part is devoted to a few elements of classification of symmetric spaces.

A29.1 Pure gauge. Maurer–Cartan equations

The notation has been defined in Section 29.1, in particular, the matrix \mathbf{R} belonging to the representation of the group G by equation (29.4), and the coordinates φ^i parametrize the coset space G/H (Section 29.1.3).

Multiplying $\mathbf{R}(\varphi)$ by an element \mathbf{R}_g of the representation of the group G on the left, we note that $\mathbf{R}^{-1}\partial_i\mathbf{R}$ ($\partial_i \equiv \partial/\partial\varphi^i$) transforms like

$$\mathbf{R}^{-1}\partial_i\mathbf{R} \mapsto \mathbf{R}_g^{-1}\mathbf{R}^{-1}\partial_i\mathbf{R}\mathbf{R}_g.$$

This implies that the matrices $\mathbf{R}^{-1}\partial_i\mathbf{R}$ transform like elements of the adjoint representation of the Lie algebra $\mathcal{L}(G)$ of the group G . Therefore, they can be expanded on the generators \mathbf{t}^α , and we set (equation (29.10))

$$\mathbf{R}^{-1}\partial_i\mathbf{R} = L_{\alpha,i}(\varphi)\mathbf{t}^\alpha. \quad (\text{A29.1})$$

It is easy to verify more directly, by parametrizing the left-hand side as

$$\mathbf{R} = e^{\mathbf{t}^\alpha \xi_\alpha},$$

that the expansion of $L_{\alpha,i}$ in powers of ξ only involves commutators of generators of $\mathcal{L}(G)$ and, therefore, $L_{\alpha,i}(\varphi)$ depends on the parametrization of the group elements, but not on the representation to which $\mathbf{R}(\varphi)$ belongs.

The definition (A29.1) implies that the quantities $L_{\alpha,i}(\varphi)$ are the components of a vector transforming under the adjoint representation, and have the φ^i dependence of a pure gauge (see Chapters 22 and 23 for details). We have shown in Section 29.5 that the corresponding curvature vanishes (equation (29.62)). In terms of the components $L_{\alpha,i}$, one finds

$$\begin{aligned} \partial_i(\mathbf{R}^{-1}\partial_j\mathbf{R}) - \partial_j(\mathbf{R}^{-1}\partial_i\mathbf{R}) &= \mathbf{t}^\alpha [\partial_i L_{\alpha,j} - \partial_j L_{\alpha,i}], \\ [\mathbf{R}^{-1}\partial_i\mathbf{R}, \mathbf{R}^{-1}\partial_j\mathbf{R}] &= L_{\alpha,i}L_{\beta,j}f_{\gamma}^{\alpha\beta}\mathbf{t}^\gamma. \end{aligned}$$

As a consequence,

$$\partial_i L_{\alpha,j} - \partial_j L_{\alpha,i} + f_{\alpha}^{\beta\gamma}L_{\beta,i}L_{\gamma,j} = 0. \quad (\text{A29.2})$$

These relations, which express that the curvature corresponding to the gauge connection L_i^α vanishes, are known as the Maurer–Cartan equations.

A29.2 Metric and curvature in homogeneous spaces

A group transformation acting on the coordinates φ^i can also be considered as a reparametrization of the manifold. The infinitesimal form is given by equation (29.36),

$$\varphi^i = \varphi'^i + D_\alpha^i(\varphi')\omega^\alpha. \quad (\text{A29.3})$$

The generator \mathcal{D}_α of $\mathcal{L}(G)$, as defined by equation (29.40), then characterizes the corresponding infinitesimal variation of scalars:

$$\mathcal{D}_\alpha S(\varphi) = D_\alpha^i(\varphi)\partial_i S(\varphi). \quad (\text{A29.4})$$

More generally, equation (28.10) defines its action on all tensors on the homogeneous space. For vectors $V_i(\varphi)$, it yields

$$\mathcal{D}_\alpha V_i = D_\alpha^j \partial_j V_i + \partial_i D_\alpha^j V_j. \quad (\text{A29.5})$$

This property can be verified by a short calculation in the case of the gauge field $L_i^a(\varphi)$, defined by equation (29.10). As explained in Section 28.1, $\mathcal{D}_\alpha V_i$ is a vector, as expected.

For general tensors, the result is

$$\mathcal{D}_\alpha V_{i_{p+1} \dots i_n}^{i_1 \dots i_p} = D_\alpha^j \partial_j V_{i_{p+1} \dots i_n}^{i_1 \dots i_p} - \sum_{\ell=1}^p \partial_j D_\alpha^{\ell} V_{i_{p+1} \dots i_n}^{i_1 \dots j \dots i_p} + \sum_{\ell=p+1}^n \partial_{i_\ell} D_\alpha^j V_{i_{p+1} \dots j \dots i_n}^{i_1 \dots i_p}. \quad (\text{A29.6})$$

With this definition, \mathcal{D}_α obeys the usual rule of differentiation for products of tensors.

The invariance of the metric, as expressed by equation (29.39), then takes the simple form

$$\mathcal{D}_\alpha g_{jk} = 0 \iff \nabla_i D_{jk} + \nabla_j D_{ik} = 0, \quad (\text{A29.7})$$

where the second equation follows from equation (28.111).

Consistency with parallel transport. The metric tensor defines uniquely a torsion-free parallel transport on the manifold. If the infinitesimal change of variables (A29.3) leaves the metric invariant, it leaves invariant all quantities that are functions only of the metric. With the definition (A29.6), we can write

$$\mathcal{D}_\alpha R_{lij}^k = 0, \quad (\text{A29.8})$$

$$\mathcal{D}_\alpha R_{ij} = 0, \quad (\text{A29.9})$$

$$\mathcal{D}_\alpha R = 0. \quad (\text{A29.10})$$

The Christoffel connection is also invariant but, since it is not a tensor, the action of \mathcal{D}_α takes the inhomogeneous form (28.35)

$$\mathcal{D}_\alpha \Gamma_{jk}^i = \partial_j D_\alpha^l \Gamma_{jk}^i - \partial_l D_\alpha^i \Gamma_{jk}^l + \partial_k D_\alpha^l \Gamma_{jl}^i + \partial_j D_\alpha^l \Gamma_{lk}^i + \partial_j \partial_k D_\alpha^i = 0.$$

The compatibility between parallel transport and symmetry can then be expressed by the commutation relation

$$[\mathcal{D}_\alpha, \nabla_i] = 0. \quad (\text{A29.11})$$

To prove the relation, it is sufficient to verify it on scalars and vectors, it then follows from forming tensor products. For scalars, it is an immediate consequence of the definition (A29.6). For vectors, one finds

$$[\mathcal{D}_\alpha, \nabla_i] V_j = V^k \mathcal{D}_\alpha \Gamma_{ki}^j = 0.$$

Equation (A29.9) implies that R_{ij} is an acceptable metric tensor, and equation (A29.10) that the scalar curvature is a constant in homogeneous spaces. More generally, all symmetric tensors with two indices constructed from the curvature tensor satisfy the equivalent of equation (A29.9), and are of the form of a metric tensor. Since, in the case of homogeneous spaces, as we have shown, the set of metrics form a finite-dimensional vector space, only a finite number of these tensors are linearly independent.

As a final remark, note that expression (29.12) for the metric shows that the quantities $L_i^a(\varphi)$ essentially play the role of a vielbein in the case of homogeneous spaces, the only difference being the constant internal metric μ_{ab} .

Remark. In Chapter 29, the coordinates φ^i are themselves fields depending on variables x^μ . Then $\nabla \varphi^i(x)$ belongs to the space tangent to the manifold at point $\varphi^i(x)$, and thus transforms like a vector. Using the metric tensor, we have then constructed scalars like the action density of (29.8). Moreover, in this situation, another covariant derivative D_μ can be defined, which involves the connection (28.22):

$$D_\mu V^i = \partial_\mu V^i + \Gamma_{kj}^i \partial_\mu \varphi^j V^k. \quad (\text{A29.12})$$

For functions only of $\varphi^i(x)$, this definition is redundant, since D_μ can be rewritten in terms of the covariant derivative, but it is useful when applied to derivatives of the field φ . Generalization to higher-order tensors is straightforward: one contracts the free index with $\nabla \varphi^i$.

The definition (A29.12) makes it possible to express the classical field equation corresponding to the action (29.119) in covariant form:

$$D_\mu \nabla \varphi^i(x) = 0. \quad (\text{A29.13})$$

A29.3 Explicit expressions for the metric

Finally, let us give several more explicit expressions for the metric.

A29.3.1 Metric tensor and transformations

For a general homogeneous space, one can find an inverse metric tensor g^{ij} of the form

$$g^{ij}(\varphi) = D_\alpha^i(\varphi) m^{\alpha\beta} D_\beta^j(\varphi), \quad (\text{A29.14})$$

where $m^{\alpha\beta}$ is a constant, symmetric, non-singular matrix. The tensor has to satisfy the equivalent of equations (29.39) or (A29.7):

$$D_\alpha^k \partial_k g^{ij} = g^{ik} \partial_k D_\alpha^j + g^{jk} \partial_k D_\alpha^i.$$

We replace g^{ij} by the form (A29.14):

$$m^{\beta\gamma} D_\alpha^k \left(\partial_k D_\beta^i D_\gamma^j + \partial_k D_\beta^j D_\gamma^i \right) = m^{\beta\gamma} D_\gamma^k \left(D_\beta^i \partial_k D_\alpha^j + D_\beta^j \partial_k D_\alpha^i \right),$$

and use the Lie algebra commutation relations (29.42).

We then obtain

$$m^{\beta\gamma} f_{\alpha\beta}^\delta \left(D_\gamma^j D_\delta^i + D_\gamma^i D_\delta^j \right) = 0. \quad (\text{A29.15})$$

Exchanging $\gamma \leftrightarrow \delta$ in one of the terms, we note that the equation is satisfied if $m^{\alpha\beta}$ is a solution to the numerical equation,

$$m^{\gamma\beta} f_{\beta\alpha}^\delta = f_{\alpha\beta}^\gamma m^{\beta\delta}.$$

In a basis in which the generators \mathbf{t}^α are orthogonal by the trace, the structure constants are antisymmetric and $m^{\alpha\beta} = m\delta_{\alpha\beta}$ is the solution to the last equation.

In the case of symmetric spaces, expressions simplify. The vector $\mathbf{t}^a |0\rangle$ transforms under an irreducible representation of H ; the matrix μ defined by equation (29.13),

$$\mu^{ab} = -\langle 0 | \mathbf{t}^a \mathbf{t}^b | 0 \rangle, \quad (\text{A29.16})$$

is diagonal

$$\mu^{ab} = \mu\delta_{ab}.$$

This provides another proof of the uniqueness of the metric.

The tensor $g^{ij}(\varphi)$ (equation (A29.14)) is a possible inverse metric tensor. In the case of symmetric spaces, the unique metric, in a basis in which the generators \mathbf{t}^α are orthogonal by the trace, is thus,

$$g^{ij}(\varphi) = D_\alpha^i(\varphi) D_\alpha^j(\varphi), \quad (\text{A29.17})$$

up to a normalization.

A29.3.2 Manifolds embedded in Euclidean space

If one knows an embedding of a Riemannian manifold \mathfrak{M} in Euclidean space, one can describe it with constrained Euclidean coordinates, as we have done in the case of homogeneous spaces in Section 29.2.

Let (σ^s, φ^i) be such a set of Euclidean coordinates. We assume that locally the σ^s can be expressed as functions of the independent coordinates φ^i :

$$\sigma^s = \sigma^s(\varphi). \quad (\text{A29.18})$$

The metric tensor in this representation is obtained from

$$g_{ij}(\varphi) d\varphi^i d\varphi^j = d\varphi^i d\varphi^j + d\sigma^s d\sigma^s \quad (\text{A29.19})$$

and, therefore,

$$g_{ij}(\varphi) = \delta_{ij} + \partial_i \sigma^s(\varphi) \partial_j \sigma^s(\varphi). \quad (\text{A29.20})$$

A short calculation shows that the connection has the simple form

$$\Gamma_{jk}^i = g^{il}(\varphi) \partial_l \sigma^s(\varphi) \partial_j \partial_k \sigma^s(\varphi), \quad (\text{A29.21})$$

and the curvature tensor is given by

$$R_{ijkl} = [\partial_i \partial_k \sigma^s(\varphi) \partial_j \partial_l \sigma^s - \partial_i \partial_l \sigma^s(\varphi) \partial_j \partial_k \sigma^s(\varphi)] [\text{tr}(\mathbf{g}^{-1}(\varphi)) - N + 1], \quad (\text{A29.22})$$

in which N is the dimension of \mathfrak{M} and $\text{tr } \mathbf{g}^{-1}$ the trace of the inverse of the metric:

$$\text{tr } \mathbf{g}^{-1} = \text{tr}_{st} [\delta_{st} + \partial_k \sigma^s(\varphi) \partial_k \sigma^t(\varphi)]^{(-1)}. \quad (\text{A29.23})$$

A29.4 Symmetric spaces: Classification

We examine now a few simple properties of symmetric spaces and describe a few important families.

A29.4.1 Definition

Let us consider a semi-simple compact Lie group G , and assume that it can be equipped with a non-trivial involutive automorphism of G , which to an element g of G associates an element \bar{g} :

$$(\overline{g_1 g_2}) = \bar{g}_1 \bar{g}_2, \quad \text{with } \bar{\bar{g}} = g. \quad (\text{A29.24})$$

We consider the coset space G/H , where H is the subgroup of elements invariant under the automorphism:

$$\overline{H} \equiv H. \quad (\text{A29.25})$$

The automorphism can be extended to the Lie algebra $\mathcal{L}(G)$. It then becomes a reflection, and each element of $\mathcal{L}(G)$ can be decomposed into a sum of an even and an odd element. By definition, even elements belong to $\mathcal{L}(H)$, and the generators of $\mathcal{L}(H)$ are denoted by τ^α . The generators of $\mathcal{L}(G)$ not belonging to $\mathcal{L}(H)$ (we denote the corresponding vector space $\mathcal{L}(G/H)$) can be chosen odd. We denote them by t^a . We choose the Lie algebra structure constants f_{ijk} to be completely antisymmetric. This leads to the rules

$$\begin{aligned} \bar{t}^a &= -t^a, & t^a \in \mathcal{L}(G/H); \\ \bar{\tau}^\alpha &= -\tau^\alpha, & \tau^\alpha \in \mathcal{L}(H). \end{aligned} \quad (\text{A29.26})$$

It follows that

$$[\tau^\alpha, \tau^\beta] = f_{\alpha\beta\gamma} \tau^\gamma, \quad (\text{A29.27})$$

$$[\tau^\alpha, t^b] = f_{\alpha bc} t^c, \quad (\text{A29.28})$$

$$[t^a, t^b] = f_{ab\gamma} \tau^\gamma. \quad (\text{A29.29})$$

In the case of a compact group, only the last set (A29.29) of commutation relations is characteristic of a symmetric space, since (A29.28) is then a consequence of (A29.27) and the antisymmetry of f_{abc} .

Preliminary remarks

(i) We will consider symmetric spaces derived from non-simple groups G , but we want to exclude the possibility that

$$\begin{aligned} G &= G_1 \times G_2, \\ H &= H_1 \times H_2 \quad \text{with} \quad H_1 \subset G_1, H_2 \subset G_2, \end{aligned} \quad (\text{A29.30})$$

because in this case the coset space decomposes into two independent spaces G_1/H_1 and G_2/H_2 . In particular, this excludes the trivial situation $G_2 \equiv H_2$ and this property will be used in what follows.

(ii) All generators of $\mathcal{L}(H)$ can be obtained as linear combinations of commutators of generators of $\mathcal{L}(G/H)$. Indeed, if we assume that a generator τ_δ cannot be obtained, we can rearrange the generators in such a way that $f_{ab\delta} = 0$. It follows that τ_δ commutes with $\mathcal{L}(G/H)$, and thus with all generators of $\mathcal{L}(H)$ that can be obtained as commutators (A29.29). This is exactly a situation that we have just excluded.

A29.4.2 A basic property

The purpose of this Appendix is not a complete presentation of the mathematical properties of symmetric spaces. However, a few of these properties are directly relevant to the problem of renormalization, and can be derived by elementary methods. A very important property is the following.

If a homogeneous space G/H is symmetric, H is a maximal proper subgroup of G .

To prove this assertion, we assume that a subgroup G' of G which contains H exists, and exhibit a contradiction,

$$G \supset G' \supset H.$$

Note first that G'/H is then also a symmetric space.

If t''^a belongs to $\mathcal{L}(G) - \mathcal{L}(G')$ and t'^b belongs to $\mathcal{L}(G') - \mathcal{L}(H)$, then, from equation (A29.29), $[t''^a, t'^b]$ belongs to $\mathcal{L}(H)$. However, the relations (A29.28) and the antisymmetry of the structure constants imply that such a commutator vanishes. Since $\mathcal{L}(H)$ is obtained from the commutators of generators in $\mathcal{L}(G'/H)$, we find again that the generators t''^a commute with $\mathcal{L}(G')$, and thus

$$G = G' \times G'',$$

in which H is a subgroup of G' —the situation we have excluded.

Therefore, the maximality of H has been derived. Several other important properties of symmetric spaces follow.

A few consequences

(i) The maximality of H has one very important consequence: the generators $\{t^a\}$ form a real irreducible representation of the group H .

To derive this result, we assume the converse, that is, that the generators $\{t^a\}$ can be divided into two representations of H — \mathcal{L}_1 and \mathcal{L}_2 :

$$\begin{aligned} t^a \in \mathcal{L}_1 &\Rightarrow [t^a, \tau^\alpha] \in \mathcal{L}_1, \\ t^a \in \mathcal{L}_2 &\Rightarrow [t^a, \tau^\alpha] \in \mathcal{L}_2. \end{aligned}$$

Since, as a consequence of equation (A29.29), the commutator of two elements of \mathcal{L}_1 belongs to $\mathcal{L}(H)$, $\mathcal{L}(H) \oplus \mathcal{L}_1$ forms a subalgebra of $\mathcal{L}(G)$ and H is not maximal. The converse is obvious: if H is not maximal the representation is reducible.

(ii) Let us assume that we have defined the space G/H in the way described in Section 29.1.1. From the preceding results, we conclude that the field $\phi(x)$ of equation (29.4),

$$\phi(x) = \exp \left(\sum_{\alpha=1}^{\ell} \xi_a(x) t^a \right) \phi_c,$$

belongs to an irreducible representation of G . Also, simple considerations show that, since the $\{t^a\}$ form an irreducible representation of H , the vector ϕ_c is unique in the following sense: given H , there exists a unique vector ϕ_c in the representation which has H as little group (stabilizer).

Therefore, a unique classical model is associated to each symmetric space. The QFT is defined in terms of a unique coupling constant, and the perturbative expansion is equally unique. The general arguments on homogeneous spaces given before imply that two renormalization constants are sufficient to renormalize the model.

We now describe symmetric spaces corresponding to orthogonal and unitary groups G .

A29.4.3 The principal chiral models

We first examine the case in which G is not simple, and factorizes into $G \equiv \mathfrak{G}_1 \times \mathfrak{G}_2$. Since we have excluded the situation (A29.30), the automorphism must map elements of \mathfrak{G}_1 into \mathfrak{G}_2 and vice versa. This implies $\mathfrak{G}_1 \equiv \mathfrak{G}_2$ and the automorphism is

$$(\overline{g_1, g_2}) = (g_2, g_1) , \quad g_1 \in \mathfrak{G}_1, \mathfrak{G}_2 \in G_2 .$$

The subgroup H is then given by elements of the form

$$H \equiv \{(g, g)\} \equiv \mathfrak{G} .$$

The symmetric space $\mathfrak{G} \times \mathfrak{G} / \mathfrak{G}$ is isomorphic to the group space \mathfrak{G} itself. A canonical realization is to consider group elements of the form $g_1^{-1}g_2$.

A29.4.4 Simple groups

We now assume that the group G is simple. We decompose the automorphism into the product of an inner automorphism and a remaining irreducible involutive automorphism,

$$\bar{g} = g_0^{-1}g^*g_0 ,$$

with

$$\begin{aligned} g_1^*g_2^* &= (g_1g_2)^* , \\ (g^*)^* &= g . \end{aligned}$$

The condition $\bar{g} = g$ then takes the form

$$g_0^{-1}(g_0^{-1})^*gg_0^*g_0 = g .$$

The element $g_0^*g_0$ commutes with all elements of the group. This implies that $g_0^*g_0$ belongs to the centre of the group:

$$g_0^*g_0 = c\mathbf{1} ,$$

$\mathbf{1}$ being the unit matrix in the defining representation, and c a phase factor which reduces to $c = \pm 1$ for orthogonal groups.

The subgroup H is defined by the invariant group elements h :

$$\bar{h} = g_0^{-1}h^*g_0 = h , \quad \text{or} \quad h^*g_0 = hg_0 .$$

A realization of G/H in the group space is given by group elements of the form

$$g = (f^{-1})^*g_0f , \quad \forall f \in G . \tag{A29.31}$$

Indeed, if we multiply f on the left by an element of H , g is not modified.

Note that these group elements satisfy

$$g^*g = c\mathbf{1} . \tag{A29.32}$$

We can now classify symmetric spaces corresponding to orthogonal and unitary groups.

Orthogonal groups. The group G is $O(N)$, the star automorphism is the identity and $c = \pm 1$.

(i) $c = +1$.

We take g_0 diagonal without loss of generality. It has only ± 1 as eigenvalues. If it possesses p eigenvalues $+1$ and $(N-p)$ eigenvalues -1 , the subgroup H is $O(p) \times O(N-p)$.

The symmetric space $O(N)/O(p) \times O(N-p)$ is called a real Grassmannian manifold. The $O(N)$ non-linear σ -model corresponds to $p = 1$.

(ii) $c = -1$.

This implies that N must be even: $N = 2N'$. Without loss of generality, we can choose g_0 of the form

$$g_0 = \begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix},$$

and the subgroup which commutes with g_0 is isomorphic to the unitary group $U(N')$.

Unitary groups. The group G is $U(N)$.

(i) The star automorphism is the identity. The phase c is irrelevant. Taking $c = 1$, one notes that g_0 has only ± 1 as eigenvalues. If it has p eigenvalues $+1$ and $(N-p)$ eigenvalues -1 , the subgroup H is $U(p) \times U(N-p)$.

The symmetric spaces $U(N)/U(p) \times U(N-p)$ are called complex Grassmannian manifolds. The case $p = 1$ corresponds to the complex projective space CP_{N-1} .

(ii) The star automorphism is the complex conjugation. The condition $g_0^* g_0 = c\mathbf{1}$ then implies $c = \pm 1$. For $c = +1$, we can diagonalize g_0 by an orthogonal transformation and then set it equal to 1 by a diagonal unitary transformation. Since the elements h of H then satisfy

$$h = h^*,$$

the subgroup H is the orthogonal subgroup $O(N)$ of $U(N)$.

If we take $c = -1$, we note again that N must be even. We set

$$N = 2N'.$$

We can then choose g_0 of the form

$$g_0 = \begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix}.$$

The subgroup H is defined by the elements h , which satisfy

$$\begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix} h = h^* \begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix}.$$

This defines the symplectic group $Sp(N)$, subgroup of $U(N)$.

30 A few solvable two-dimensional quantum field theories (QFT)

This chapter is devoted to a discussion of several two-dimensional quantum field theories (QFT), which can be solved, in the sense that their properties can be determined by non-perturbative methods.

First, we consider the Schwinger model [319], a model of two-dimensional quantum electrodynamics (QED) with massless fermions, which illustrates the properties of confinement, spontaneous chiral symmetry breaking, asymptotic freedom, and anomalies, properties one also expects in particle physics from quantum chromodynamics (QCD). We discuss the sine-Gordon (sG) model [320, 321]. We prove the equivalence between the massive Thirring model [322], a fermion model with current-current interaction, and the sG model [323].

In Chapter 19, we have discussed the generic $O(N)$ -non-linear σ -model. We have pointed out that the Abelian case $N = 2$ is special, because the renormalization group (RG) β -function vanishes in two dimensions. The corresponding $O(2)$ -invariant spin model is especially interesting: it provides an example of the celebrated Kosterlitz–Thouless phase transition and will be studied in Chapter 31. This chapter also provides the necessary technical background for such an investigation.

The *bosonization* technique, based on an identity for Cauchy determinants, plays a major role in this chapter. It establishes relations, *specific to two dimensions*, between fermion and boson local field theories. For example, the relation between the massive Thirring model and the sG model relies on the bosonization technique.

Critical (thus, massless) unitary field theories are *conformal invariant*. In two dimensions, this is a powerful property, because the conformal group is infinite dimensional. A sophisticated technology has been developed to study them. A general discussion of quantum integrable systems and *conformal field theory* goes beyond the scope of this work, and the interested reader is referred to the abundant literature [324].

In the spirit of this work, field integral techniques are used everywhere, although, in some cases, the operator formulation may lead to more elegant derivations.

The appendix contains additional remarks concerning the *anomaly* in the Schwinger model and *solitons* in the sG model.

Notation. Implicit summation over *repeated lower Greek indices* is meant everywhere.

30.1 The free massless scalar field

To discuss the peculiar properties of the massless, free, scalar field in two dimensions, we consider it as the limit of the massive field. The action for a free massive scalar field $\varphi(x)$ of mass m has the form

$$\mathcal{S}(\varphi) = \frac{1}{2} \int d^2x \left[(\nabla\varphi(x))^2 + m^2\varphi^2(x) \right]. \quad (30.1)$$

The propagator of the field φ is ($\langle \bullet \rangle$ means expectation value with the weight e^{-S}),

$$\langle \varphi(x)\varphi(0) \rangle = \Delta(x, m) = \frac{1}{(2\pi)^2} \int d^2p \frac{e^{ipx}}{p^2 + m^2}. \quad (30.2)$$

It has a singularity at $x = 0$, which corresponds to an ultraviolet (UV) divergence. After introduction of a UV cut-off Λ (see, e.g. Section 8.4.2), one finds ($\gamma = -\psi(1)$ is Euler's constant)

$$\Delta(0, m) = \frac{1}{2\pi} [\ln(2\Lambda/m) - \gamma] + K + o(\Lambda^{-1}), \quad (30.3)$$

where the constant K depends on the cut-off function.

In two dimensions, the propagator of the massless scalar field is infrared (IR) divergent. The fluctuations which translate the field $\varphi(x)$ by an almost constant function are not damped enough by the action (30.1), and are responsible for this divergence.

Evaluating the massive propagator (30.2) for $m \rightarrow 0$, one obtains

$$\Delta(x, m) = -\frac{1}{4\pi} (\ln(m^2 x^2/4) + 2\gamma) + O(m). \quad (30.4)$$

30.1.1 $m = 0$: IR finite correlation functions

Although the massless scalar field itself has IR divergent correlation functions, some local functions of the field have IR finite correlation functions. For instance, the correlation functions of exponentials of the field φ are given by

$$\begin{aligned} & \left\langle \prod_{i=1}^n e^{i\kappa_i \varphi(x_i)} \right\rangle \\ & \equiv \int [d\varphi] \exp \left\{ -\frac{1}{2} \int d^2x \left[(\nabla \varphi(x))^2 + m^2 \varphi^2(x) \right] + i \sum_i \kappa_i \varphi(x_i) \right\}, \end{aligned} \quad (30.5)$$

where the κ_i 's are arbitrary real coefficients. We introduce the source

$$J(x) = i \sum_i \kappa_i \delta^{(2)}(x - x_i). \quad (30.6)$$

The field integral then takes the form

$$\int [d\varphi] \exp \left\{ - \int d^2x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} m^2 \varphi^2(x) - J(x) \varphi(x) \right] \right\}.$$

We recognize the basic Gaussian field integral

$$\begin{aligned} & \left\langle \prod_{i=1}^n e^{i\kappa_i \varphi(x_i)} \right\rangle = \exp \left[\frac{1}{2} \int d^2x d^2y J(x) \Delta(x - y, m) J(y) \right] \\ & = \exp \left[-\frac{1}{2} \sum_{i,j} \kappa_i \kappa_j \Delta(x_i - x_j, m) \right]. \end{aligned} \quad (30.7)$$

Replacing the propagator by its small m expansion (equations (30.4, 30.3)), and collecting all divergent contributions that do not vanish with m , one finds

$$\begin{aligned} \sum_{i,j} \kappa_i \kappa_j \Delta(x_i - x_j, m) & = -\frac{1}{2\pi} \left[\left(\sum_i \kappa_i \right)^2 (\ln m - \ln 2 + \gamma) - \sum_i \kappa_i^2 \ln \Lambda \right. \\ & \quad \left. + \sum_{i \neq j} \kappa_i \kappa_j \ln |x_i - x_j| \right] + O(m). \end{aligned} \quad (30.8)$$

Therefore, due to the IR divergences, when m goes to 0, only the products such that $\sum_i \kappa_i = 0$ survive.

This result has a simple interpretation in terms of the two-dimensional Coulomb gas: the potential created by point-like charges decreases at large distance only if the total system is neutral (for a discussion of the Coulomb gas, see Section 31.3.1).

From equation (30.8), one also learns that the UV divergences can be removed by a multiplicative renormalization ζ_i of the composite fields $e^{i\kappa_i\varphi(x)}$:

$$e^{i\kappa_i\varphi(x)} = \zeta_i \left[e^{i\kappa_i\varphi(x)} \right]_r, \quad \text{with } \zeta_i = (\Lambda/\mu)^{-\kappa_i^2/(4\pi)}, \quad (30.9)$$

where μ is a renormalization scale. The dimension of the operator $e^{i\kappa\varphi(x)}$ follows:

$$[e^{i\kappa\varphi(x)}] = \kappa^2/4\pi. \quad (30.10)$$

Thus, in the massless limit $m = 0$, for $\sum_i \kappa_i = 0$, the renormalized correlation functions have the form,

$$\left\langle \prod_{i=1}^n e^{i\kappa_i\varphi(x_i)} \right\rangle_r = \prod_{i < j} (\mu |x_i - x_j|)^{\kappa_i \kappa_j / 2\pi}. \quad (30.11)$$

One can verify that these correlation functions of exponentials of the field decay algebraically at large distance and satisfy, as they should, the *cluster property* (Section 7.6).

The two-charge example. The situation where the κ_i 's only take two values $\pm\kappa$ is of special interest. Then, the number of plus and minus signs must be equal. One finds,

$$\left\langle \prod_{i=1}^n e^{i\kappa[\varphi(x_i) - \varphi(y_i)]} \right\rangle_r = \frac{\prod_{i < j} (\mu |x_i - x_j|)^{\kappa^2/2\pi} (\mu |y_i - y_j|)^{\kappa^2/2\pi}}{\prod_{i,j} (\mu |x_i - y_j|)^{\kappa^2/2\pi}}. \quad (30.12)$$

In particular, the two-point function becomes

$$\left\langle e^{i\kappa\varphi(x)} e^{-i\kappa\varphi(0)} \right\rangle \propto x^{-\kappa^2/2\pi}. \quad (30.13)$$

For $\kappa^2 \geq 4\pi$, the two-point function is singular at short distance in the sense of distributions. It has to be renormalized as the UV divergence of its Fourier transform,

$$\int d^2x \frac{e^{ipx}}{x^{\kappa^2/2\pi}} = \pi \frac{\Gamma(1 - \kappa^2/4\pi)}{\Gamma(\kappa^2/4\pi)} \left(\frac{p^2}{4} \right)^{\kappa^2/4\pi - 1},$$

also shows. This expression is valid for $\kappa^2 < 4\pi$. At $\kappa^2 = 4\pi$, a divergent constant has to be subtracted. For $4\pi \leq \kappa^2 < 8\pi$, it is thus defined only up to an arbitrary additive renormalization. For $\kappa^2 = 8\pi$, a second additive renormalization proportional to p^2 is required. We meet again these two special values when we discuss the sG model.

30.1.2 Symmetries and currents

In the massless limit $m = 0$, the action (30.1) becomes invariant under constant translations of the field. In the translation

$$\varphi(x) \mapsto \varphi(x) + \omega, \quad (30.14)$$

the field $e^{i\kappa_i\varphi}$ becomes

$$e^{i\kappa_i\varphi(x)} \mapsto e^{i\omega\kappa_i} e^{i\kappa_i\varphi(x)}.$$

All non-vanishing correlation functions (30.5) are invariant under this transformation, a result consistent with the analysis in Section 19.3, indicating the absence of spontaneous breaking of continuous symmetries in two dimensions.

To this symmetry corresponds the current ($\partial_\mu \equiv \partial/\partial x_\mu$),

$$J_\mu^V(x) = \partial_\mu \varphi(x), \quad (30.15)$$

which is classically conserved. In two dimensions, another current exists,

$$J_\mu^A(x) = \epsilon_{\mu\nu} \partial_\nu \varphi(x) = \epsilon_{\mu\nu} J_\nu^V(x), \quad (30.16)$$

which is trivially conserved ($\epsilon_{\mu\nu}$ is the antisymmetric tensor with $\epsilon_{12} = 1$).

The currents (30.15) and (30.16) provide other examples of fields with IR-finite correlation functions. For instance, in the Fourier representation, the two-point functions are

$$\langle \tilde{J}_\mu^V(k) \tilde{J}_\nu^V(-k) \rangle = k_\mu k_\nu / k^2, \quad (30.17a)$$

$$\langle \tilde{J}_\mu^A(k) \tilde{J}_\nu^A(-k) \rangle = \delta_{\mu\nu} - k_\mu k_\nu / k^2. \quad (30.17b)$$

After Fourier transformation, the two-point functions become

$$\langle J_\mu^V(x) J_\nu^V(0) \rangle = \frac{1}{2\pi x^2} \left(\delta_{\mu\nu} - 2 \frac{x_\mu x_\nu}{x^2} \right), \quad (30.18a)$$

$$\langle J_\mu^A(x) J_\nu^A(0) \rangle = -\frac{1}{2\pi x^2} \left(\delta_{\mu\nu} - 2 \frac{x_\mu x_\nu}{x^2} \right). \quad (30.18b)$$

The two functions are proportional (as implied by current conservation), although, in the Fourier representation, they are different: the right-hand side involves a singular distribution (in the mathematical sense), ambiguous at $x = 0$, defined up to the addition of a $\delta^{(2)}(x)$ function (really distribution).

It is simple to calculate the effect of the insertion of the current in a correlation function of exponentials, since this involves calculating a one-point function in an external source. One finds

$$\begin{aligned} \left\langle J_\mu^V(x) \prod_{i=1}^n e^{i\kappa_i \varphi(x_i)} \right\rangle &= i \sum_i \kappa_i \partial_\mu^x \Delta(x - x_i) \left\langle \prod_{i=1}^n e^{i\kappa_i \varphi(x_i)} \right\rangle \\ &= -\frac{i}{2\pi} \sum_i \kappa_i \frac{(x - x_i)_\mu}{(x - x_i)^2} \left\langle \prod_{i=1}^n e^{i\kappa_i \varphi(x_i)} \right\rangle. \end{aligned}$$

30.1.3 Complex coordinates

A peculiarity of the dimension 2 is that, in real time, massless fields can be decomposed into left and right moving components. In Euclidean space, this corresponds to a description of the real plane (x^1, x^2) in terms of the complex coordinates,

$$z = x^1 + ix^2, \quad \bar{z} = x^1 - ix^2. \quad (30.19)$$

The action then takes the form

$$\mathcal{S}(\varphi) = \int dz d\bar{z} \partial_z \varphi(z, \bar{z}) \partial_{\bar{z}} \varphi(z, \bar{z}). \quad (30.20)$$

The field can be decomposed into the sum of an analytic and an anti-analytic component,

$$\varphi(x^1, x^2) = \varphi_+(z) + \varphi_-(\bar{z}),$$

with the propagators

$$\begin{aligned} \Delta(x^1, x^2) &= \langle \varphi_+(z)\varphi_+(0) \rangle + \langle \varphi_-(\bar{z})\varphi_-(0) \rangle, \\ \langle \varphi_+(z)\varphi_+(0) \rangle &= -\frac{1}{4\pi} \ln z + \text{const.}, \quad \langle \varphi_-(\bar{z})\varphi_-(0) \rangle = -\frac{1}{4\pi} \ln \bar{z} + \text{const.}. \end{aligned}$$

These complex variables are particularly well-suited for exploring the consequences of *conformal symmetry*: the action (30.20) is obviously invariant in the change $z = f(z')$, $\bar{z} = \bar{f}(z')$. They will appear naturally in the discussion of massless fermions.

30.2 The free massless Dirac fermion

The action for a massless Dirac fermion can be written as

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x \bar{\psi}(x) \not{\partial} \psi(x). \quad (30.21)$$

In two dimensions, the three Dirac matrices γ_μ are identical to the three Pauli matrices σ_μ (Section A12.1.4). In equation (30.21), $\not{\partial} \equiv \gamma_1 \partial_1 + \gamma_2 \partial_2$, and $\psi, \bar{\psi}$ are two two-component spinors.

The massless action has two $U(1)$ symmetries corresponding to phase and chiral phase transformations, with angles θ and θ_3 ,

$$\psi(x) = e^{i(\theta_3 \gamma_3 + \theta)} \psi'(x), \quad \bar{\psi}(x) = \bar{\psi}'(x) e^{i(\gamma_3 \theta_3 - \theta)}. \quad (30.22)$$

We define the corresponding currents, generated by space-dependent transformations, by the variation $\delta \mathcal{S}$ of the action as

$$\delta \mathcal{S} = -i \int d^2x [J_\mu(x) \partial_\mu \theta(x) + J_\mu^3(x) \partial_\mu \theta_3(x)].$$

A peculiarity of dimension 2 is that the corresponding vector and axial currents,

$$J_\mu(x) = \bar{\psi}(x) \gamma_\mu \psi(x), \quad J_\mu^3(x) = -\bar{\psi}(x) \gamma_3 \gamma_\mu \psi(x), \quad (30.23)$$

are related, since equation (A12.23) implies (see also equation (30.16)),

$$i\gamma_3 \gamma_\mu = -\epsilon_{\mu\nu} \gamma_\nu \Rightarrow J_\mu^3(x) = -i\epsilon_{\mu\nu} J_\nu(x). \quad (30.24)$$

We can decompose the massless fermion into chiral components $\psi_\pm = \frac{1}{2}(1 \pm \gamma_3)\psi$, $\bar{\psi}_\pm = \frac{1}{2}\bar{\psi}(1 \mp \gamma_3)$ (equation (A12.36)). In the representation in which γ_3 is diagonal, this leads to the representation,

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \quad \bar{\psi} = \begin{pmatrix} \bar{\psi}_- \\ \bar{\psi}_+ \end{pmatrix}. \quad (30.25)$$

With this parametrization, the components of the vector current become,

$$\begin{cases} J_1(x) = \bar{\psi}(x) \gamma_1 \psi(x) = \bar{\psi}_+(x) \psi_+(x) + \bar{\psi}_-(x) \psi_-(x), \\ J_2(x) = \bar{\psi}(x) \gamma_2 \psi(x) = i(\bar{\psi}_+(x) \psi_+(x) - \bar{\psi}_-(x) \psi_-(x)) . \end{cases} \quad (30.26)$$

30.2.1 Fermion correlation functions

The propagator of the free massless fermion is (equation (12.41))

$$[\Delta_\psi]_{\alpha\beta}(x) \equiv \langle \bar{\psi}'_\alpha(x) \psi'_\beta(0) \rangle = \frac{1}{4\pi^2} \int d^2 p e^{-ipx} [i\slashed{p}]_{\beta\alpha}^{-1} = -\frac{1}{2\pi} [\not{x}]_{\beta\alpha}^{-1}. \quad (30.27)$$

In the complex notation (30.19), the action (30.21) takes the form

$$\mathcal{S}(\bar{\psi}, \psi) = - \int dz d\bar{z} [\bar{\psi}_+(z, \bar{z}) \partial_{\bar{z}} \psi_+(z, \bar{z}) + \bar{\psi}_-(z, \bar{z}) \partial_z \psi_-(z, \bar{z})]. \quad (30.28)$$

In the Euclidean formulation, the left and right movers become analytic and anti-analytic fields. The corresponding propagators Δ_ψ^+ and Δ_ψ^- read,

$$\Delta_\psi^+ \equiv \langle \bar{\psi}_+(\bar{z}, z) \psi_+(0, 0) \rangle = -\frac{1}{2\pi z}, \quad (30.29)$$

$$\Delta_\psi^- \equiv \langle \bar{\psi}_-(\bar{z}, z) \psi_-(0, 0) \rangle = -\frac{1}{2\pi \bar{z}}. \quad (30.30)$$

The ψ_\pm correlation functions. Using Wick's theorem for Grassmann variables (equation (1.78)), we can calculate the $2n$ -point correlation functions of $\psi_\pm, \bar{\psi}_\pm$:

$$\left\langle \prod_{i=1}^n \bar{\psi}_+(x_i) \psi_+(x'_i) \right\rangle = \det \Delta_\psi^+(z_i - z'_j) = \left(\frac{-1}{2\pi} \right)^n \det \frac{1}{z_i - z'_j}, \quad (30.31a)$$

$$\left\langle \prod_{i=1}^n \bar{\psi}_-(x_i) \psi_-(x'_i) \right\rangle = \det \Delta_\psi^-(\bar{z}_i - \bar{z}'_j) = \left(\frac{-1}{2\pi} \right)^n \det \frac{1}{\bar{z}_i - \bar{z}'_j}, \quad (30.31b)$$

where here x_i stands for the pair $\{z_i, \bar{z}_i\}$.

Cauchy determinants. The right-hand sides of equations (30.31) involve *Cauchy determinants*. They satisfy a general identity,

$$(-1)^{n(n-1)/2} \det \frac{1}{a_i - b_j} = \frac{\prod_{i < j} (a_i - a_j)(b_i - b_j)}{\prod_{i,j} (a_i - b_j)}, \quad (30.32)$$

which is at the basis of a relation between boson and fermion QFT in two dimensions.

30.2.2 Bilinear operator correlations and boson–fermion correspondence

We now consider the two composite fields

$$\sigma_+(x) = \bar{\psi}_-(x) \psi_+(x), \quad \sigma_-(x) = \bar{\psi}_+(x) \psi_-(x), \quad (30.33)$$

linear combinations of the scalar and pseudoscalar components,

$$\bar{\psi}(x) \psi(x) = \sigma_+(x) + \sigma_-(x), \quad \bar{\psi}(x) \gamma_3 \psi(x) = \sigma_+(x) - \sigma_-(x). \quad (30.34)$$

From the form of the action (30.28) it follows that only correlation functions of a product of an equal number of σ_+ and σ_- fields do not vanish (conservation of chirality). Furthermore, they factorize into the form

$$\left\langle \prod_{i=1}^n \sigma_+(x_i) \sigma_-(x'_i) \right\rangle = (-1)^n \left\langle \prod_{i=1}^n \bar{\psi}_+(x'_i) \psi_+(x_i) \right\rangle \left\langle \prod_{i=1}^n \bar{\psi}_-(x_i) \psi_-(x'_i) \right\rangle, \quad (30.35)$$

where x_i again stands for the pair $\{z_i, \bar{z}_i\}$.

Using the identities (30.31) and (30.32), we infer

$$\left\langle \prod_{i=1}^n \sigma_+(x_i) \sigma_-(x'_i) \right\rangle = \left(\frac{1}{2\pi} \right)^{2n} \frac{\prod_{i < j} |x_i - x_j|^2 |x'_i - x'_j|^2}{\prod_{i,j} |x_i - x'_j|^2}. \quad (30.36)$$

The right-hand side has short-distance divergences associated with the required additive renormalization of the $\sigma_+ \sigma_-$ two-point function.

Comparing with equation (30.12), we discover, for $\kappa^2 = 4\pi$, the identity between the expectation values corresponding to the free massless fermion and boson actions,

$$\left\langle \prod_{i=1}^n \mu^2 e^{i\kappa(\varphi(x_i) - \varphi(x'_i))} \right\rangle_\varphi \Big|_r = (2\pi)^{2n} \left\langle \prod_{i=1}^n \sigma_+(x_i) \sigma_-(x'_i) \right\rangle_\psi.$$

The identity can be translated into the correspondence between quantum operators,

$$2\pi\sigma_\pm(x) \mapsto \mu \left[e^{\pm i\sqrt{4\pi}\varphi(x)} \right]_r = \Lambda e^{\pm i\sqrt{4\pi}\varphi(x)}, \quad (30.37)$$

with the definition (30.3). This remarkable relation between local theories of bosons and fermions is special to two dimensions: in higher dimensions, spin degrees of freedom distinguish between bosons and fermions. In three dimensions, boson–fermion transformations require the presence of gauge fields.

30.2.3 The massive free fermion and the sG model

We consider the partition function

$$\mathcal{Z}_\psi(M_\pm) = \int [d\psi d\bar{\psi}] \exp[-\mathcal{S}(\bar{\psi}, \psi)], \quad (30.38)$$

where the action is

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x [\bar{\psi}(x) \not{\partial} \psi(x) + M_+(x) \sigma_+(x) + M_-(x) \sigma_-(x)]. \quad (30.39)$$

In an expansion of \mathcal{Z}_ψ in powers of the sources M_\pm , the term of degree n in M_\pm is the free-field expectation value, corresponding to the action (30.21),

$$\frac{1}{n!} \left\langle \left[\int d^2x (M_+(x) \sigma_+(x) + M_-(x) \sigma_-(x)) \right]^n \right\rangle. \quad (30.40)$$

Evaluating the expectation values, we obtain the expansion of the partition function,

$$\mathcal{Z}_\psi(M) = \sum_n \frac{1}{(n!)^2} \int \prod_i d^2x_i d^2y_i \frac{M_+(x_i) M_-(y_i)}{(2\pi)^2} \frac{\prod_{i < j} |x_i - x_j|^2 |y_i - y_j|^2}{\prod_{i,j} |x_i - y_j|^2}. \quad (30.41)$$

Alternatively, we can use the correspondence (30.37). The term (30.40) is replaced by

$$\frac{1}{n!} \left(\frac{\mu}{2\pi\zeta} \right)^n \left\langle \left[\int d^2x (M_+(x) e^{i\sqrt{4\pi}\varphi(x)} + M_-(x) e^{-i\sqrt{4\pi}\varphi(x)}) \right]^n \right\rangle_\varphi, \quad (30.42)$$

where ζ is the renormalization constant (30.9) for $\kappa^2 = 4\pi$:

$$\zeta = \mu/\Lambda \Rightarrow \mu/\zeta = \Lambda. \quad (30.43)$$

Summing the expansion in the boson theory, we find that the partition function $\mathcal{Z}_\psi(M)$ is identical to the partition function of a boson field φ corresponding to the action

$$\mathcal{S}(\varphi) = \int d^2x \left\{ \frac{1}{2} [\nabla \varphi(x)]^2 - \frac{\Lambda}{2\pi} [M_+(x) e^{i\sqrt{4\pi}\varphi(x)} + M_-(x) e^{-i\sqrt{4\pi}\varphi(x)}] \right\}. \quad (30.44)$$

In the limit $M_+(x) = M_-(x) = M$ constant, we obtain a remarkable correspondence: the free massive fermion theory is equivalent to the SG model for a special value of the coupling constant. This relation will be further discussed in Section 30.4 and generalized in Section 30.6.

30.3 The gauge-invariant fermion determinant and the anomaly

We now consider the fermion action in an external gauge field B_μ ,

$$\mathcal{S}(\bar{\psi}, \psi, B) = - \int d^2x \bar{\psi}(x) [\partial + iB(x)] \psi(x). \quad (30.45)$$

Because the vector and axial current are related, the classical action (30.45) is invariant not only under phase gauge transformations, but also under chiral gauge transformations:

$$\psi(x) = e^{i\gamma_3\varphi(x)} \psi'(x), \quad \bar{\psi}(x) = \bar{\psi}'(x) e^{i\gamma_3\varphi(x)}. \quad (30.46)$$

Indeed, the transformation generates the term

$$\bar{\psi}'(x) i\partial \varphi(x) \gamma_3 \psi'(x) = \bar{\psi}'(x) \epsilon_{\mu\nu} \gamma_\nu \partial_\mu \varphi(x) \gamma_3 \psi'(x), \quad (30.47)$$

which can be cancelled by the transformation

$$B_\mu(x) = B'_\mu(x) - i\epsilon_{\mu\nu} \partial_\nu \varphi(x). \quad (30.48)$$

The field B_μ is a gauge field for two sets of gauge transformations. Since it has only two components, it can be completely eliminated from the classical action by gauge transformations. Indeed, we parametrize B_μ as

$$B_\mu(x) = -[\partial_\mu \chi(x) + i\epsilon_{\mu\nu} \partial_\nu \varphi(x)]. \quad (30.49)$$

Then,

$$i\epsilon_{\mu\nu} \nabla^2 \varphi(x) = \partial_\mu B_\nu(x) - \partial_\nu B_\mu(x), \quad \text{or} \quad \nabla^2 \varphi(x) = -i\epsilon_{\mu\nu} \partial_\mu B_\nu(x). \quad (30.50)$$

We also perform the gauge transformation,

$$\psi(x) = e^{i(\chi(x)+\gamma_3\varphi(x))} \psi'(x), \quad \bar{\psi}(x) = \bar{\psi}'(x) e^{i(-\chi(x)+\gamma_3\varphi(x))}. \quad (30.51)$$

The action (30.45) then reduces to the free-field fermion action,

$$\mathcal{S}(\bar{\psi}, \psi, B) = - \int d^2x \bar{\psi}'(x) \partial \psi'(x). \quad (30.52)$$

Quantum theory: The anomaly. As we have more generally shown in Section 23.6, the chiral gauge symmetry is broken at the quantum level by the chiral anomaly. Here, we recover the form of the anomaly by a direct calculation. In Section A30.2.1, we verify the consistency of the result we will obtain with the general form derived in Section 23.6 (equation (23.92)).

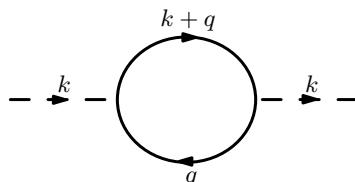


Fig. 30.1 The one-loop contribution to the current two-point function

The integral over fermion fields yields a determinant, which has an expansion in terms of B_μ , and generates one-loop diagrams. In two dimensions, only the term quadratic in B_μ , corresponding to the diagram displayed in Fig. 30.1, is divergent.

We regularize the determinant by introducing a massive spinor boson ϕ as a regulator field:

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi, B_\mu, \phi) \\ = - \int d^2x [\bar{\psi}(x)(\not{\partial} + i\not{B}(x))\psi(x) - \bar{\phi}(x)(\not{\partial} + i\not{B}(x) + M)\phi(x)]. \end{aligned} \quad (30.53)$$

With this modification, the theory is finite, but the ϕ mass breaks chiral gauge invariance explicitly (but not global chiral symmetry). We parametrize the two-component gauge field in terms of two scalar fields as in equation (30.49), and perform the corresponding gauge transformations (30.51) both on $\bar{\psi}, \psi$ and $\bar{\phi}, \phi$. The action becomes

$$\mathcal{S}(\bar{\psi}', \psi', B_\mu, \phi') = - \int d^2x [\bar{\psi}'(x)\not{\partial}\psi'(x) - \bar{\phi}'(x)(\not{\partial} + M e^{2i\gamma_3\varphi(x)})\phi'(x)]. \quad (30.54)$$

An integration over the fields $\bar{\phi}', \phi'$ yields the inverse determinant

$$\mathcal{D}^{-1}(\varphi) = \det(\not{\partial} + M e^{2i\gamma_3\varphi(x)}), \quad (30.55)$$

which has to be evaluated for M large. It can be expressed as the fermion field integral:

$$\frac{1}{\mathcal{D}} = \int [d\psi d\bar{\psi}] \exp \int d^2x [\bar{\psi}(x)\not{\partial}\psi(x) + M e^{2i\varphi(x)}\sigma_+(x) + M e^{-2i\varphi(x)}\sigma_-(x)]. \quad (30.56)$$

We now use the *bosonization* technique. In Section 30.2.3, we have derived a fermion–boson equivalence. We can thus replace the fermion expectation value with the action (30.39) by a boson expectation value with the action (30.44) and $M_\pm = e^{\pm 2i\varphi} M$. Denoting the sG field (Section 30.4) by $\vartheta(x)$, we find a modified sG model (30.44) with ϑ replaced by $\vartheta + \varphi/\sqrt{\pi}$ in the interaction:

$$\mathcal{S}(\vartheta, \varphi) = \int d^2x \left\{ \frac{1}{2} [\nabla\vartheta(x)]^2 - \frac{M\Lambda}{\pi} \cos(\sqrt{4\pi}\vartheta(x) + 2\varphi(x)) \right\}. \quad (30.57)$$

We now change field variables, with $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. The action becomes

$$\mathcal{S}'(\vartheta) = \int d^2x \left\{ \frac{1}{2} [\nabla\vartheta(x) - \nabla\varphi(x)/\sqrt{\pi}]^2 - \frac{M\Lambda}{\pi} \cos \sqrt{4\pi}\vartheta(x) \right\}. \quad (30.58)$$

In the large M limit, the ϑ -field becomes very massive and has thus vanishing fluctuations around $\vartheta(x) = 0$. At leading order, we can set $\vartheta = 0$, and thus obtain the finite result,

$$\mathcal{D}^{-1}(\varphi) \propto \exp \left[-\frac{1}{2\pi} \int d^2x (\nabla\varphi(x))^2 \right]. \quad (30.59)$$

Since $\mathcal{D}(\varphi)$ does not go to a constant for M large, the chiral gauge symmetry is broken by quantum fluctuations: this is the simplest example of a chiral anomaly.

The action (30.45) is thus equivalent to

$$\mathcal{S}(\bar{\psi}', \psi', B_\mu) = - \int d^2x \left[\bar{\psi}'(x)\not{\partial}\psi'(x) + \frac{1}{2\pi} \int d^2x (\nabla\varphi(x))^2 \right]. \quad (30.60)$$

30.3.1 Current correlation functions

Since the expressions (30.31) are singular when $x_i \rightarrow x'_i$, they cannot be used to define the current without a short-distance regularization. We can avoid the problem by noting that, in the action (30.45), the gauge field acts as a source for the current: differentiating with respect to B_μ yields iJ_μ .

Current two-point function. From the result (30.59) expanded at second order in φ and the relation (30.50), we obtain the current two-point function in momentum and space representations:

$$\langle \tilde{J}_\mu(k) \tilde{J}_\nu(-k) \rangle = \frac{1}{\pi} \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right), \quad \langle J_\mu(x) J_\nu(0) \rangle = -\frac{1}{2\pi^2 x^2} \left(\delta_{\mu\nu} - 2 \frac{x_\mu x_\nu}{x^2} \right),$$

expressions proportional to the two-point functions (30.17b) and (30.18b) of the boson current. The second expression is a distribution singular at $x = 0$, where it is ambiguous.

Then, using equation (30.24), we obtain the axial current two-point function

$$\langle \tilde{J}_\mu^3(k) \tilde{J}_\nu^3(-k) \rangle = \frac{1}{\pi} \frac{k_\mu k_\nu}{k^2},$$

a result identical to (30.17a). In the space variables, both currents are proportional but their different Fourier representations show that only the vector current is conserved.

Current insertion. We now calculate σ_\pm correlation functions with one current $J_\mu(x)$ insertion. In the transformation (30.51), the operators $\sigma_\pm(x)$ become $\sigma_\pm(x) e^{\pm 2i\varphi(x)}$. To obtain the current insertion, we expand at first order in φ . This yields a factor

$$2\epsilon_{\mu\nu} \int d^2x \partial_\mu B_\nu(x) \sum_i [\Delta(x - x_i) - \Delta(x - x'_i)].$$

Finally, differentiating with respect to $B_\mu(x)$, we obtain

$$\begin{aligned} \left\langle J_\mu(x) \prod_{i=1}^n \sigma_+(x_i) \sigma_-(x'_i) \right\rangle &= \left\langle \prod_{i=1}^n \sigma_+(x_i) \sigma_-(x'_i) \right\rangle \\ &\times \sum_i \frac{i}{\pi} \epsilon_{\mu\nu} \left(\frac{(x - x_i)_\nu}{(x - x_i)^2} - \frac{(x - x'_i)_\nu}{(x - x'_i)^2} \right). \end{aligned} \quad (30.61)$$

The expression is proportional to the boson current insertion $J_\mu^A(x)$ for $\kappa_i = \pm\sqrt{4\pi}$. In the same way, the axial current J_μ^A is proportional to J_μ^3 :

$$J_\mu^A(x) \mapsto \sqrt{\pi} J_\mu(x), \quad J_\mu^V(x) \mapsto \sqrt{\pi} J_\mu^3(x). \quad (30.62)$$

The second result is not surprising: a translation of φ , $\varphi \mapsto \varphi + \theta$ multiplies $e^{\pm i\sqrt{4\pi}\varphi}$ by $e^{\pm i\sqrt{4\pi}\theta}$. To J_μ^3 is associated the transformation $\sigma_\pm \mapsto \sigma_\pm e^{\pm 2i\theta}$.

30.4 The sG model

The sG model is a QFT for a scalar boson field $\vartheta(x)$ defined by the action

$$\mathcal{S}(\vartheta) = \int d^2x \left\{ \frac{1}{2} [\nabla \vartheta(x)]^2 - \frac{\alpha_0}{\kappa^2} \cos(\kappa \vartheta(x)) \right\}. \quad (30.63)$$

One can choose $\alpha_0 > 0$ without loss of generality. Depending on the issue one wants to investigate, one sometimes normalizes the sG field differently, setting $\theta = \kappa \vartheta$ and thus,

$$\mathcal{S}(\theta) = \frac{1}{\tau} \int d^2x \left\{ \frac{1}{2} [\nabla \theta(x)]^2 - \alpha_0 \cos \theta(x) \right\}, \quad \text{with } \tau = \kappa^2. \quad (30.64)$$

The model has been extensively studied. The classical field equations are completely integrable. This makes it possible to obtain *solitons*, that is, finite energy solutions of the real-time equations of motion, and to infer the semi-classical spectrum [320, 321]. This integrability survives quantization and thus, for example, exact expressions for the spectrum and the S -matrix can be obtained, confirming the semi-classical analysis. However, the relevant techniques are outside the scope of this work. We simply describe the model, and derive some of its algebraic and RG properties.

30.4.1 Perturbative expansion

For τ small, the field θ fluctuates around one of the minima $\theta = 2n\pi$ or $-\cos \theta$. We can choose one of them, for example $\theta = 0$, to expand perturbation theory, because they are all equivalent and are connected by tunnelling (see Section 39.4). However, the degeneracy of the classical minimum is responsible for the presence of solitons. An expansion of $\cos \theta$ in powers of θ shows that the θ -field is massive, with a mass $\alpha^{1/2} + O(\tau)$.

The model is super-renormalizable. From the discussion of Section A9.2, we know that divergences, in the perturbative expansion in powers of τ arise only from the self-contractions of the interaction term $\cos \theta$. These divergences have been calculated in Section 30.1. Using equation (30.9), we can set (μ is the renormalization scale)

$$\begin{aligned} e^{\pm i\theta(x)} &= e^{\pm i\kappa\vartheta(x)} = Z_\theta^{1/2} \left[e^{\pm i\theta(x)} \right]_r, \\ \alpha_0 &= Z_\theta^{-1/2} \alpha, \quad \text{with } Z_\theta = (\Lambda/\mu)^{-\tau/(2\pi)}. \end{aligned} \quad (30.65)$$

30.4.2 RG equations

Since the theory is super-renormalizable, the RG β -function vanishes. The field $e^{\pm i\theta(x)}$ and the coupling constant α RG functions are given in terms of Z_θ . One finds

$$\eta(\tau = \kappa^2) = \mu \frac{\partial}{\partial \mu} \ln Z_\theta(\tau, \Lambda/\mu) = \frac{\tau}{2\pi}. \quad (30.66)$$

The renormalized n -point correlation function $\tilde{W}^{(n)}$ of $e^{\pm i\theta(x)}$, in the Fourier representation, satisfies the RG equation,

$$\left[\mu \frac{\partial}{\partial \mu} + \frac{n}{2} \eta(\tau) + \frac{1}{2} \eta(\tau) \alpha \frac{\partial}{\partial \alpha} \right] \tilde{W}^{(n)}(p_i, \tau, \alpha) = 0. \quad (30.67)$$

Combining its solution with the dimensional relation

$$\mu \frac{\partial}{\partial \mu} + 2\alpha \frac{\partial}{\partial \alpha} + p_i \frac{\partial}{\partial p_i} = 2(1-n),$$

one infers the scaling relation,

$$\tilde{W}^{(n)}(p_i, \tau, \alpha) = \alpha^{\tau/(8\pi-\tau)} \alpha^{1-n} F^{(n)}(p_i \alpha^{-4\pi/(8\pi-\tau)}, \tau). \quad (30.68)$$

In particular, for the one-point function and the mass scale or θ -mass, one obtains

$$\langle e^{\pm i\theta(x)} \rangle = W^{(1)} \sim \alpha^{\tau/(8\pi-\tau)}, \quad m_\theta \sim \alpha^{4\pi/(8\pi-\tau)}. \quad (30.69)$$

In these relations, two values of $\tau = \kappa^2$ play a special role:

(i) For $\tau = 4\pi$, the propagator (30.13) becomes singular at short distance, in the sense of distributions (or UV divergent in momentum space). Therefore, the term of order α^2 in the expansion of the sG partition function requires a new additive renormalization. This yields an additive renormalization of the free energy proportional to α^2 .

Also the mass scale m_θ and $\langle e^{i\theta} \rangle$ become linear in α , results which have simple interpretations in terms of the correspondence (30.44) with a free massive fermion model. Indeed, $m_\theta \propto m_\psi = \alpha$ and

$$\langle e^{i\theta(x)} \rangle \propto \langle \bar{\psi}(x)\psi(x) \rangle = \frac{1}{2\pi} \text{tr} \int d^2 p \frac{1}{ip + \alpha} \sim 2\alpha \ln(\Lambda/\alpha).$$

The UV divergence is of the same nature as the divergence in Fig. 30.1 (see also equation (30.3)), since by differentiating with respect to α , one obtains a two-point function at zero-momentum. It leads to a logarithmic correction to the linear behaviour. We further discuss this relation in Section 30.6. For $\tau > 4\pi$, generically the leading contribution to $W^{(1)}$ comes from short-distance effects, and remains linear in α .

(ii) For $\tau = 8\pi$, the quantities $\langle e^{i\theta} \rangle$ and m_θ in equations (30.69) vanish identically for α small. We can find an interpretation of the result by calculating the dimension of the operator $\cos l\theta(x)$ and more generally of $\cos l\theta(x)$:

$$[\cos l\theta] = \tau l^2 / 4\pi. \quad (30.70)$$

Therefore, these operators, which are relevant for τ small and give a mass to the θ -field, become irrelevant beyond a finite value of τ given by

$$\tau l^2 / 4\pi = 2 \implies \tau = 8\pi/l^2. \quad (30.71)$$

For $\tau > 8\pi$, all operators $\cos l\theta$ are irrelevant, which explains why no mass is generated. At $\tau = 8\pi$, the interaction $\int d^2 x \cos \theta(x)$ is marginal. In the field theory terminology, this means that the theory is just renormalizable; for $\tau > 8\pi$ the theory is no longer renormalizable. We discuss this peculiar transition in Chapter 31, in more detail. However, the analysis is valid only at leading order in α , since at $\tau = 8\pi$ new divergences are generated at each order in α , and thus the dimension of operators may be modified.

30.5 The Schwinger model

We now consider two-dimensional QED, with one charged Dirac fermion coupled to an Abelian gauge field. With the help of the bosonization technique, the massless model can be solved exactly, and some properties of the massive model can be discovered.

30.5.1 The massless model

The action reads

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int d^2x \left[\frac{1}{4} F_{\mu\nu}^2(x) - \bar{\psi}(x)(\partial + ieA(x))\psi(x) \right], \quad (30.72)$$

where e is the charge. The model, first discussed by Schwinger, exhibits the simplest example of a *chiral anomaly*, illustrates both *confinement* (see Chapter 25) and *spontaneous chiral symmetry breaking* in two dimensions.

In a linear covariant gauge, the field theory is super-renormalizable by power counting. The only divergent diagram corresponds to the one-loop contribution to the gauge field two-point function (Fig. 30.1), and comes from the expansion of the fermion determinant which we have already discussed in Section 30.3. In Section 30.3, we have shown that the fermion part of the action with $B_\mu = eA_\mu$, after the gauge transformations (30.51), and with the parametrization (30.49), is equivalent to the free field action (30.60). We have here simply to add the contribution coming from $F_{\mu\nu}^2$:

$$\frac{1}{4} F_{\mu\nu}^2(x) = \frac{1}{2e^2} \nabla\varphi(x) \cdot \nabla\nabla^2\varphi(x).$$

In terms of the fields $\bar{\psi}', \psi', \varphi$, the Schwinger model action then takes the free-field form (in the gauge $\partial_\mu A_\mu = \nabla^2\chi/e = 0$)

$$\mathcal{S}(\bar{\psi}', \psi', \varphi) = \int d^2x \left[-\bar{\psi}'(x)\partial\psi'(x) - \frac{1}{2\pi} (\nabla\varphi(x))^2 + \frac{1}{2e^2} \nabla\varphi(x) \nabla\nabla^2\varphi(x) \right]. \quad (30.73)$$

We infer the φ -field propagator in the Fourier representation,

$$\tilde{\Delta}_\varphi(p) = \pi \left(\frac{1}{p^2 + e^2/\pi} - \frac{1}{p^2} \right). \quad (30.74)$$

This expression shows that the boson φ -field propagates two neutral particles, one with a positive residue and a mass

$$m = e/\sqrt{\pi}, \quad (30.75)$$

and another one, massless and non-physical, since with a negative residue. The appearance of a non-vanishing mass is a direct consequence of the chiral anomaly.

Finally, using the representation (30.49), we can calculate the gauge field transverse two-point function. In the Fourier representation, we find

$$\tilde{\Delta}_{\mu\nu}^{(2)}(p) = (\delta_{\mu\nu} - p_\mu p_\nu/p^2) \frac{1}{p^2 + m^2}, \quad (30.76)$$

a result that we verify in Section A30.2.2 by a direct one-loop calculation.

We also bosonize the free fermion $\bar{\psi}', \psi'$, introducing a free massless boson ϑ . The action becomes

$$\mathcal{S}(\vartheta, \varphi) = \int d^2x \left[\frac{1}{2} (\nabla\vartheta(x))^2 - \frac{1}{2\pi} (\nabla\varphi(x))^2 + \frac{1}{2\pi m^2} \nabla\varphi(x) \nabla^2\nabla\varphi(x) \right]. \quad (30.77)$$

We translate ϑ : $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. The action takes the form

$$\mathcal{S}(\vartheta, \varphi) = \int d^2x \left[\frac{1}{2} (\nabla\vartheta(x))^2 - \frac{1}{\sqrt{\pi}} \nabla\varphi(x) \nabla\vartheta(x) + \frac{1}{2\pi m^2} \nabla\varphi(x) \nabla^2\nabla\varphi(x) \right].$$

Finally, we integrate over φ and obtain

$$\mathcal{S}(\vartheta) = \frac{1}{2} \int d^2x \left[(\nabla \vartheta(x))^2 + m^2 \vartheta^2(x) \right]. \quad (30.78)$$

It is remarkable that the fermion contribution just cancels the massless boson field, leaving only the massive boson. Starting from the representation (30.51), it is easy to verify that these transformations lead to the mapping

$$\sigma_{\pm}(x) = \bar{\psi}_{\mp}(x) \psi_{\pm}(x) \mapsto \frac{\Lambda}{2\pi} e^{\pm i \sqrt{4\pi} \vartheta(x)} \quad (30.79)$$

of the chiral components σ_{\pm} of the neutral composite field $\bar{\psi}\psi$.

Remark. Before integration over φ , the action is formally invariant when one translates ϑ by a constant. Therefore, the form (30.78) corresponds to a choice of boundary conditions for the inverse of the Laplace operator. A different choice leads to the replacement $\vartheta(x) \mapsto \vartheta(x) - \vartheta_{\infty}$. From the correspondence (30.79), we see that, in general, space reflection symmetry is then broken, except for $\vartheta_{\infty} = 0 \pmod{\sqrt{\pi}/2}$. Finally, this modification formally corresponds to adding a topological term proportional to $\nabla^2 \varphi(x) \propto \epsilon_{\mu\nu} F_{\mu\nu}(x)$ to the action density, that is, a constant electric field. For a discussion of the physical effects of such a modification, we refer to the literature.

30.5.2 Confinement and chiral symmetry breaking

We note that the gauge field (30.76) two-point function has no cut corresponding to fermion intermediate states. This property is a sign of *confinement* [325], because it implies that no charged particles are emitted by the neutral particle: the electromagnetic forces between particles of opposite charge are strong enough to prevent the separation of charged particles: these cannot be observed as free particles (for details see Chapter 25).

This observation is confirmed by the structure of $\sigma_{\pm}(x)$ correlation functions. Indeed, they can be calculated with the massive free action (30.78), where no massless field appears.

We now calculate the expectation values of σ_{\pm} . From the correspondence (30.79) and equation (30.7), we obtain

$$\langle \sigma_{\pm} \rangle = \frac{\Lambda}{2\pi} \langle e^{\pm i \sqrt{4\pi} \vartheta(x)} \rangle = \frac{\Lambda}{2\pi} e^{-2\pi \Delta(0,m)}.$$

Taking into account the definition (30.3), we find the finite result

$$\langle \sigma_{\pm} \rangle = \frac{e^{\gamma}}{4\pi} m \quad \Rightarrow \quad \langle \bar{\psi}(x) \psi(x) \rangle = \frac{e^{\gamma}}{2\pi} m. \quad (30.80)$$

Since $\bar{\psi}\psi$ is a composite field that is not chiral invariant, the non-vanishing result shows that *global chiral symmetry is spontaneously broken*. Note that spontaneous symmetry breaking with order is possible in this model, because the electromagnetic interaction generates long-range forces.

From equation (30.7), we infer the two-point functions of σ_{\pm} :

$$\langle \sigma_{\epsilon_1}(x_1) \sigma_{\epsilon_2}(x_2) \rangle = \langle \sigma \rangle^2 e^{-4\pi \epsilon_1 \epsilon_2 \Delta(x_1 - x_2, m)}. \quad (30.81)$$

As a consequence,

$$\langle \bar{\psi}(x)\psi(x)\bar{\psi}(0)\psi(0) \rangle = \langle \bar{\psi}\psi \rangle^2 \cosh(4\pi\Delta(x, m)), \quad (30.82a)$$

$$\langle \bar{\psi}(x)\gamma_3\psi(x)\bar{\psi}(0)\gamma_3\psi(0) \rangle = \langle \bar{\psi}\psi \rangle^2 \sinh(4\pi\Delta(x, m)). \quad (30.82b)$$

These expressions have several remarkable properties: the two-point functions have only singularities associated with the massive field. If we expand the exponentials in powers of the propagator and Fourier transform, we find that the expression (30.82b) has a pole at $k^2 = -m^2$ and cuts at $k^2 = -(2n-1)^2 m^2$, $n > 1$, in momentum space, while the expression (30.82a) has cuts at $k^2 = -(2n)^2 m^2$. Only the massive neutral boson appears in the intermediate states, but no charged fermions (the confinement property), and, moreover, the boson is a pseudoscalar, since it appears as a simple pole only in the $\bar{\psi}\gamma_3\psi$ two-point function.

Finally, the short-distance behaviour of the propagator $\Delta(m, x)$ is given by the asymptotic form (30.4), since it is a function only of mx . Therefore, the two-point functions $\langle \sigma_+(x)\sigma_+(0) \rangle$ and $\langle \sigma_-(x)\sigma_-(0) \rangle$ go to 0 as x^2 , while

$$\langle \sigma_+(x)\sigma_-(0) \rangle \sim \frac{1}{4\pi^2 x^2},$$

that is, like in a free massless fermion theory. This property reflects the *asymptotic freedom* (at large momentum) of super-renormalizable theories.

These are all properties we also expect in the true physical world with quarks and gluons (Section 23.4.1).

30.5.3 The massive Schwinger model

Let us now briefly discuss the effect of the addition of a fermion mass term. Chiral symmetry is no longer an issue, since the mass term explicitly breaks chiral symmetry, but there is still the question of confinement: will charged particles appear in the spectrum?

We consider the action

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int d^2x \left[\frac{1}{4} F_{\mu\nu}^2(x) - \bar{\psi}(x) (\not{D} + ie\mathcal{A}(x) + M) \psi(x) \right]. \quad (30.83)$$

We perform the transformations (30.46) and (30.49) ($B_\mu = eA_\mu$) and obtain

$$\begin{aligned} \mathcal{S}(\bar{\psi}', \psi', \varphi) = & - \int d^2x \left[\bar{\psi}'(x) \left(\not{D} + M e^{2i\gamma_3\varphi(x)} \right) \psi'(x) \right. \\ & \left. + \nabla\varphi(x) (m^2 - \nabla^2) \nabla\varphi(x) / 2e^2 \right]. \end{aligned} \quad (30.84)$$

The model is no longer free and cannot be solved exactly. However, the action can be further transformed, as shown in Section 30.2.3. We call ϑ the boson field associated by the bosonization with $\bar{\psi}', \psi'$. We then use the equivalence between the actions (30.39) and (30.44). The fermion action is replaced by the action (30.57), and thus,

$$\begin{aligned} \mathcal{S}(\vartheta, \varphi) = & \int d^2x \left\{ \frac{1}{2} [\nabla\vartheta(x)]^2 - \frac{M\Lambda}{\pi} \cos \left(\sqrt{4\pi}\vartheta(x) + 2\varphi(x) \right) \right. \\ & \left. - \frac{1}{2e^2} \nabla\varphi(x) (m^2 - \nabla^2) \nabla\varphi(x) \right\}. \end{aligned}$$

We now translate ϑ . We change $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. After translation, the action becomes quadratic in φ . We can thus integrate over φ , and find the ‘massive’ sG action for ϑ ,

$$\mathcal{S}(\vartheta) = \int d^2x \left\{ \frac{1}{2} [\nabla \vartheta(x)]^2 + \frac{1}{2} m^2 \vartheta^2(x) - \alpha_0 \cos(\sqrt{4\pi}\vartheta(x)) \right\}, \quad (30.85)$$

with the correspondence (30.79) and

$$\alpha_0 = M\Lambda/\pi.$$

Physics interpretation. In the action (30.85), we note that, at least for M small, the result obtained for $M = 0$ survives: the spectrum consists in a massive neutral boson of mass squared $m^2 + O(mM)$. No charged particles appear in the spectrum. On the other hand, for M large, we expect a non-relativistic analysis to be valid. Then, we have a set of fermions interacting through a one-dimensional Coulomb potential. At large distance, the one-dimensional Coulomb potential raises linearly with the distance between charged particles, and thus charged particles can never be separated [326].

30.6 The massive Thirring model

In Section 30.2, comparing equations (30.11) and (30.36), we have discovered identities between free massless fermion expectation values of powers of $\bar{\psi}\psi$ and expectation values of exponentials of free massless boson fields. This has led to the correspondence derived in Section 30.2.3 between the free massive fermion and the sG model for $\tau = \kappa^2 = 4\pi$. The same identities make it possible to establish a relation between the general quantum sG and the massive Thirring model [323]. The Thirring model is described in terms of a Dirac fermion with the action:

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x [\bar{\psi}(x) (\not{\partial} + m_0) \psi(x) - \frac{1}{2} g J_\mu(x) J_\mu(x)], \quad (30.86)$$

where

$$J_\mu(x) = \bar{\psi}(x) \gamma_\mu \psi(x). \quad (30.87)$$

For $m_0 = 0$, the Thirring model retains the $U(1)$ symmetries (30.22) of the free massless action. Note that the interaction is the only local interaction possible, since the model involves only four fermion variables. Finally, power counting shows that the model is renormalizable in two dimensions, and the coupling constant g thus is dimensionless.

The massive Thirring model can be mapped onto the general sG model. The correspondence between the two models can be summarized by the relations:

$$1 + g/\pi = 4\pi/\kappa^2, \quad (30.88a)$$

$$J_\mu(x) \mapsto \frac{1}{2\pi} \epsilon_{\mu\nu} \partial_\nu \theta(x), \quad (30.88b)$$

$$\bar{\psi}\psi(x) \mapsto \frac{\Lambda}{\pi} \cos \theta(x). \quad (30.88c)$$

The chiral-invariant model, $m_0 = 0$, is mapped onto a free boson theory, as we have already discussed in Section 30.4, and is thus exactly soluble.

We will derive this correspondence, using the bosonization identities established in previous sections.

30.6.1 Bosonization: The sG model

The first step of the bosonization is to introduce a vector field A_μ and rewrite the interaction term as resulting from a Gaussian integral over A_μ :

$$A_\mu^2(x)/2g + iA_\mu(x)J_\mu(x) \mapsto \frac{1}{2}gJ_\mu(x)J_\mu(x).$$

The fermion action becomes a quadratic action of charged fermions coupled to an Abelian gauge field:

$$\mathcal{S}(A_\mu, \bar{\psi}, \psi) = - \int d^2x [\bar{\psi}(x)(\partial + i\mathcal{A}(x) + m_0)\psi(x) - A_\mu^2(x)/2g]. \quad (30.89)$$

Bosonization. The action then differs from the massive Schwinger model action (30.83) only by the gauge-field kinetic term. It follows from the arguments given in Section 30.2.3 that the action (30.89) can be replaced by a scalar boson action. Parametrizing the two-component field A_μ in terms of two scalar fields χ and φ , as in equation (30.49), one finds,

$$\int d^2x A_\mu^2(x) = \int d^2x [(\nabla\chi(x))^2 - (\nabla\varphi(x))^2].$$

After the gauge transformations (30.51), the action (30.89) becomes

$$\begin{aligned} \mathcal{S}(\chi, \varphi, \bar{\psi}', \psi') = & - \int d^2x \bar{\psi}'(x) (\partial + m_0 e^{2i\gamma_3\varphi(x)}) \psi'(x) \\ & + \frac{1}{2} \int d^2x \left[\frac{1}{g} (\nabla\chi(x))^2 - \left(\frac{1}{g} + \frac{1}{\pi} \right) (\nabla\varphi(x))^2 \right], \end{aligned} \quad (30.90)$$

where we have taken into account the anomaly (30.59).

The field χ decouples from neutral correlation functions. The kinetic term of the field φ has the wrong sign. Expectation values of products of $e^{\pm i\varphi}$ are IR divergent when the numbers of + and - signs are different. Fortunately, they always appear multiplied by fermion expectation values which vanish due to chiral symmetry, as we have indicated in Section 30.2.2. .

The correspondence between the massive Thirring and sG models then follows directly from the analysis of Section 30.2.3. We have shown that the fermion part of the action can be replaced by a sG action for $\kappa^2 = \tau = 4\pi$ of the form (30.57):

$$\mathcal{S}(\bar{\psi}', \psi') \mapsto \int d^2x \left[\frac{1}{2} (\nabla\vartheta(x))^2 - \frac{m_0\Lambda}{\pi} \cos(\sqrt{4\pi}\vartheta(x) + 2\varphi(x)) \right]. \quad (30.91)$$

We then change field variables $\varphi \mapsto \theta$, setting

$$\sqrt{4\pi}\vartheta + 2\varphi = \theta.$$

The action becomes

$$\mathcal{S}(\vartheta, \theta) = \frac{1}{2} \int d^2x \left[(\nabla\vartheta(x))^2 - \frac{2m_0\Lambda}{\pi} \cos\theta(x) - \frac{\pi+g}{4\pi g} (\nabla\theta(x) - \sqrt{4\pi}\nabla\vartheta(x))^2 \right].$$

The integral over ϑ then is Gaussian and can be explicitly performed. One obtains the action of the sG model,

$$\mathcal{S}(\theta) = \int d^2x \left[\frac{1}{8\pi} \left(1 + \frac{g}{\pi} \right) (\nabla\theta(x))^2 - \frac{m_0\Lambda}{\pi} \cos\theta(x) \right]. \quad (30.92)$$

This establishes a correspondence between the massive Thirring model and the sG model. The relation with the coupling constant appearing in the action (30.64) is ($\tau = \kappa^2$)

$$\frac{4\pi}{\tau} = 1 + \frac{g}{\pi},$$

proving relations (30.88a, c). Since the coupling constant τ is not renormalized, the correspondence establishes that the coupling constant g is not renormalized either and, thus, the RG $\beta(g)$ -function vanishes.

By contrast, since the field $e^{\pm 2i\varphi}$ has to be renormalized, the fermion mass also needs a renormalization. With the regularization and renormalization at mass scale μ of Section 30.1, one finds

$$m_0 = m(\mu/\Lambda)^{g/(g+\pi)}.$$

The relation between renormalized parameters is then

$$\alpha/\tau = m\mu/\pi.$$

30.6.2 The massless Thirring model

We now briefly discuss the special massless case. In the case of the massless Thirring model, $m_0 = 0$, one obtains the equivalent free boson theory, with the action

$$\mathcal{S}(\vartheta, \varphi, \chi) = \frac{1}{2} \int d^2x \left[(\nabla \vartheta(x))^2 + \frac{1}{g} (\nabla \chi(x))^2 - \left(\frac{1}{g} + \frac{1}{\pi} \right) (\nabla \varphi(x))^2 \right]. \quad (30.93)$$

The massless Thirring model can thus be solved exactly. All ψ -field correlation functions can be calculated explicitly. In particular, the bosonized form of the $\bar{\psi}(x)\psi(x)$ correlation functions only depends on $\theta(x)$, and thus can be calculated with the action (30.92) for $m_0 = 0$. The one-point function $\bar{\psi}(x)\psi(x)$ vanishes. The $\bar{\psi}\psi$ two-point function is given by

$$\langle \bar{\psi}(x)\psi(x)\bar{\psi}(0)\psi(0) \rangle \propto \langle \cos \theta(x) \cos \theta(0) \rangle \propto x^{-\lambda}$$

with $\lambda = \tau/2\pi = 2\pi/(g + \pi)$.

Current two-point function. The massless Thirring model has a $U(1)$ chiral symmetry. After adding a source term B_μ for the current J_μ , one finds the non-local action for the remaining A_μ field,

$$\mathcal{S}(A_\mu) = \frac{1}{2} \int d^2x \left[\frac{1}{g} A_\mu^2(x) + \frac{1}{\pi} (A_\mu(x) - iB_\mu(x)) \left(\delta_{\mu\nu} - \frac{\partial_\mu \partial_\nu}{\nabla^2} \right) (A_\nu(x) - iB_\nu(x)) \right].$$

After integration over the field A_μ , one reads off the current two-point correlation function. In the Fourier representation,

$$\langle \tilde{J}_\mu(k) \tilde{J}_\nu(-k) \rangle = \frac{1}{\pi + g} \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right). \quad (30.94)$$

All other connected correlation functions vanish. Alternatively, if one calculates the two-point correlation function of the field $K_\mu = \epsilon_{\mu\nu} \partial_\nu \theta$ with the free action (30.92) for $m_0 = 0$, one finds

$$\langle \tilde{K}_\mu(k) \tilde{K}_\nu(-k) \rangle = \frac{4\pi^2}{\pi + g} \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right), \quad (30.95)$$

a result consistent with the equivalence relations (30.88a, b).

30.6.3 RG properties. Mass spectrum

RG properties. The massless Thirring model exhibits a scaling behaviour for all values of the coupling constant g . This is a consequence of the vanishing of the $\beta(g)$ -function: this model provides an example of a *line of IR fixed points*. Since the $\bar{\psi}\psi$ two-point function decreases at large distance, chiral symmetry is not spontaneously broken, although the correlation function only decreases algebraically.

As we have discussed in the case of the sG model, at the value $\kappa^2 = 8\pi$, that is, $g = \pi/2$, the mass term becomes irrelevant. This suggests the possibility of a phase transition, a question we examine in Chapter 31. However, the transition point is neither in the perturbative domain of the Thirring model nor of the sG model. The properties of the bosonization method suggest that it should be possible to map the sG model with $\kappa^2 = 8\pi$ onto a free model by doubling the number of fermions. Indeed, we exhibit in the next section a two-fermion model that is also equivalent to the sG model.

Mass spectrum. The correspondence between the Thirring and sG models makes it possible, in particular, to calculate physical quantities in the sG model for g small, that is, for τ close to 4π , or in the Thirring model for g large. Moreover, we see that for $g > 0$, that is, $\tau < 4\pi$, where the potential between fermions is attractive, the spectrum of the theory consists at least in one Dirac fermion and one boson bound state corresponding to the field of the sG model (the fermion appears semi-classically in the sG model as a soliton). Actually, the exact bound-state mass spectrum is given by

$$m_n = 2m(\alpha, \tau) \sin(n\gamma/16), \quad n = 1, 2, \dots < 8\pi/\gamma, \\ \text{with } \gamma = \frac{\tau}{1 - \tau/8\pi} = \frac{8\pi}{1 + 2g/\pi},$$

where $m(\alpha, \tau)$ is the fermion mass (see also equation (31.20)), which is of the form,

$$m(\alpha, \tau) = f(\tau) \frac{8}{\gamma} \alpha^{1/2 + \gamma/16\pi}, \quad f(0) = 1. \quad (30.96)$$

The lowest mass $n = 1$ corresponds to the sG field; the remaining bound states can also be considered as collective excitations of the sG particle.

30.7 A generalized Thirring model with two fermions

The derivation of the correspondence between the massive Thirring and the sG model suggests that it should be possible to construct a fermion model equivalent to the sG model, such that $\tau = 8\pi$ corresponds to a free-field theory, by introducing two Dirac fermions and by choosing the four-point interaction appropriately.

30.7.1 The model

We consider the action for a generalized massless Thirring model with two fermions $\psi^a, \bar{\psi}^a$, $a = 1, 2$, coupled through current-current interactions,

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x \left[\bar{\psi}(x) \cdot \not{\partial} \psi(x) - \frac{1}{4} (g_0 J_\mu^0(x) J_\mu^0(x) + g_3 J_\mu^3(x) J_\mu^3(x)) \right. \\ \left. - \frac{1}{4} f \sum_i J_\mu^i(x) J_\mu^i(x) \right], \quad (30.97)$$

where J_μ^0 and J_μ^i are $U(1)$ and $SU(2)$ currents, respectively, defined by,

$$J_\mu^0(x) = \sum_a \bar{\psi}^a(x) \gamma_\mu \psi^a(x), \quad J_\mu^i(x) = \sum_{a,b} \bar{\psi}^a(x) \tau_{ab}^i \gamma_\mu \psi^b(x).$$

The three Pauli matrices τ^i act on the $SU(2)$ indices.

The action (30.97) is the most general renormalizable action with the following symmetries:

- (i) A $U(1)$ chiral symmetry corresponding to the transformations,

$$\psi^a(x) \mapsto e^{i\gamma_3\theta} \psi^a(x), \quad \bar{\psi}^a(x) \mapsto \bar{\psi}^a e^{i\gamma_3\theta},$$

which prevents the addition of a fermion mass term.

- (ii) Separate fermion number conservation for each type of fermions, together with a symmetry between fermions 1 and 2. When $g_3 = 0$, the model has an additional $SU(2)$ symmetry.

Due to the $U(1)$ chiral invariance of the action (30.97), the coupling g_0 is not renormalized, as we have shown in Section 30.7.2.

For what follows, it is convenient to set

$$g = g_3 + f.$$

The RG β -functions at one-loop order are (see Section A30.1),

$$\beta_g = -2f^2/\pi, \quad \beta_f = -2fg/\pi. \quad (30.98)$$

The equivalence with the sG model ($\tau = \kappa^2$), which is summarized by the relations,

$$\frac{8\pi}{\tau} = 1 + \frac{g}{\pi}, \quad \frac{\alpha_0}{\tau} = \frac{f\Lambda^2}{\pi^2}, \quad (30.99)$$

$$\cos\theta(x) \mapsto \frac{\pi^2}{\Lambda^2} \bar{\psi}^1(x)\gamma_\mu\psi^2(x)\bar{\psi}^2(x)\gamma_\mu\psi^1(x), \quad (30.100)$$

will be derived in Section 30.7.2. These expressions show that, for τ close to 8π and α small, f and g are both small. The study of the phase transition is reduced to standard perturbation theory with fermion four-point renormalizable interactions.

Since the transition, which takes place at $g = f = 0$, is discussed extensively in Chapter 31, here we only note that, in the fermion language, the transition occurs between two phases, a massless phase as seen in perturbation theory, and a massive phase which exhibits the property of asymptotic freedom at large momentum.

30.7.2 Derivation

The action (30.97) is the sum of the two terms,

$$\begin{aligned} \mathcal{S}_0(\bar{\psi}, \psi) &= - \int d^2x [\bar{\psi}(x) \cdot \not{\partial} \psi(x) - \frac{1}{4}(g_0 J_\mu^0(x) J_\mu^0(x) + g J_\mu^3(x) J_\mu^3(x))], \\ \mathcal{S}_1(\bar{\psi}^a, \psi^a) &= \frac{1}{4}f \int d^2x (J_\mu^1(x) J_\mu^1(x) + J_\mu^2(x) J_\mu^2(x)). \end{aligned}$$

The first term \mathcal{S}_0 contains all interaction terms that have a separate chiral invariance for each species of fermions, because τ^3 is diagonal, and have an interaction of Thirring-model type. Like in the massless Thirring model, they can be transformed into a free-field action.

The remainder \mathcal{S}_1 is only invariant with respect to identical chiral transformations for ψ_1 and ψ_2 . It can be expanded in perturbation theory like the mass term in the Thirring model and, eventually, yields the interaction.

We thus express the interaction terms of \mathcal{S}_0 as resulting from a Gaussian integration over two vector fields $A_\mu^{0,3}$. The action then becomes

$$\begin{aligned}\mathcal{S}_0(\bar{\psi}, \psi, A^0, A^3) = & -\int d^2x \left[\bar{\psi}(x) \cdot (\partial + i\mathcal{A}^0(x) + i\mathcal{A}^3(x)\tau^3) \psi(x) \right. \\ & \left. - (A_\mu^0(x))^2/g_0 - (A_\mu^3(x))^2/g \right].\end{aligned}$$

Parametrizing the vector fields as in equation (30.49),

$$A_\mu^{0,3}(x) = -\partial_\mu \chi^{0,3}(x) - i\epsilon_{\mu\nu} \partial_\nu \varphi^{0,3}(x), \quad (30.101)$$

and performing the corresponding gauge transformations (30.51) on both fermion fields, we obtain (see equation (30.90)),

$$\begin{aligned}\mathcal{S}_0(\bar{\psi}, \psi, A^0, A^3) = & -\int d^2x \left[\bar{\psi}(x) \cdot \partial \psi(x) + \frac{\pi + g}{\pi g} (\nabla \varphi^3(x))^2 + \frac{\pi + g_0}{\pi g_0} (\nabla \varphi^0(x))^2 \right. \\ & \left. - \frac{1}{g} (\nabla \chi^3(x))^2 - \frac{1}{g_0} (\nabla \chi^0(x))^2 \right] \quad (30.102)\end{aligned}$$

(we have omitted the primes on the fermions). Note that

$$\frac{1}{4} [J_\mu^1(x) J_\mu^1(x) + J_\mu^2(x) J_\mu^2(x)] = \bar{\psi}^1(x) \gamma_\mu \psi^2(x) \bar{\psi}^2(x) \gamma_\mu \psi^1(x). \quad (30.103)$$

For \mathcal{S}_1 to be able to directly use the identities derived for the Thirring model, it is convenient to rewrite the f -term, using Fierz transformation (Section A12.3),

$$\frac{1}{2} \sum_i \tau_{ab}^i \tau_{dc}^i = \delta_{ac} \delta_{bd} - \frac{1}{2} \delta_{ab} \delta_{cd}.$$

Thus,

$$\bar{\psi}^1 \gamma_\mu \psi^2 \bar{\psi}^2 \gamma_\mu \psi^1 = -\bar{\psi}^1 \psi^1 \bar{\psi}^2 \psi^2 + \bar{\psi}^1 \gamma_3 \psi^1 \bar{\psi}^2 \gamma_3 \psi^2. \quad (30.104)$$

Introducing the chiral components of the mass operators,

$$\varsigma_\pm^1(x) = \frac{1}{2} \bar{\psi}^1(x) (1 \pm \gamma_3) \psi^1(x), \quad \varsigma_\pm^2(x) = \frac{1}{2} \bar{\psi}^2(x) (1 \pm \gamma_3) \psi^2(x),$$

we can rewrite the right-hand side as,

$$-\bar{\psi}^1 \psi^1 \bar{\psi}^2 \psi^2 + \bar{\psi}^1 \gamma_3 \psi^1 \bar{\psi}^2 \gamma_3 \psi^2 = -2 (\varsigma_+^1 \varsigma_-^2 + \varsigma_+^2 \varsigma_-^1). \quad (30.105)$$

Therefore, in a perturbative expansion in powers of f , the integrals over the two fermions factorize, and each term is just the square of the corresponding term in the expansion in powers of m in the Thirring model.

In the transformations (30.51), the chiral components ς_\pm^a become

$$\varsigma_\pm^1 \mapsto \varsigma_\pm^1 e^{\pm 2i(\varphi^0 + \varphi^3)}, \quad \varsigma_\pm^2 \mapsto \varsigma_\pm^2 e^{\pm 2i(\varphi^0 - \varphi^3)}.$$

Then, using identity (30.105), we can rewrite the f term as

$$\bar{\psi}^1 e^{2i\gamma_3\varphi^3} \gamma_\mu \psi^2 \bar{\psi}^2 e^{-2i\gamma_3\varphi^3} \gamma_\mu \psi^1 = -2 \left(e^{4i\varphi^3} \varsigma_+^1 \varsigma_-^2 + e^{-4i\varphi^3} \varsigma_+^2 \varsigma_-^1 \right). \quad (30.106)$$

We now associate two bosons ϑ^a with the fermions $\bar{\psi}^a, \psi^a$. In the expansion in powers of f , we note that we can simply replace the quantities ς_{\pm}^a by their boson counterparts. We then obtain the boson action

$$\mathcal{S}_{\text{bos.}} = \int d^2x \left\{ \frac{1}{2} \nabla \boldsymbol{\vartheta}(x) \cdot \nabla \boldsymbol{\vartheta}(x) - \frac{f\Lambda^2}{\pi^2} \cos \left[\sqrt{4\pi} (\vartheta^1(x) - \vartheta^2(x)) + 4\varphi^3(x) \right] \right. \\ \left. - \frac{\pi+g}{\pi g} (\nabla \varphi^3(x))^2 - \frac{\pi+g_0}{\pi g_0} (\nabla \varphi^0(x))^2 + \frac{1}{g} (\nabla \chi^3(x))^2 + \frac{1}{g_0} (\nabla \chi^0(x))^2 \right\}.$$

We change field variables $\vartheta^i \mapsto \theta^{0,3}$, with

$$\theta^0 = \sqrt{4\pi}(\vartheta^1 + \vartheta^2) + 4\varphi^0, \quad \theta^3 = \sqrt{4\pi}(\vartheta^1 - \vartheta^2) + 4\varphi^3.$$

Then,

$$\frac{1}{2} \nabla \boldsymbol{\vartheta}(x) \cdot \nabla \boldsymbol{\vartheta}(x) = \frac{1}{16\pi} \left[(\nabla \theta^0(x) - 4\nabla \varphi^0(x))^2 + (\nabla \theta^3(x) - 4\nabla \varphi^3(x))^2 \right].$$

Integrating over the fields $\varphi^{0,3}$ and $\chi^{0,3}$, we obtain the boson action

$$\mathcal{S}(\theta) = \int d^2x \left[\frac{(\pi+g)}{16\pi^2} (\nabla \theta^3(x))^2 - \frac{f\Lambda^2}{\pi^2} \cos \theta^3(x) + \frac{(\pi+g_0)}{16\pi^2} (\nabla \theta^0(x))^2 \right], \quad (30.107)$$

which is the sum of a sG action and a massless decoupled scalar action. This confirms that g_0 indeed is not renormalized.

A comparison with the form (30.64) shows that the sG model has a parameter τ :

$$8\pi/\tau = 1 + g/\pi. \quad (30.108)$$

Therefore, the value $\tau = 8\pi$ corresponds to a free massless fermion theory.

Finally, for $\tau < 8\pi$, although the remaining spectrum is massive, chiral symmetry is not broken, simply the corresponding field θ^0 is massless and free, and thus decouples.

30.8 The $SU(N)$ Thirring model

In Section 30.7, we have shown that the $SU(2)$ -invariant Thirring model is equivalent to a boson model, the sG model. It is clear from the derivation that more general models can be bosonized. As another example, we consider the $SU(N)$ Thirring model, a model described in terms of N Dirac fermions with current-current interactions, corresponding to the action

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x \left[\bar{\psi}(x) \cdot \partial \psi(x) - \frac{1}{2N} \left(g J_\mu^0(x) J_\mu^0(x) + f \sum_\alpha J_\mu^\alpha(x) J_\mu^\alpha(x) \right) \right], \quad (30.109)$$

$$= - \int d^2x \left[\bar{\psi}(x) \cdot \partial \psi(x) - \frac{1}{2N} (g-f)(\bar{\psi}(x) \cdot \gamma_\mu \psi(x))^2 \right. \\ \left. - \frac{1}{2} f \sum_{i,j} \bar{\psi}^i(x) \gamma_\mu \psi^j(x) \bar{\psi}^j(x) \gamma_\mu \psi^i(x) \right], \quad (30.110)$$

where J_μ^0 and J_μ^α are the $U(1)$ and $SU(N)$ currents, respectively,

$$J_\mu^0(x) = \bar{\psi}(x) \cdot \gamma_\mu \psi(x), \quad J_\mu^\alpha(x) = \bar{\psi}^i(x) \gamma_\mu t_{ij}^\alpha \psi^j(x),$$

and t^α is a generator of the $SU(N)$ Lie algebra.

The action has a $SU(N) \times U(1)$ vector together with a chiral $U(1)$ symmetry. After a Fierz transformation, the second interaction term in the action (30.110) becomes

$$\bar{\psi}^i \gamma_\mu \psi^j \bar{\psi}^j \gamma_\mu \psi^i = -\bar{\psi}^i \psi^i \bar{\psi}^j \psi^j + \bar{\psi}^i \gamma_3 \psi^i \bar{\psi}^j \gamma_3 \psi^j.$$

In the special case $g' = f'$, we recognize the action of the Nambu–Jona-Lasinio model, but such an equality is not preserved by renormalization. The two coupling constants are needed for renormalization purpose.

Due to the global chiral invariance of the model, the coupling constant g' associated with the current J_μ^0 is not renormalized. The remaining RG β_f -function in $d = 2 + \varepsilon$, at one-loop order, then is

$$\beta_f = \varepsilon f - \frac{N}{\pi} f^2, \quad (30.111)$$

showing that the two-dimensional model is asymptotically free, and has a non-perturbative spectrum for f positive and small. By contrast, for $f < 0$, the model is IR free: the fermions remain massless.

The model can be bosonized in the same way as the Thirring model. This yields a generalized sG model of the form [327] (Λ is the cut-off),

$$\begin{aligned} \mathcal{S}(\boldsymbol{\theta}) = \int d^2x \Bigg\{ & \frac{1}{4\pi^2} \left[(\pi + f) \sum_i (\nabla \theta^i(x))^2 + \frac{1}{N} (g - f) \left(\sum_i \nabla \theta^i(x) \right)^2 \right] \\ & - f \frac{\Lambda^2}{\pi^2} \sum_{i,j} \cos(\theta^i(x) - \theta^j(x)) \Bigg\}. \end{aligned} \quad (30.112)$$

The physics of the $SU(N)$ model is again the physics of spontaneous mass generation. In perturbation theory, we only see massless fermions but for $g > f$ (where the forces between fermions are attractive) the spectrum contains massive particles. However, chiral symmetry is not broken and $\langle \bar{\psi} \cdot \psi \rangle = 0$. The would-be Goldstone boson φ , which is associated with $U(1)$ chiral transformations, and thus, with the translation $\theta^i \mapsto \theta^i + \varphi$ (the centre of mass), is massless and free. The discussion of Section 30.1 then applies. Only chiral-invariant correlation functions are non-vanishing. Still, because the field φ decouples from other fields, in the massive phase of the model, correlation functions of $\theta_i - \theta_j$ only involve massive particles. This is another example of the Kosterlitz–Thouless mechanism (see Chapter 31). Note that the spectrum has been obtained exactly, $m_n \propto \frac{N}{\pi} \sin(n\pi/N)$, for $n = 1, 2, \dots < N$, n odd corresponding to fermions, and n even to bosons. This result confirms that fermions become massive.

30.8.1 Derivation

The method is a simple extension of the method explained in Section 30.7, in the $SU(2)$ example. Introducing the currents,

$$j_\mu^i(x) = \bar{\psi}^i(x) \gamma_\mu \psi^i(x), \quad (30.113)$$

we can write the action (30.110) as

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi) = - \int d^2x \Bigg[& \bar{\psi}(x) \cdot \partial \psi(x) - \frac{1}{2} \sum_{i,j} g_{ij} j_\mu^i(x) j_\mu^j(x) \\ & - \frac{1}{2} f \sum_{i \neq j} \bar{\psi}^i(x) \gamma_\mu \psi^j(x) \bar{\psi}^j(x) \gamma_\mu \psi^i(x) \Bigg], \end{aligned}$$

with

$$g_{ij} = f \delta_{ij} + (g - f)/N. \quad (30.114)$$

Introducing vector fields A_μ^i , we transform the action into

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi, \mathbf{A}_\mu) = & - \int d^2x \left[\sum_i \bar{\psi}^i(x) \left(\not{\partial} + i \not{A}^i(x) \right) \psi^i(x) - \frac{1}{2} \sum_{i,j} \tilde{g}_{ij} A_\mu^i(x) A_\mu^j(x) \right. \\ & \left. - \frac{1}{2} f \sum_{i \neq j} \bar{\psi}^i(x) \gamma_\mu \psi^j(x) \bar{\psi}^j(x) \gamma_\mu \psi^i(x) \right], \end{aligned} \quad (30.115)$$

with

$$\tilde{g}_{ij} \equiv [g^{-1}]_{ij} = \frac{1}{f} \left(\delta_{ij} - \frac{g-f}{Ng} \right). \quad (30.116)$$

Combining the usual $U(1)$ gauge transformations on the fermions with a parametrization of the vector fields,

$$\begin{aligned} \psi^i(x) &\mapsto e^{i\chi^i(x) + i\gamma_3\varphi^i(x)} \psi^i(x), \\ \bar{\psi}^i(x) &\mapsto \bar{\psi}^i e^{-i\chi^i(x) + i\gamma_3\varphi^i(x)}, \\ A_\mu^i(x) &= -\partial_\mu \chi^i(x) - i\epsilon_{\mu\nu}\partial_\nu \varphi^i(x), \end{aligned}$$

we find

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi, \varphi, \chi) = & \int d^2x \left[-\bar{\psi}(x) \cdot \not{\partial} \psi(x) + \frac{1}{2} \sum_{i,j} \tilde{g}_{ij} \nabla \chi^i(x) \nabla \chi^j(x) - \frac{1}{2} \sum_{i,j} \left(\tilde{g}_{ij} + \frac{\delta_{ij}}{\pi} \right) \nabla \varphi^i(x) \nabla \varphi^j(x) \right. \\ & \left. - \frac{1}{2} f \sum_{i \neq j} \bar{\psi}^i(x) e^{i\gamma_3(\varphi^i(x) - \varphi^j(x))} \gamma_\mu \psi^j(x) \bar{\psi}^j(x) e^{-i\gamma_3(\varphi^i(x) - \varphi^j(x))} \gamma_\mu \psi^i(x) \right]. \end{aligned} \quad (30.117)$$

In terms of the chiral components of the mass operator $\bar{\psi}^i \psi^i$:

$$\varsigma_\pm^i(x) = \bar{\psi}_\pm^i(x) \psi_\pm^i(x),$$

and using the identity (30.106), we obtain the identity,

$$\bar{\psi}^i e^{i\gamma_3(\varphi^i - \varphi^j)} \gamma_\mu \psi^j \bar{\psi}^j e^{-i\gamma_3(\varphi^i - \varphi^j)} \gamma_\mu \psi^i \underset{i \neq j}{=} -2 \left(e^{2i(\varphi^i - \varphi^j)} \varsigma_+^i \varsigma_-^j + e^{2i(\varphi^j - \varphi^i)} \varsigma_+^j \varsigma_-^i \right).$$

We then expand the four-fermion interaction in perturbation theory. As in the $SU(2)$ model, the expectation value over the fermions $\bar{\psi}^i, \psi^i$ can be replaced by an expectation value over bosons ϑ^i . After summation, this leads to a generalized sG interaction of the form

$$-f \frac{\Lambda^2}{\pi^2} \sum_{i \neq j} \cos \left[\sqrt{4\pi} (\vartheta^i - \vartheta^j) + 2(\varphi^i - \varphi^j) \right].$$

Then, changing variables, $\varphi^i \mapsto \theta^i$, setting

$$2\varphi^i + \sqrt{4\pi}\vartheta^i = \theta^i,$$

and integrating on all fields but the θ^i 's, we obtain the action (30.112).

A30 Two-dimensional models: A few additional results

A30.1 Four-fermion current interactions: RG β -function

A calculation of the RG β -function for an action with a general four-fermion, current-current interaction of the form

$$\mathcal{S}(\bar{\psi}, \psi) = - \int d^2x \left[\bar{\psi}(x) \not{\partial} \psi(x) - \frac{1}{2} \sum_{a,b,c,d} g_{abcd} \bar{\psi}^a(x) \gamma_\mu \psi^b(x) \bar{\psi}^c(x) \gamma_\mu \psi^d(x) \right], \quad (A30.1)$$

where g_{abcd} is chosen to be symmetric in the exchange $(ab) \leftrightarrow (cd)$, yields,

$$\pi \beta_{abcd}(g) = \sum_{i,j} g_{aicj} g_{ibjd} - g_{aijd} g_{ibcj} + O(g^3). \quad (A30.2)$$

A30.2 The Schwinger model: The anomaly

Here, in the Schwinger model, we relate the anomaly to the mass generation by two different methods.

A30.2.1 The general anomaly

The anomaly of the axial current $J_\mu^3 = -i\bar{\psi}\gamma_3\gamma_\mu\psi$ is (equation (23.92))

$$\partial_\mu J_\mu^3(x) = \frac{ie}{2\pi} \epsilon_{\mu\nu} F_{\mu\nu}(x). \quad (A30.3)$$

By contrast, the vector current $J_\mu = \bar{\psi}\gamma_\mu\psi$ is exactly conserved:

$$\partial_\mu J_\mu(x) = 0.$$

In dimension 2, the two currents are related (equation (30.24)),

$$J_\mu^3(x) = -i\epsilon_{\mu\nu} J_\nu(x),$$

and thus

$$\partial_\mu J_\nu(x) - \partial_\nu J_\mu(x) = \frac{ie}{\pi} F_{\mu\nu}(x).$$

The gauge field equation is

$$ieJ_\mu(x) + \partial_\nu F_{\nu\mu}(x) = 0.$$

Combining these equations, one finds

$$[-\nabla^2 + (e^2/\pi)] J_\mu(x) = 0.$$

This equation shows that the current J_μ , and thus the curvature $F_{\mu\nu}$, are free-fields of mass $m = e/\sqrt{\pi}$, in agreement with the result derived in Section 30.5.1.

A30.2.2 The two-point function: One-loop calculation

From the general analysis (Section 23.6), we know that we can regularize the fermion determinant while preserving the QED gauge invariance, but not always the chiral gauge invariance. Here, we explicitly work out the one-loop contribution to the A_μ field two-point function (Fig. 30.1) in the Feynman gauge. Using dimensional regularization, which preserves QED gauge invariance, and giving temporarily a mass M to the fermion field to avoid IR problems, we find

$$\tilde{\Gamma}_{\mu\nu}^{(2)}(k) = k^2 \delta_{\mu\nu} - e^2 \int \frac{d^d q}{(2\pi)^d} \text{tr} \left[\gamma_\mu \frac{M - iq}{M^2 + q^2} \gamma_\nu \frac{M - i(q+k)}{M^2 + (q+k)^2} \right]. \quad (\text{A30.4})$$

It is simple to verify that the dimensionally regularized one-loop contribution is transverse, as expected. Setting

$$\int \frac{d^d q}{(2\pi)^d} \text{tr} \left[\gamma_\mu \frac{M - iq}{M^2 + q^2} \gamma_\nu \frac{M - i(q+k)}{M^2 + (q+k)^2} \right] = D \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right), \quad (\text{A30.5})$$

and taking the trace of both sides, we obtain

$$D(k)(d-1) = \text{tr} \mathbf{1} \int \frac{d^d q}{(2\pi)^d} \frac{dM^2 + (d-2)(q^2 + q \cdot k)}{(q^2 + M^2)[(q+k)^2 + M^2]}. \quad (\text{A30.6})$$

The right-hand side contains a term proportional to M^2 which is UV convergent and which goes to 0 with M for $\mathbf{k} \neq 0$. The second term is proportional to $(d-2)$. In the $d=2$ limit, only the divergent part of the integral survives:

$$\int \frac{d^d q}{(2\pi)^d} \frac{q^2}{(q^2 + M^2)^2} \underset{d \rightarrow 2}{\sim} \frac{1}{2\pi(2-d)} \Rightarrow D = -1/\pi. \quad (\text{A30.7})$$

The gauge field inverse two-point function at one-loop order follows:

$$\tilde{\Gamma}_{\mu\nu}^{(2)}(k) = (k^2 \delta_{\mu\nu} - k_\mu k_\nu) \left(1 + \frac{e^2}{\pi k^2} \right) + k_\mu k_\nu, \quad (\text{A30.8})$$

a result consistent with the exact expression (30.76).

A30.3 Solitons in the sG model

Solitons correspond to finite energy solutions of the real-time classical equations of motion. Soliton calculus should be thought as the field theory generalization of the WKB method. Solitons have a particle interpretation, the energy of the soliton being its mass in the semi-classical limit. The sG model being classically integrable, the whole soliton spectrum can be derived, since time-dependent solutions can also be obtained [320, 321]. The semi-classical results have been confirmed by the exact quantum analysis. A simple example is provided by the mass (which is the rest energy) calculated from the static solution. The field equation for time-independent solutions derived from the action (30.63) is a differential equation in the space coordinate x :

$$\theta''(x) = \alpha \sin \theta(x).$$

Finite energy solutions necessarily connect minima of the potential. Thus,

$$\theta = 4 \arctan e^{\sqrt{\alpha}x}.$$

The corresponding energy then is the space integral of the Lagrangian density (after subtraction of the vacuum energy):

$$M_{\text{sol.}} = \frac{1}{\tau} \int dx \left[\frac{1}{2} (\partial_x \theta(x))^2 + \alpha (1 - \cos \theta(x)) \right] = 8\sqrt{\alpha}/\tau + O(1),$$

which coincides with the Thirring fermion mass (30.96) for τ small.

31 O(2) spin model and the Kosterlitz–Thouless’s (KT) phase transition

Having established, in Chapter 30, a few properties of some two-dimensional models, we now have the necessary technical background to discuss the Abelian $O(2)$ spin model near and in two dimensions. We have shown in Chapter 19 that, at low temperature, the large distance properties of the spin model can be described by the $O(2)$ non-linear σ -model. We recall that the $O(2)$ model is special, because the renormalization group (RG) β -function reduces, in the low-temperature expansion, to a dimensional linear term and, therefore, the properties, from the RG point of view, are quite different. In particular, in two dimensions, the β -function vanishes, and the $O(2)$ model is not asymptotically free. The origin of this difference can be found in the local structure of the manifold: for $N = 2$, the $O(N)$ sphere reduces to a circle, which cannot be distinguished locally from a straight line. Therefore, if we parametrize the spin $\mathbf{S}(x)$ as

$$\mathbf{S}(x) = \{\cos \theta(x), \sin \theta(x)\},$$

the Euclidean action (19.82), in zero field, becomes the free massless action,

$$\mathcal{S}(\theta) = \frac{\Lambda^{d-2}}{2t} \int d^d x [\nabla \theta(x)]^2, \quad (31.1)$$

where the coupling constant t is proportional to the temperature. This explains the form of the RG β -function. Nevertheless, because the physical fields are $\sin \theta$ or $\cos \theta$, or equivalently $e^{\pm i\theta}$, instead of θ , a field renormalization is necessary, and thus temperature-dependent anomalous dimensions are generated (equations (19.97) and (30.10)) [328].

However, the simplicity of the action (31.1) leads to a mystery: since the field θ is massless, the correlation length remains infinite for all temperatures. By contrast, a high-temperature analysis of the corresponding spin model on the lattice shows that the correlation length is finite for large enough temperature. A phase transition at finite temperature is necessary to explain this phenomenon. Therefore, the action (31.1) cannot describe the long-distance properties of the lattice model for all temperatures.

As the one-dimensional example of Section 3.4 indicates, it is necessary to somehow take into account the property that θ is a cyclic variable. This condition, which, obviously, is not incorporated in the action (31.1), is irrelevant at low temperatures, but when the temperature increases, classical configurations with singularities at isolated points, around which θ varies by a multiple of 2π , become important. The action of these configurations (vortices) can be identified with the energy of a *Coulomb gas*. The neutral Coulomb gas exhibits a transition between a low temperature of bound neutral molecules, and a high-temperature phase of a plasma of free charges.

Remarkably enough, the Coulomb gas can be mapped onto the sine-Gordon (sG) model, mapping in which the low- and high-temperature regions of the two models are exchanged. This correspondence helps to understand some properties of the transition, the famous *KT* phase transition [329], which separates an infinite correlation length phase *without order* (the low-temperature phase of the $O(2)$ spin model), from a finite correlation length phase (the high-temperature phase of the $O(2)$ spin model). Additional information about the transition can be obtained from the equivalence between the sG model and several fermion models, established in Sections 30.6 and 30.7.

31.1 The spin correlation functions at low temperature

The spin correlation functions in d dimensions with the free action (31.1) are given by

$$\left\langle \prod_{i=1}^n e^{i\epsilon_i \theta(x_i)} \right\rangle = \int [d\theta] \exp \left[-\frac{\Lambda^{d-2}}{2t} \int d^d x (\nabla \theta(x))^2 + i \sum_i \epsilon_i \theta(x_i) \right], \quad (31.2)$$

with $\epsilon_i = \pm 1$. They can be calculated by the method explained in Section 30.1. The result is

$$\left\langle \prod_{i=1}^n e^{i\epsilon_i \theta(x_i)} \right\rangle = \exp \left[-\frac{1}{2} t \Lambda^{2-d} \sum_{i,j} \epsilon_i \epsilon_j \Delta(x_i - x_j) \right], \quad (31.3)$$

in which $\Delta(x)$ is the massless propagator. For $d > 2$,

$$\Delta(x) = \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2} e^{ipx} = \Gamma(d/2 - 1) \frac{x^{2-d}}{4\pi^{d/2}}. \quad (31.4)$$

For $d \geq 2$, $\Delta(0)$ diverges with the cut-off Λ as

$$\int \frac{d^d p}{p^2} \propto \Lambda^{d-2}.$$

The limit $d = 2$ is singular, as we have discussed in Section 30.1, because the field is massless and this leads, in two dimensions, to infrared (IR) divergences.

Renormalization. To define renormalized correlation functions, we introduce a renormalization scale μ , and the corresponding renormalized coupling constant t_r :

$$t \Lambda^{2-d} = t_r \mu^{2-d},$$

in order to cancel the divergent terms in the exponential in the right-hand side of equation (31.3):

$$\sum_{i,j} \epsilon_i \epsilon_j \Delta(x_i - x_j) = n \Delta(0) + 2 \sum_{i < j} \epsilon_i \epsilon_j \Delta(x_i - x_j). \quad (31.5)$$

As noted in Section 30.1 (equation (30.9)), the fields $e^{\pm i\theta(x)}$, which are composite fields in terms of $\theta(x)$, require the field renormalization

$$Z = e^{-t_r \mu^{2-d} \Delta(0)}. \quad (31.6)$$

Dimension 2. We now examine more closely what happens when the dimension approaches 2 (at fixed cut-off). The propagator Δ has an IR divergence. Setting $d = 2 + \varepsilon$, one finds

$$\Delta(x) = \frac{1}{2\pi\varepsilon} - \frac{1}{4\pi} (\ln x^2 + \ln \pi + \gamma) + O(\varepsilon), \quad (31.7)$$

$$\Delta(0) = \frac{1}{2\pi\varepsilon} + \frac{1}{2\pi} \ln \Lambda + \text{terms finite when } \varepsilon \rightarrow 0 \text{ or } \Lambda \rightarrow \infty, \quad (31.8)$$

where $\gamma = -\psi(1)$ is Euler's constant. It follows that the sum of the divergent contributions in the exponential of equation (31.3) takes the form

$$-t \frac{\Lambda^{2-d}}{4\pi\varepsilon} \sum_{i,j} \epsilon_i \epsilon_j = -t \frac{\Lambda^{2-d}}{4\pi\varepsilon} \left(\sum_i \epsilon_i \right)^2. \quad (31.9)$$

Therefore, the only correlation functions that do not vanish in the $d = 2$ limit, are those for which the sum $\sum \epsilon_i$ vanishes. This result has a simple interpretation in the $O(2)$ -symmetric spin model: *all non $O(2)$ -invariant correlation functions vanish*. In particular,

$$\langle \mathbf{S}(x) \rangle = 0.$$

This result is consistent with the absence of spontaneous symmetry breaking with order in two dimensions. However, this property does not depend on the compact nature of the $O(2)$ group. All correlation functions that are invariant under a translation of $\theta(x)$ by a constant have a finite limit. In particular, as shown in Section 30.1, one can replace the signs ϵ_i by any set of numbers with a vanishing sum.

The non-vanishing limit is then

$$\left\langle \prod_{i=1}^n e^{i\epsilon_i \theta(x_i)} \right\rangle \propto \prod_{i < j} (\Lambda |x_i - x_j|)^{\epsilon_i \epsilon_j t / 2\pi}. \quad (31.10)$$

The result is rather surprising: although the $O(2)$ symmetry is not broken, the correlation functions decay algebraically at large distance, and the correlation length thus is infinite. Moreover, the power-law behaviour depends continuously on the temperature: in the RG sense, this signals a line of fixed points, and is consistent with the property that the RG β -function vanishes identically. Specializing equation (31.10) to the two-point function,

$$\langle e^{i\theta(x)} e^{-i\theta(0)} \rangle \sim x^{-t/2\pi}, \quad (31.11)$$

one infers the value of the critical exponent

$$\eta = t/2\pi. \quad (31.12)$$

The question now is how to modify the free-field theory (31.1) to generate the necessary phase transition. In particular, no invariant relevant operator can be constructed that would modify the action (31.1). Since we do not have yet all the necessary ingredients to examine this question, we postpone it until Section 31.4. Finally, we note that $t > 4\pi$ implies $\eta > 2$. This is a non-physical result from the point of view of the $O(2)$ spin model, since it implies that, in momentum space, the two-point correlation function vanishes at low momentum. In the corresponding lattice model, the correlation function is then dominated by a regular constant (non-critical) contribution. $t = 4\pi$ is thus a priori the largest possible value of t for which the action (31.1) can describe the $O(2)$ spin model below T_c .

31.2 Correlation functions in a field

In presence of a magnetic field, the action (31.1) becomes

$$\mathcal{S}(\theta) = \frac{\Lambda^{d-2}}{t} \int d^d x \left\{ \frac{1}{2} [\nabla \theta(x)]^2 - h \cos \theta(x) \right\}. \quad (31.13)$$

We recognize the sG action studied in Section 30.4. Since we know the two RG functions

$$\beta(t) = (d-2)t, \quad \zeta(t) = t/2\pi, \quad (31.14)$$

we can use the expressions of Chapter 19 to determine the scaling form of correlation functions.

For $d > 2$, the two functions $M_0(t)$ and $\xi(t)$ defined by equations (19.105, 19.106) are

$$M_0(t) = \exp \left[-\frac{1}{2} \int_0^t \frac{\zeta(t')}{\beta(t')} dt' \right] = \exp \left(-\frac{t}{4\pi\varepsilon} \right), \quad (31.15)$$

$$\xi(t) = \Lambda^{-1} t^{1/\varepsilon} \exp \left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{\varepsilon t'} \right) dt' \right] = \Lambda^{-1} t^{1/\varepsilon}. \quad (31.16)$$

This leads to the scaling form (equation (19.116))

$$\tilde{W}^{(n)}(p_i, t, h) = t^{(n-1)d/(d-2)} e^{-nt/4\pi\varepsilon} F^{(n)}(p_i t^{1/\varepsilon}, h e^{-t/4\pi\varepsilon} t^{-2/\varepsilon}). \quad (31.17)$$

Dimension 2. In dimension 2, the situation again is different, since the β -function vanishes. In perturbation theory, the action (31.13) is super-renormalizable. This corresponds to the relevance of the magnetic field for large-distance properties. The RG equations have been written in Sections 19.11 and 30.4. For any temperature t , the correlation functions have a scaling behaviour. The RG equation (19.93), in this particular case, reduces to

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \frac{n}{2} \zeta(t) + \frac{1}{2} \zeta(t) h \frac{\partial}{\partial h} \right) \tilde{W}^{(n)}(p_i, t, h) = 0. \quad (31.18)$$

Taking into account the dimensional relation

$$\Lambda \frac{\partial}{\partial \Lambda} + 2h \frac{\partial}{\partial h} + p_i \frac{\partial}{\partial p_i} = 2(1-n), \quad (31.19)$$

one infers the scaling form of correlation functions:

$$\tilde{W}^{(n)}(p_i, t, h) = h^{t/(8\pi-t)} h^{1-n} F^{(n)}(p_i h^{-4\pi/(8\pi-t)}). \quad (31.20)$$

This scaling form is consistent with the general scaling form of the ϕ^4 field theory at T_c in a magnetic field, and the value (31.12) of the exponent η . In particular, the magnetization M has the scaling form,

$$M = W^{(1)} \sim h^{t/(8\pi-t)}. \quad (31.21)$$

We have already noticed in Section 30.4, the special values $t = 4\pi$ and $t = 8\pi$. At $t = 4\pi$, the relation between M and h is linear up to a logarithmic correction: $M \propto -h \ln h$. This relation has been explained by the equivalence with the free massive Thirring model. As argued previously when examining the behaviour of the spin two-point function, for $t > 4\pi$, the model can no longer represent the large-distance physics of the lattice $O(2)$ spin model.

31.3 The Coulomb gas in two dimensions

We now postpone a further discussion of the $O(2)$ model, and study a model that is not obviously related, the two-dimensional Coulomb gas. We first show that, in any dimension, the Coulomb gas is related to the sG field theory [330].

31.3.1 Coulomb gas and sine-Gordon field theory

We first note that, in any dimension d , the Coulomb potential $V(x)$ is identical to the propagator (31.4). Therefore, the right-hand side of equation (31.3) is, up to a multiplicative factor, the Boltzmann weight for a gas of particles with charges $\epsilon_i = \pm 1$, and temperature $2\Lambda^{d-2}/t$. The identities of Section 30.2.3 then suggest that the sG field theory and the Coulomb gas are related. To establish the correspondence, we generalize the identities of Section 15.9.1 to two kind of particles, corresponding to the two possible charges, and apply them to the Coulomb potential. Actually for the physical interpretation of the results, it is convenient to add to the gas charge density an external background charge density $\rho_{\text{ext.}}$. In terms of the *charge density* $\rho(x)$ of the particles,

$$\rho(x) = \sum_a \epsilon_a \delta^{(d)}(x - x_a), \quad \epsilon_a = \pm 1, \quad (31.22)$$

the potential energy of the gas can then be written as

$$\mathcal{V}(\rho) = \frac{1}{2} \int d^d x d^d y (\rho(x) + \rho_{\text{ext.}}(x)) V(x - y) (\rho(y) + \rho_{\text{ext.}}(y)), \quad (31.23)$$

provided the neutrality condition is satisfied:

$$\int d^d x (\rho(x) + \rho_{\text{ext.}}(x)) = 0. \quad (31.24)$$

The neutrality condition implies that the total background charge takes integer values.

In the partition function, we implement the definition (31.22) by the field integral

$$\int [d\phi] \exp \left[i \int d^d x \phi(x) \rho(x) - i \sum_a \epsilon_a \phi(x_a) \right],$$

where we integrate over real fields ϕ . The partition function in the grand canonical formalism is defined in terms of a fugacity. Due to the neutrality condition, we can give the same fugacity z to the two kind of particles. At $\rho(x)$ fixed, the sum corresponds to independent particles in a potential $\pm i\phi$ and can be performed. The partition function at temperature T then becomes

$$\mathcal{Z} = \int [d\phi][d\rho] \exp [-\mathcal{S}(\rho, \phi)],$$

with

$$\mathcal{S}(\rho, \phi) = \mathcal{V}(\rho)/T - \int d^d x [i\phi(x)\rho(x) + 2z \cos \phi(x)], \quad (31.25)$$

because, in the expansion of $\exp[2z \int d^d x \cos \phi(x)]$ in powers of z , the terms that violate charge neutrality vanish, having infinite potential energy.

The integral over ρ is Gaussian and can be performed. The density ρ is then related to ϕ by the equation

$$\int d^d y V(x - y) (\rho(y) + \rho_{\text{ext.}}(y)) - iT\phi(x) = 0,$$

which can be inverted to yield

$$\rho(x) + \rho_{\text{ext.}}(x) = -iT\nabla^2 \phi(x). \quad (31.26)$$

This equation induces relations between charge density and ϕ -field correlation functions.

It also ensures the global electric neutrality of system. The action then becomes

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} T (\nabla \phi(x))^2 - 2z \cos \phi(x) + i\rho_{\text{ext.}}(x)\phi(x) \right]. \quad (31.27)$$

From the action (31.27), one infers relations between the change in free energy produced by background charges and ϕ -field correlation functions. For example, when the charges are localized and have integer values $\pm q$, the variation is obtained from correlation functions of $e^{\pm iq\phi(x)}$. For two infinitesimal, localized, opposite charges, the variation of the free energy is directly related to the ϕ two-point function.

The relation between Coulomb gas and sG model holds in any dimension, but we now discuss only the case of dimension 2. The properties of the sG model then provide important information about the physics of the Coulomb gas. The correspondence with the notation of Section 30.4 is

$$T = 1/\tau, \quad z = \alpha/2\tau.$$

Remark. We have shown that for $\tau = 4\pi$, the free energy becomes infinite in the absence of a short-distance cut-off. This implies that for $\tau \geq 4\pi$, that is, $T \leq 1/4\pi$, the Coulomb gas is stable only if the charged particles have a *hard core*.

The phase transition. We have shown in Section 30.4 that the quantum sG model must undergo a phase transition at $\tau = 8\pi$ for α small, that is, low fugacity and thus low-particle density, between a phase with a finite correlation length at low τ , that is, at high-temperature T in the Coulomb gas, and a phase with infinite correlation length at high τ , that is, at low T in the Coulomb gas. In the Coulomb gas terminology, the nature of the phases is clear (but not the existence of a transition at finite strictly positive temperature). At high T , the gas is composed of free charges. At low T , the system approaches the classical ground state: pairs of positive and negative charges are tightly bound. In the Coulomb gas terminology, the $i\nabla^2\phi$ correlation functions are the charge density correlation functions. Therefore, the correlation length characterizes the decay of the correlation between the charges. It also characterizes the decay of the effective potential between two infinitesimal external charges. In the free-charge phase, the correlation length is finite, which means that the electrostatic potential is screened, the correlation length being the screening length. By contrast, in the phase of molecular bound states no screening occurs, the effective potential is proportional to the bare potential, and the effective potential between integer background charges has a power law decay, the two-point charge density correlation function does not vanish only at coinciding points.

31.3.2 Renormalization and RG

We now examine precisely the properties of the phase transition. Since the transition point corresponds to a finite value of the coupling constant T^{-1} , the renormalization and RG properties of the model do not follow from simple perturbative considerations. The derivation of the RG equations within the sG framework thus involves a series of intuitive arguments, which are not easily made rigorous. However, in Chapter 30 we have derived a remarkable relation between the sG model and a two-fermion model, which is a free-field theory just at the transition point. This makes it possible to use perturbative arguments in the fermion model to confirm the RG equations obtained more directly, and to gain a further insight into the problem.

At $T = 1/8\pi$ (and $z \rightarrow 0$), the sG model is renormalizable (the interaction is marginal). To study its renormalization properties for T close to $1/8\pi$, we note that the deviation $(T^{-1} - 8\pi)$ plays a role analogous to the difference $(d-4)$ in the ϕ^4 field theory. Therefore, we can try to calculate RG quantities in a double z and $(8\pi - T^{-1})$ expansion. It is convenient to introduce the two dimensionless coupling constants,

$$u = 1 - 1/8\pi T, \quad v = 2z/T\Lambda^2. \quad (31.28)$$

We assume that the set $\{u, v\}$ is multiplicatively renormalizable. To these coupling constants correspond two RG β -functions. The property that the free-expectation value of the product of an odd number of fields $e^{\pm i\phi}$ vanishes (chiral symmetry in the equivalent fermion formulation) implies a reflection symmetry in v . The RG function β_v is given at leading order in v in terms of the function $\zeta(t)$ of equations (31.14). Thus,

$$\beta_v = v \left[\frac{1}{2} \zeta(T^{-1}) - 2 \right] + O(v^3) = -2uv + O(v^3). \quad (31.29)$$

We know that the function β_u vanishes at $v = 0$. Therefore, β_u starts at order v^2 . The sign of the coefficient of v^2 is fixed by the requirement that $u > 0$ and $u < 0$ cannot be connected, for v infinitesimal, by RG transformations. The exact value is normalization dependent. We choose

$$\beta_u = -2v^2, \quad \beta_v = -2uv. \quad (31.30)$$

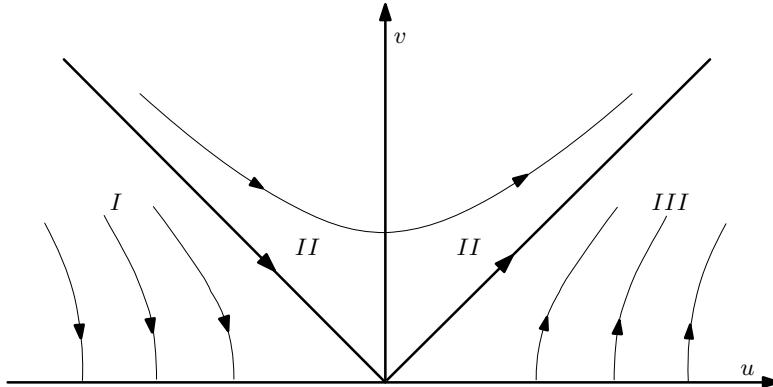


Fig. 31.1 The RG flow of the two-dimensional Coulomb gas

The RG flow. At this order, it is easy to find RG-invariant curves:

$$\frac{d}{d\lambda} (u^2(\lambda) - v^2(\lambda)) = 2u\beta_u - 2v\beta_v = 0. \quad (31.31)$$

For $|u| > |v|$, we parametrize these hyperbolas by

$$u = a \frac{1+s^2}{1-s^2}, \quad v = a \frac{2s}{1-s^2}, \quad \text{with } |s| < 1.$$

The RG equation for the parameter s is then

$$\lambda \frac{d}{d\lambda} s(\lambda) = -2as(\lambda), \Rightarrow s(\lambda) = s(1)\lambda^{-2a}. \quad (31.32)$$

In the large-distance limit, λ goes to 0. Therefore, if a is positive, that is, if initially $u > 0$ or $T > 1/8\pi$, then $s(\lambda)$ increases, as well as $u(\lambda)$ and $v(\lambda)$, until the RG trajectory leaves the perturbative regime. This is consistent with the relevance of v for large T . By contrast, if we start with $T < 1/8\pi$, which implies $u < 0$ and $a > 0$, then $s(\lambda)$ goes to 0 for $\lambda \rightarrow 0$. $v(\lambda)$ then goes to 0, and $u(\lambda)$ to a finite limit. This is consistent with the irrelevance of v for $T < 1/8\pi$. Thus, the half-line $T < 1/8\pi$, $z = 0$ is a half-line of IR fixed points.

By contrast, for $|u| < |v|$, we parametrize the hyperbolas by

$$u = a \frac{2s}{1 - s^2}, \quad v = a \frac{1 + s^2}{1 - s^2}, \quad \text{for } |s| < 1.$$

The flow equation then becomes,

$$\lambda \frac{d}{d\lambda} s(\lambda) = -a(1 + s^2) \Rightarrow \arctan s(\lambda) = -a \ln \lambda + \arctan s(1).$$

Therefore, irrespective of the sign of a , that is, the initial sign of u , $u(\lambda)$ and $v(\lambda)$ always increase until they leave the perturbative regime. This again is a region in which v is relevant. In Fig. 31.1, the regions are separated by the lines $u = \pm v$. Region I is an infinite correlation length phase, the low-temperature phase of the Coulomb problem. Regions II and III are both finite correlation phases with free charges. The line $u + v = 0$, thus, is the line of phase transition. Regions II and III differ by the property that, in region III, the field theory is asymptotically free at short distance while, in region II, the field theory is non-trivial at both long and short distances.

31.3.3 The correlation length near the phase transition

We now characterize the behaviour of the correlation length in the high-temperature phase, when one approaches the phase transition. We set

$$v + u = \tau, \quad v - u = 2w. \quad (31.33)$$

With this parametrization, τ plays the role of the deviation $T - T_c$ from the critical temperature.

The correlation length is a RG invariant of mass dimension -1 . Therefore, it satisfies

$$\xi(\tau(\lambda), w(\lambda)) = \lambda \xi(\tau(1), w(1)). \quad (31.34)$$

To find the relation between τ and ξ , we look for a value of λ such that $\xi(\lambda)$ is of order 1 when $\xi(1) = \xi$ is large:

$$\lambda \sim 1/\xi.$$

We have shown that the quantity τw is an RG invariant:

$$\tau(\lambda)w(\lambda) = \tau w.$$

The RG equation for $\tau(\lambda)$ then reads

$$\lambda \frac{d\tau}{d\lambda} = -2\tau w - \tau^2(\lambda). \quad (31.35)$$

A short calculation gives the function $\tau(\lambda)$:

$$\arctan\left(\tau(\lambda)/\sqrt{2w\tau}\right) - \arctan\left(\tau/\sqrt{2w\tau}\right) = -\sqrt{2w\tau} \ln \lambda. \quad (31.36)$$

We evaluate the left-hand side when τ is small and $\tau(\lambda)$ is of order 1. Then, the equation becomes

$$\pi/2 = -\sqrt{2w\tau} \ln \lambda.$$

We thus obtain

$$\xi \sim \exp\left(\frac{\pi}{2\sqrt{2w\tau}}\right), \quad (31.37)$$

a result that is characteristic of the 2D Coulomb gas phase transition.

Remark. The RG β -functions have also been calculated to third order. The result is

$$\beta_u = -2v^2 - 4v^2u, \quad (31.38)$$

$$\beta_v = -2uv + 5v^3. \quad (31.39)$$

None of the conclusions drawn from the analysis at leading order is qualitatively affected.

31.4 $O(2)$ spin model and Coulomb gas

Notation. In what follows, we again use the convention of implicit summation over repeated *Greek* indices.

We have argued that the action (31.1) cannot describe the $O(2)$ lattice spin model for all temperatures, and that one has to find a method to implement the condition that the $\theta(x)$ is a cyclic variable. We also know from the analysis of Section 19.10.2 that, at higher temperature, the non-linear σ -model approximation fails, because there may be many points where the field $\phi(x)$ in the $(\phi^2)^2$ field theory vanishes. The cost in energy is minimized when these points are isolated because this corresponds to point defects rather than line defects. Then, after a number of turns around these points, the direction of the field changes by an angle multiple of 2π . Therefore, we consider a configuration of the field $\theta(x)$ that is the sum of a smooth background $\theta_1(x)$, and a configuration $\theta_V(x, x_i)$ solution of the classical field equation, and regular everywhere except at isolated points x_i where $\theta_V(x, x_i)$ changes by a multiple of 2π ,

$$\theta(x) = \theta_1(x) + \theta_V(x, x_i), \quad (31.40)$$

with

$$\theta_V(x, x_i) = \sum_i n_i \arctan \frac{(x - x_i)_2}{(x - x_i)_1}, \quad n_i \in \mathbb{Z}. \quad (31.41)$$

The terminology is that $\theta_V(x, x_i)$ is a sum of vortex excitations located at points x_i and of vorticity n_i . Vortices are topological excitations in the sense that they cannot be removed by a continuous deformation of the field $\theta(x)$. The energy of the configuration is (setting $\Lambda = 1$)

$$\mathcal{S}(\theta) = \frac{1}{2t} \int d^2x [\nabla \theta_1(x) + \nabla \theta_V(x)]^2. \quad (31.42)$$

We now use the identities,

$$\partial_\mu \theta_V(x) = \sum_i n_i \epsilon_{\mu\nu} \frac{(x - x_i)_\nu}{(x - x_i)^2} = \sum_i n_i \epsilon_{\mu\nu} \partial_\nu \ln |x - x_i|. \quad (31.43)$$

We observe from expression (31.42) that the energy can be finite only if

$$\sum_i n_i = 0.$$

An integration by parts shows that the cross-term in the right-hand side of equation (31.42) vanishes, and yields

$$\begin{aligned} \int d^2x [\nabla \theta_V(x)]^2 &= \int d^2x \sum_{i,j} n_i n_j \nabla \ln |x - x_i| \nabla \ln |x - x_j| \\ &= -2\pi \sum_{i,j} n_i n_j \ln |x_i - x_j|. \end{aligned} \quad (31.44)$$

We recognize the energy of a neutral Coulomb gas of charges n_i and temperature $T = t/4\pi^2$. We know from the analysis of Section 30.4.2 (equation (30.70)) that the most relevant terms correspond to $n_i = \pm 1$, a restriction that we assume from now on. Of course, the fugacity A of the equivalent Coulomb gas has to be calculated from a microscopic model. Only results which are independent of its explicit value can be obtained by this method. The relation between the Coulomb gas temperature $t/4\pi^2$ and the temperature t' of the equivalent sG model is

$$t't = 4\pi^2. \quad (31.45)$$

In particular, by introducing an auxiliary field $\theta_2(x)$ and using the identity (31.3), we can write an effective action of the form

$$\mathcal{S}(\theta_1, \theta_2) = \int d^2x \left[\frac{1}{2t} (\nabla \theta_1(x))^2 + \frac{t}{8\pi^2} (\nabla \theta_2(x))^2 - \frac{At}{4\pi^2} \cos(\theta_2(x)) \right]. \quad (31.46)$$

The analysis of the Coulomb transition then shows that $t = \pi/2$ is the transition temperature. For $t > \pi/2$, the $\cos \theta$ interaction is relevant and the correlation length finite. For $t < \pi/2$, $\cos \theta$ is irrelevant and no mass is generated.

31.5 The critical two-point function in the $O(2)$ model

To determine the two-point correlation function of the $O(2)$ model near $t = \pi/2$, we have to calculate the expectation value (we follow here Ref. [332]),

$$\langle e^{i(\theta_1(y)+\theta_V(y))} e^{-i(\theta_1(x)+\theta_V(x))} \rangle = |x - y|^{-t/2\pi} \langle e^{i(\theta_V(y)-\theta_V(x))} \rangle_{\text{CG}}. \quad (31.47)$$

To find an interpretation to this expectation value in terms of a modification of the energy of the Coulomb gas, we make some transformations:

$$\begin{aligned} \theta_V(y) - \theta_V(x) &= \int_x^y ds \cdot \nabla \theta_V(s) \\ &= \frac{2\pi}{t} \sum_i n_i \frac{t}{2\pi} \int_x^y ds_\mu \epsilon_{\mu\nu} \partial_\nu \ln |s - x_i|. \end{aligned} \quad (31.48)$$

To calculate the Coulomb gas expectation value in the right-hand side of equation (31.47), we want to express the sum of the term (31.48) and the Coulomb gas energy (31.47), as arising from the free-field expectation value in a source.

The source $J_2(s)$ for the free-field $\theta_2(s)$,

$$J_2(s) = J(s) + K(s), \quad \begin{cases} J(s) = i \sum_i n_i \delta^{(2)}(s - x_i), \\ K(s) = \frac{t}{2\pi} \int_x^y ds'_\mu \epsilon_{\mu\nu} \partial_\nu \delta^{(2)}(s - s'), \end{cases} \quad (31.49)$$

generates the two terms we need. After integration over θ_2 , the JJ term gives the Coulomb gas energy, the JK cross-term gives the term (31.48) but, in addition, a KK contribution $D(x, y)$ arises of the form

$$\begin{aligned} D(x, y) &\equiv \frac{2\pi^2}{t} \int ds ds' K(s) K(s') \Delta(s - s') \\ &= -\frac{t}{4\pi} \int_x^y ds_\mu ds'_\nu \epsilon_{\mu\rho} \epsilon_{\nu\sigma} \partial_\rho^s \partial_\sigma^{s'} \ln |s - s'|. \end{aligned} \quad (31.50)$$

We can thus express the Coulomb expectation value (31.47) as a free-field expectation value in the source (31.49), provided we multiply the expression by e^{-D} . To evaluate $D(x, y)$, we use the identity

$$\epsilon_{\mu\rho} \epsilon_{\nu\sigma} = \delta_{\mu\nu} \delta_{\rho\sigma} - \delta_{\mu\sigma} \delta_{\nu\rho},$$

and find

$$D(x, y) = -\frac{t}{2} \int_x^y ds \cdot ds' \delta(s - s') + \frac{t}{4\pi} \int_x^y ds_\mu ds'_\nu \partial_\mu^s \partial_\nu^{s'} \ln |s - s'|. \quad (31.51)$$

The integration over s and s' can be performed. Only one term coming from the second integral, gives a contribution that depends on $x - y$, the others are cut-off dependent. The final result is

$$D(x, y) = -\frac{t}{2\pi} \ln |x - y| + \text{const.} \quad (31.52)$$

If we replace the Coulomb gas expectation value in equation (31.47) by the free-field expectation value with the source (31.49), multiplied by e^{-D} , something remarkable happens: the last factor just cancels the factor coming from the integration over θ_1 . Therefore, the spin correlation functions can be entirely calculated from the sG model, the spin field being represented by the non-local expression, related to $K(s)$ in the equations (31.49), depending on a field attached to a string, (see also equation (30.88b)):

$$\exp [i\theta(x)] \mapsto \exp \left[-\frac{t}{2\pi} \int_x^x ds_\mu \epsilon_{\mu\nu} \partial_\nu \theta_2(s) \right]. \quad (31.53)$$

The critical two-point function. To calculate the two-point function at the critical temperature, we now use the RG considerations of Section 31.3.1. Criticality corresponds to the special line $u = -v$, in the notation of Section 31.3.1. Note that the relation between t and u on the critical line is

$$t = \pi/2(1 - u). \quad (31.54)$$

The origin $u = v = 0$ is an IR fixed point. The behaviour of the correlation function can thus be obtained from perturbation theory. At leading order in the parameter A of equation (31.46), the two-point function is still given by equation (31.11).

RG equations. The critical spin two-point correlation function $W^{(2)}$ satisfies the RG equation,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t} + \zeta(t) \right] W^{(2)}(x, t, \Lambda) = 0. \quad (31.55)$$

The functions $\zeta(t)$ and $\beta(t)$ are (equations (31.14), (31.30), and (31.54))

$$\begin{aligned} \zeta(t) &= t/2\pi + O\left(t - \frac{\pi}{2}\right)^2, \\ \beta(t) &= -\frac{4}{\pi} \left(t - \frac{\pi}{2}\right)^2 + O\left(t - \frac{\pi}{2}\right)^3. \end{aligned} \quad (31.56)$$

For $t < \pi/2$, the theory is IR-free. The effective coupling constant at scale λ behaves like

$$t(\lambda) = \frac{\pi}{2} \left(1 + \frac{1}{2 \ln \lambda}\right) + o\left(\frac{1}{\ln \lambda}\right), \text{ for } \lambda \rightarrow 0. \quad (31.57)$$

$W^{(2)}(x, t)$ being dimensionless, the solution of the RG equation can be written as

$$W^{(2)}(x/\lambda, t) = Z^2(\lambda) W^{(2)}(x, t(\lambda)), \quad (31.58)$$

with

$$Z(\lambda) = \exp \left[\frac{1}{2} \int_1^\lambda \frac{d\sigma}{\sigma} \zeta(t(\sigma)) \right] \sim \lambda^{1/8} |\ln \lambda|^{1/16}. \quad (31.59)$$

Therefore, at large distance, the critical two-point correlation function $W^{(2)}$ behaves like

$$W^{(2)}(x, t) \underset{x \rightarrow \infty}{\sim} x^{-1/4} (\ln x)^{1/8}. \quad (31.60)$$

This is a celebrated KT result for the phase transition of the classical two-dimensional XY model [329, 331].

31.6 The generalized Thirring model

We now use the results of Section 30.7 to justify the RG functions (31.30), and thus the RG flow near the phase transition. The correspondence with the sG model (see equation (30.107)), in the notation of the effective $O(2)$ spin-model action (31.46) is

$$A \sim f, \quad t = \frac{1}{2}\pi(1 + g/\pi). \quad (31.61)$$

From these expressions, we infer that, for t close to $\pi/2$ and A small, f and g are both small. The study of the phase transition is reduced to standard perturbation theory with fermion four-point renormalizable interactions. The β -functions at one-loop order are (equation (A30.2))

$$\beta_g = -2f^2/\pi, \quad \beta_f = -2fg/\pi. \quad (31.62)$$

We recognize the RG functions (31.30) in a different parametrization.

The KT phase transition then has an interpretation in the fermion language. The phase diagram is simplest in the case of the $SU(2)$ -symmetric model. For g positive, the force between fermions is attractive. The model is asymptotically free and the spectrum, which is non-perturbative, contains massive particles and a massless boson corresponding to chiral transformations, which decouples. Note that, as we have discussed in Section 30.8, this does not mean that chiral symmetry in the fermion model is spontaneously broken; the expectation value of the order parameter $\langle \bar{\psi}\psi \rangle$ vanishes (as well as all non-chiral invariant correlation functions) in agreement with the Mermin–Wagner theorem. By contrast, for $g < 0$, the force between fermions is repulsive, the model is IR-free and the spectrum is perturbative with massless fermions.

The same picture is valid for the $SU(N)$ Thirring model (see Section 30.8).

For the two-parameter model, we have the three regions of Fig. 31.1, a region I in which particles are massless, and regions II and III where particles are massive.

Finally, note that in the framework of the simple massive Thirring model of Section 30.6, the situation is reversed. We have a transition at $g = -\pi/2$ (for m small) between a perturbative massive phase and a massless phase where the mass operator $\bar{\psi}\psi$ becomes irrelevant.

32 Finite-size effects in field theory. Scaling behaviour

A number of numerical calculations, like Monte Carlo or transfer matrix calculations, are performed with systems in which the size in several or all dimensions is finite. To extrapolate the results to the infinite system, it is thus necessary to understand how the infinite-size limit is reached. In particular, in a system in which the forces are short-range, no phase transition can occur in a finite volume, or in a geometry in which the size is infinite only in one dimension. This indicates that the infinite-size extrapolation is somewhat non-trivial [333] (or reviews see *e.g.* Ref. [334]). We expect finite-size effects to become important when the correlation length becomes comparable to the finite size. We thus present in this chapter an analysis of the problem in the case of second order phase transitions, in the framework of the N -vector model. We first establish the existence of a finite-size scaling, extending renormalization group (RG) arguments to this new situation. We then distinguish between the finite-volume geometry (in explicit calculations we take the example of the hypercube), and the cylindrical geometry in which the size is finite in all dimensions except one. We explain how to adapt the methods used in the case of infinite systems to calculate the new universal quantities appearing in finite-size effects, for example, in $d = 4 - \varepsilon$ or $d = 2 + \varepsilon$ dimensions. Special properties of the commonly used (although non-physical) periodic boundary conditions are emphasized.

Note that, from the point of view of classical statistical field theory, finite-temperature quantum field theory (QFT) can be considered as a theory with one finite size $1/T$, where T is the temperature. Finite temperature QFTs are discussed mainly in Chapter 33, although the two-dimensional non-linear σ -model, with one infinite size is relevant to both situations. Moreover, for a class of systems, zero temperature quantum transitions in d dimensions share properties with classical transitions in $(d + 1)$ dimensions.

The appendix contains a few remarks about finite-size effects in the ordered phase, when the correlation length is finite, and about the calculation of one-loop finite-size Feynman diagrams.

32.1 RG in finite geometries

In this chapter, we consider systems that have sizes characterized by one length L , which is large in the microscopic scale, for example, much larger than the lattice spacing in lattice models. When the correlation length is also large, the universal properties of the system can be described by a continuum field theory. We consider only boundary conditions that *do not break translation symmetry* to avoid surface effects which are of a different nature. Periodic boundary conditions satisfy such a criterion. Depending on the specific symmetries of a model, other boundary conditions may also be available (like anti-periodic boundary conditions for Ising-like systems).

The crucial observation that explains finite-size scaling is that renormalization theory, which leads to RG equations is *insensitive to finite-size effects*, since renormalizations are due to *short-distance singularities*. Therefore, RG equations are not modified, but the solutions are different, because correlation functions now depend on the additional dimensional parameter L [335, 336].

We will discuss the solutions of RG equations both in the examples of the $(\phi^2)^2$ QFT, and of the $O(N)$ -symmetric non-linear σ -model.

32.1.1 The $(\phi^2)^2$ field theory for $d < 4$

In the continuum limit, the N -vector model can be described by a $(\phi^2)^2$ QFT with the Euclidean action (Chapter 16),

$$\mathcal{S}(\phi) = \int \left\{ \frac{1}{2} [\nabla \phi(x)]^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} u (\phi^2(x))^2 \right\} d^d x, \text{ with } u > 0, \quad (32.1)$$

where ϕ is a N -component field. For $r = r_c$, the theory is critical (massless) and we set $r = r_c + \tau$, where τ characterizes the deviation from the critical temperature. Feynman diagrams of perturbation theory are calculated with a large momentum cut-off Λ inverse of the microscopic scale (like the lattice spacing in lattice models). For $d < 4$, in terms of the dimensionless coupling constant $g = u\Lambda^{d-4}$, the corresponding vertex functions (in the Fourier representation) satisfy the RG equations (Section 16.5),

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \eta(g) \left(n + M \frac{\partial}{\partial M} \right) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}^{(n)}(p_i; \tau, M, g, L, \Lambda) = 0. \quad (32.2)$$

In terms of a scaling parameter λ , the solutions can be written as

$$\tilde{\Gamma}^{(n)}(p_i; \tau, M, g, L, \Lambda) = Z^{-n/2}(\lambda) \tilde{\Gamma}^{(n)}(p_i; \tau(\lambda), M/\sqrt{Z(\lambda)}, g(\lambda), L, \lambda\Lambda), \quad (32.3)$$

the various functions of λ being defined by (equations (16.26–16.28)),

$$\begin{aligned} \lambda \frac{d}{d\lambda} g(\lambda) &= \beta(g(\lambda)), & g(1) &= g, \\ \lambda \frac{d}{d\lambda} \ln \tau(\lambda) &= -\eta_2(g(\lambda)), & \tau(1) &= \tau, \\ \lambda \frac{d}{d\lambda} \ln Z(\lambda) &= \eta(g(\lambda)), & Z(1) &= 1. \end{aligned}$$

However, the presence of the new length scale L modifies the dimensional relations, which become,

$$\tilde{\Gamma}^{(n)}(p_i; \tau, M, g, L, \Lambda) = \Lambda^{d-n(d-2)/2} \tilde{\Gamma}^{(n)}(p_i/\Lambda; \tau/\Lambda^2, M\Lambda^{(2-d)/2}, g, L\Lambda, 1). \quad (32.4)$$

We use this relation in the right-hand side of equation (32.3), and choose $\lambda = 1/L\Lambda$. The dilatation parameter λ goes to zero for ΛL large and, therefore, $g(\lambda)$ approaches the infrared (IR) fixed point g^* . This implies for $Z(\lambda)$ and $t(\lambda)$, the behaviour ($\eta_2(g^*) = 1/\nu - 2$),

$$Z(\lambda) \propto (L\Lambda)^{-\eta}, \quad \frac{\tau(\lambda)}{\lambda^2 \Lambda^2} \propto \frac{\tau}{\Lambda^2} (L\Lambda)^{1/\nu}. \quad (32.5)$$

The scaling behaviour of finite-size correlation functions follows:

$$\tilde{\Gamma}^{(n)}(p_i; \tau, M, g, L, 1) \propto L^{-d+n(d-2+\eta)/2} \tilde{\Gamma}^{(n)}(Lp_i; \tau L^{1/\nu}, M L^{\beta/\nu}, g^*, 1, 1). \quad (32.6)$$

It is characterized by the presence of a new scaling variable $\tau L^{1/\nu} \propto (\xi/L)^{1/\nu}$ where, for $\tau > 0$, $\xi(\tau)$ is the correlation length and, more generally, an RG-invariant physical scale.

From equation (32.6), the usual infinite-size scaling form is recovered by expressing that $\tilde{\Gamma}^{(n)}$ has a limit for $L \gg \xi(\tau)$. In the opposite limit $\xi(\tau) \gg L$, correlation functions have a regular expansion in powers of τ , even for vanishing magnetization in a finite volume or for a cylindrical geometry, because no phase transitions can occur in both situations (for short-range interactions).

Note that all combinations that are independent of the normalization of the field ϕ , of the temperature τ , and of the magnetic field are universal for the reasons explained in the infinite volume case, once the geometry and boundary conditions are fixed.

Adapting the usual analysis of corrections to scaling to the finite-size situation, one concludes that the leading corrections to the scaling form (32.6) have, near $d = 4$, the form of a scaling function multiplied by a factor $L^{-\omega}$ ($\omega = \beta'(g^*)$).

Examples

(i) We consider a d -dimensional system in a hypercube of side length L , and the corresponding distribution of the average of the field ϕ in the hypercube

$$\varphi = \frac{1}{L^d} \int \phi(x) d^d x. \quad (32.7)$$

The distribution $e^{-\Sigma(\varphi)}$ of the φ variables can be inferred from the distribution $e^{-\mathcal{S}(\phi)}$ of the field ϕ . We consider the moments of the φ distribution, which we denote by

$$m_{2p} = \mathcal{Z}^{-1} \int d\varphi (\varphi^2)^p e^{-\Sigma(\varphi)} \equiv \langle |\varphi|^{2p} \rangle, \quad (32.8)$$

where \mathcal{Z} is the partition function,

$$\mathcal{Z} = \int d\varphi e^{-\Sigma(\varphi)}.$$

One can then define ratio of moments, for example,

$$\mathcal{R}_4 = \frac{m_4}{m_2^2}. \quad (32.9)$$

At leading order, \mathcal{R}_4 has the scaling form

$$\mathcal{R}_4(\tau, L) \sim g(\tau L^{1/\nu}). \quad (32.10)$$

The ratio can be used to determine the critical temperature. Indeed, by calculating $\mathcal{R}_4(\tau, L)$ for different values of L , and looking for a temperature at which it is independent of L , one finds the critical temperature $\tau = 0$ provided the corrections in $L^{-\omega}$ can be neglected. The quantity $g(0)$ is a universal number which, in principle, can be calculated with the continuum field theory. We examine this problem later.

(ii) In a cylindrical geometry, the correlation length ξ_L in the infinite direction is another quantity of interest. From equation (32.6), one concludes that, at leading order, it has the functional form,

$$\xi_L \sim LX(\tau L^{1/\nu}). \quad (32.11)$$

In particular, at $\tau = 0$, ξ_L grows linearly with L , and the ratio $\xi_L/L = X(0)$ is *universal*. Since for $\tau > 0$, ξ_L goes to a constant for large L and, since for $\tau < 0$, it grows faster than L , as one can verify, the ratio ξ_L/L can be used to determine the critical temperature in transfer matrix calculations.

Since ξ_L is related to the ratio of the two largest eigenvalues of the transfer matrix (equation (14.10)) λ_0 and λ_1 , we infer also that

$$\lambda_0/\lambda_1 - 1 \sim 1/\xi_L = 1/LX(0).$$

With this knowledge, it is instructive to return to the analysis of the existence of phase transitions in Chapter 14.

32.1.2 Low-temperature expansion and finite-size effects

In models with *continuous symmetries*, the large-distance behaviour, in the ordered low-temperature phase is described by an effective interaction between Goldstone modes. Therefore, we also examine the problem of finite-size effects in this context.

In the case of the $O(N)$ -symmetric N -vector model, universal physical observables can be derived from the low-temperature, or low-coupling expansion of the non-linear σ -model (see Chapter 19), whose partition function is given by (equations (19.3, 19.4))

$$\mathcal{Z} = \int [d\phi(x)] \prod_x \delta(\phi^2(x) - 1) \exp[-\mathcal{S}(\phi)], \quad (32.12)$$

with

$$\mathcal{S}(\phi) = \frac{\Lambda^{d-2}}{2t} \int d^d x (\nabla \phi(x))^2. \quad (32.13)$$

With this parametrization, t is dimensionless, and Λ is a large momentum cut-off.

The arguments of Section 32.1.1 also apply to the RG equations derived for the σ -model: equations remain unchanged, only solutions are modified by the finite size. General solutions (19.114) and (19.116) of the RG equations (19.93) now depend on an additional scaling variable $L/\xi(t)$, where, for $d > 2, t < t_c$, the RG-invariant length $\xi(t)$ is defined by equation (19.106):

$$\xi(t) = \Lambda^{-1} t^{1/(d-2)} \exp \left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{(d-2)t'} \right) dt' \right] \quad (32.14)$$

and, for $d = 2$, by

$$\xi(t) \propto \Lambda^{-1} \exp \left[\int_0^t \frac{dt'}{\beta(t')} \right]. \quad (32.15)$$

Alternatively, the solutions can be parametrized in terms of a size-dependent temperature t_L , obtained by solving the equation

$$\lambda \frac{d}{d\lambda} t(\lambda) = \beta[t(\lambda)], \quad t(1) = t, \quad (32.16)$$

and choosing $\lambda = 1/\Lambda L$, such that

$$t_L \equiv t(1/\Lambda L). \quad (32.17)$$

Then,

$$\ln(\Lambda L) = \int_{t_L}^t \frac{dt'}{\beta(t')}, \quad (32.18)$$

equation which, in particular, shows that t_L is a function of t and L only through the expected combination $L/\xi(t)$.

At one-loop order, in a cut-off scheme the RG $\beta(t)$ function has the expansion (Section 19.11)

$$\beta(t) = (d-2)t + \beta_2(d)t^2 + O(t^3), \quad (32.19)$$

with

$$\beta_2(d) = -(N-2)/2\pi + O(d-2).$$

For $d > 2$ and $t < t_c$ fixed (and thus the length $\xi(t)$ is of order $1/\Lambda$), when ΛL increases t_L approaches the IR fixed point $t = 0$:

$$t_L \sim (\xi(t)/L)^{d-2} \ll 1. \quad (32.20)$$

Therefore, finite-size effects can be calculated from the low-temperature expansion and RG.

At t_c , and more generally in the critical domain, physical quantities can be calculated in an $\varepsilon = d - 2$ expansion, as shown in Chapter 19. Since t_c is an RG fixed point, $t_L(t_c) = t_c$.

Calculations can also be performed in *two dimensions*, even in zero magnetic field h , because L provides an IR cut-off. However, because $t = 0$ is then an ultraviolet (UV) fixed point, t_L goes to 0 for $L/\xi(t)$ small:

$$t_L \sim \frac{2\pi}{(N-2)\ln(\xi(t)/L)}, \quad (32.21)$$

and this is the limit in which physical quantities can be calculated.

Finally, solving equation (32.18) perturbatively, one finds

$$\frac{1}{t_L} = \frac{(\Lambda L)^{d-2}}{t} + \frac{\beta_2(d)}{d-2} [(\Lambda L)^{d-2} - 1] + O(t). \quad (32.22)$$

Note that a statistical (classical) field theory in two dimensions with a finite size in one dimension, and periodic boundary conditions, is also a finite-temperature QFT (Chapter 33). Therefore, the discussion of Section 32.5.2 is also relevant for such a quantum theory.

32.2 Momentum quantization

The scaling properties (32.6) do not depend on the specific choice of boundary conditions, but the explicit universal finite-size expressions do. Even the technical details of the calculation, when the temperature approaches T_c , vary. Indeed, a characteristic feature of all finite geometries is that, in the Fourier representation, momenta corresponding to directions in which the size of the system is finite are quantized. The precise momentum spectrum depends on the boundary conditions. Periodic boundary conditions, which we use throughout this chapter, and other non-periodic boundary conditions which do not break translation symmetry (twisted boundary conditions), have different properties. We briefly discuss twisted boundary conditions in Section 32.2.2. An example has actually been worked out in Section 3.5.

32.2.1 Periodic boundary conditions and the zero mode

In the example of a d -dimensional hypercube of linear size L with periodic boundary conditions, the quantized momenta p have the form

$$p = 2\pi k/L, \quad k \in \mathbb{Z}^d. \quad (32.23)$$

When the product $\tau L^{1/\nu} = (L/\xi)^{1/\nu}$ is positive and not small, the usual methods of calculation of the infinite volume are applicable, and finite size effects due to momentum quantization are only quantitative, decreasing exponentially like $\exp[-\text{const. } L/\xi]$. When the product $\tau L^{1/\nu}$ is negative and not small (the ordered phase), the physics of the infinite and finite systems are quite different, and this problem will be examined later.

Finally, in a finite volume, at T_c ($\tau = 0$) the propagator has an isolated pole at $p = 0$, which generates IR divergences in perturbation theory, although we expect physical quantities to be regular in T at T_c . These divergences simply reflect a disease of the Gaussian model. More generally, straightforward perturbation theory is badly behaved for $\xi \gg L$.

As a consequence, even in high dimensions, for which in the infinite geometry mean-field theory is exact, IR divergences appear at T_c . To overcome this difficulty, it is necessary to deal separately with the zero-momentum Fourier component of the field. The components $p \neq 0$ can be treated by the methods developed for the infinite geometry (perturbation theory and RG); the component $k = 0$, whose fluctuations at T_c are only damped by interaction terms has to be treated exactly. Therefore, in the case of a finite volume, we calculate exactly an effective integral over the component $\tilde{\phi}(p = 0)$, obtained by integrating perturbatively over all other components.

In the cylindrical geometry, one component of the momentum to be denoted by ω varies continuously, and the other components are quantized. At T_c , Feynman diagrams contain divergent contributions of the form $\int d\omega/\omega^2$. The integration over all components except $\tilde{\phi}(p_T = 0, z)$, denoting by z the coordinate in the infinite direction, leads to an effective quantum Hamiltonian. Note that similar considerations apply to the zero modes of instanton calculations (see Chapters 37–42).

Finally, a geometry in which the sizes in two or three dimensions among d are infinite still leads to IR divergences. Some examples of such situations appear in Chapter 33, devoted to finite-temperature QFT. Then, the IR problem can no longer be solved exactly.

We examine, in Section 32.3, the two first geometries separately, beginning with the simplest case of the periodic hypercube.

32.2.2 Twisted boundary conditions

For systems with symmetries, additional boundary conditions do not break translation invariance: conditions such that the values of the order parameter at both boundaries (for each direction in space) differ by a constant group transformation (often called twisted boundary conditions). For instance, for Ising-like systems, one can use anti-periodic boundary conditions; for the N -vector model with $O(N)$ symmetry, one can impose a rotation of a given angle around some axis. In such situations, the quantized momenta p_μ are shifted by some additional constants, as we now show.

We consider a scalar field theory invariant under the transformations of a unitary representation of a Lie group G . We impose to the field ϕ , the boundary conditions

$$\phi(x_1, x_2, \dots, x_\mu + L, \dots, x_d) = e^{A_\mu} \phi(x_1, x_2, \dots, x_\mu, \dots, x_d),$$

where the A_μ 's are constant (*i.e.* space-independent) commuting anti-Hermitian matrices, $[A_\mu, A_\nu] = 0$, elements of the Lie algebra of G (A_μ is a curvature-free gauge field).

To return to the situation of periodic boundary conditions, we perform a gauge transformation on the field ϕ , setting

$$\phi(x) = U(x)\phi'(x), \quad \text{with } U(x) = e^{A \cdot x/L}.$$

The new field ϕ' then satisfies periodic boundary conditions. The action is invariant except for the derivatives, which now become covariant derivatives:

$$U^{-1}(x)\partial_\mu\phi(x) = (\partial_\mu + A_\mu/L)\phi'(x) \Rightarrow p_\mu\tilde{\phi}(p) \mapsto (p_\mu - iA_\mu/L)\tilde{\phi}'(p).$$

Thus, after Fourier transformation the derivatives yield quantized momenta of the form $p_\mu = (2\pi k_\mu - iA_\mu)/L$. The anti-Hermitian matrices A_μ have imaginary eigenvalues $i\theta_\mu^\alpha$. Expanding the field ϕ' on a basis in which all A_μ (which commute) are diagonal, one finds that the effect of twisted boundary conditions is to generate a set of quantized momenta of the form

$$p_\mu = (2\pi k_\mu + \theta_\mu^\alpha)/L, \quad k_\mu \in \mathbb{Z}^d,$$

for each component ϕ_α of the field. One can choose all angles θ_μ^α to belong to the interval $[-\pi, \pi]$.

The calculation of the correlation length [337] provides an example.

32.3 The ϕ^4 field theory in a periodic hypercube

We first study the effective $(\phi^2)^2$ -field theory in the dimensions $d > 4$ and $d = 4 - \varepsilon$. As we have explained in Section 32.2.1, we expand $\phi(x)$ in Fourier components, separating the zero mode:

$$\begin{aligned} \phi(x) &= \varphi + \chi(x), \\ \chi(x) &= (2\pi/L)^d \sum_{p \neq 0} e^{ip \cdot x} \phi(p), \quad \text{with } p = 2\pi k/L, \quad k \in \mathbb{Z}^d. \end{aligned} \quad (32.24)$$

The integration over the field $\chi(x)$ is performed as in the infinite geometry limit: this generates a perturbative expansion, which has RG properties. An integral over the last $k = 0$ modes remains, which must be calculated exactly. Note that the first part of the procedure is formally equivalent to the shift of the expectation value of the field $\phi(x)$ in the infinite geometry. The main difference, apart from the replacement of integrals by discrete sums in Feynman diagrams, is that here the average

$$\varphi = L^{-d} \int \phi(x) d^d x = (2\pi/L)^d \tilde{\phi}(0),$$

remains a fluctuating variable (see also the discussion of Section 7.11).

As an illustration, we calculate expectation values of the form (32.8), moments of the distribution of the average spin per unit volume in a spin system. We set

$$\exp[-\Sigma(\varphi)] = \mathcal{N}^{-1} \int [d\chi] \exp[-\mathcal{S}(\varphi + \chi)], \quad (32.25)$$

where $\mathcal{S}(\phi)$ is the action (32.1) and the normalization \mathcal{N} is chosen such that $\Sigma(0) = 0$, for $t = 0$. Moments then are given by

$$m_\sigma = \mathcal{Z}^{-1} \int d\varphi |\varphi|^\sigma \exp[-\Sigma(\varphi)], \quad (32.26)$$

where \mathcal{Z} is the partition function,

$$\mathcal{Z} = \int d\varphi \exp[-\Sigma(\varphi)].$$

From the discussion of Section 7.11, it follows that, in the infinite volume limit, $\Sigma(\varphi) = \Gamma(\varphi) - \Gamma(0)$, where $\Gamma(\varphi)$ is the thermodynamic potential as obtained in perturbation theory. It satisfies the same RG equation:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \eta(g) \varphi \frac{\partial}{\partial \varphi} - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] \Sigma(\varphi, \tau, g, L, \Lambda) = R(\tau, g, \Lambda), \quad (32.27)$$

where R has been shown to be a second degree polynomial in τ (this follows from the equation for the specific heat (17.22)).

The moments m_{2s} , with s integer, involve only powers of φ^2 and thus are related to zero momentum correlation functions:

$$\varphi^2 = L^{-2d} \int d^d x d^d y \phi(x) \cdot \phi(y) = (2\pi/L)^{2d} \tilde{\phi}^2(0). \quad (32.28)$$

Although for generic values of σ this is no longer the case, m_σ nevertheless satisfies a usual RG equation, which can easily be derived from equation (32.27):

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \frac{1}{2} \sigma \eta(g) - \eta_2(g) \tau \frac{\partial}{\partial \tau} \right] m_\sigma(\tau, g, L, \Lambda) = 0. \quad (32.29)$$

Another quantity of interest is the specific heat

$$C(L, \tau) = L^{-d} \frac{\partial^2}{(\partial \tau)^2} \ln \mathcal{Z}. \quad (32.30)$$

32.3.1 Dimension $d > 4$

We first discuss the case $d > 4$, which in the infinite geometry is simple. To calculate the zero-mode integral, at leading order we neglect the fluctuations of all non-zero modes ($\chi = 0$). The function (32.25) then reduces to

$$\Sigma(\varphi) = \mathcal{S}(\varphi) = L^d \left(\frac{1}{2} \tau \varphi^2 + \frac{1}{4!} u (\varphi^2)^2 \right), \quad (32.31)$$

where \mathcal{S} is the action (32.1). After the change of variables

$$\varphi \mapsto (uL^d)^{-1/4} \varphi, \quad (32.32)$$

the moments m_σ take the form

$$m_\sigma(L, \tau) = (uL^d)^{-\sigma/4} \mu_\sigma(\tau L^{d/2} u^{-1/2}), \quad (32.33)$$

in which $\mu_\sigma(v)$ is given by

$$\mu_\sigma(v) = \frac{g_{\sigma+N}(v)}{g_N(v)}, \quad (32.34)$$

with

$$g_\sigma(v) = \int_0^\infty d\varphi \varphi^{\sigma-1} \exp \left[- \left(\frac{1}{2} v \varphi^2 + \frac{1}{4!} \varphi^4 \right) \right]. \quad (32.35)$$

Equation (32.33) shows that, above four dimensions, the finite-size scaling relations, proven for a non-trivial fixed point, and which, by contrast predict for the moment m_σ the behaviour

$$m_\sigma = L^{-\sigma(d-2)/2} \tilde{\mu}_\sigma(\tau L^2),$$

do not hold. In particular, in place of the argument $\tau L^2 \propto (L/\xi)^2$, one finds $\tau L^2 L^{(d-4)/2}$. The extra factor $L^{(d-4)/2}$ arises because the leading order result depends explicitly on u , which has a dimension $(4-d)$, and characterizes the violation of the naive scaling (see the question of hyperscaling in Section 17.1).

The result (32.33) leads to universal predictions. For instance, dimensionless ratios of moments like

$$\mathcal{R}_\sigma(\tau) = m_\sigma / (m_2)^{\sigma/2}, \quad (32.36)$$

at $T = T_c$ ($\tau = 0$) are universal. Calculating explicitly the integral (32.35) at $v = 0$,

$$g_\sigma(0) = \frac{1}{4} (24)^{\sigma/4} \Gamma(\sigma/4),$$

one finds

$$\mathcal{R}_\sigma(0) = \Gamma((\sigma + N)/4) [\Gamma(N/4)]^{\sigma/2-1} [\Gamma((N+2)/4)]^{-\sigma/2}. \quad (32.37)$$

In particular, for Ising-like systems, the quantity \mathcal{R}_4 is

$$\mathcal{R}_4 = \Gamma^4(1/4)/(8\pi^2) = 2.1884 \dots . \quad (32.38)$$

Another quantity of interest is

$$\mathcal{R}_1^{-2} = m_2/m_1^2 = \sqrt{2}. \quad (32.39)$$

At leading order, the specific heat C (equation (32.30)) is given by

$$C(L, \tau) = L^d (m_4 - m_2^2) / 4 = (\mu_4 - \mu_2^2) / 4u + \text{const.},$$

where the constant comes from the regular part of the free energy.

Remarks. The separation of the zero mode makes a priori only sense when the correlation length is much larger than the system size, that is, $\tau L^2 \ll 1$. When the scaling variable $\tau L^{d/2}$, which appears in expression (32.33), is finite the condition is satisfied for $d > 4$. One verifies that, in addition, the expression (32.33) has both for $\tau < 0$ and $\tau > 0$ fixed the correct $L \rightarrow \infty$ behaviour. Indeed, for $\tau < 0$, one finds

$$m_\sigma(L, \tau) \rightarrow (-6\tau/u)^{\sigma/2} \equiv [M_0(\tau)]^\sigma,$$

where $M_0(\tau)$ is the infinite-size spontaneous magnetization at this order. By contrast, for $\tau > 0$, one obtains

$$m_\sigma(L, \tau) \propto (\chi(\tau)L^{-d})^{\sigma/2},$$

where $\chi(\tau) = 1/\tau$ is the infinite-size magnetic susceptibility.

However, the correction terms to this leading behaviour are incorrect. Their determination involves additional higher order contributions.

32.3.2 Higher order corrections

It remains to show that higher order contributions do not invalidate the leading order results. The contributions to $\Sigma(\varphi)$ coming from loop diagrams can be split into two classes which have different properties: parts divergent for large cut-off, which are the same as in the infinite volume limit, and the remaining finite contributions.

Remarks

(i) According to the analysis of Chapter 17, divergences can be cancelled by adding terms local in ϕ to the action $\mathcal{S}(\phi)$. This means that a summation of the divergent terms adds to $\mathcal{S}(\phi)$ all possible monomials in ϕ and its derivatives, with coefficients proportional to powers of Λ dictated by dimensional analysis. After the substitution $\phi(x) = \varphi$, and an expansion in powers of τ , the divergent terms yield a contribution $\delta\Sigma_{\text{div.}}(\varphi, \tau)$ to Σ of the form

$$\delta\Sigma_{\text{div.}}(\varphi, \tau) = L^d \sum_{l,m} \Sigma_{lm} \Lambda^{d-2m-l(d-2)} \tau^m \varphi^{2l},$$

where the coefficients Σ_{lm} are numbers. After the change of variables (32.32), and taking into account that τ is of order $L^{-d/2}$ (equation (32.33)), one finds that a term proportional to $\varphi^{2l}\tau^m$ gives a contribution of order $(\Lambda L)^{d(2-m-l)/2}$. Therefore, only terms with $m+l \leq 2$ survive. The terms $l > 0$ shift r_c and yield finite renormalizations of τ and u . The two remaining terms are proportional to τ and τ^2 . They cancel in the moments (32.26). The τ^2 term yields a constant non-universal contribution to the specific heat. Finally, the leading corrections ($l+m=3$) are of order $(\Lambda L)^{-d/2}$.

(ii) After renormalization, the loop corrections can be formally expanded in powers of τ and φ^2 , because the size L provides an IR cut-off, and the zero mode has been removed. Because all contributions are UV finite, the coefficients are proportional to powers of L given by dimensional analysis.

First, we examine the one-loop corrections, generated in equation (32.25) by an integration over χ , in the Gaussian approximation,

$$\begin{aligned}\Sigma_{\text{1 loop}}(\varphi, L, \tau, u) &= \frac{1}{2} \sum_{k \neq 0} \text{tr} \ln \left[\delta_{ij} + \frac{1}{(2\pi k/L)^2} \left(\left(\tau + \frac{u}{6} \varphi^2 \right) \delta_{ij} + \frac{u}{3} \varphi_i \varphi_j \right) \right] \\ &= \frac{1}{2} \sum_{k \neq 0} \left[\ln \left(1 + \frac{\tau + \frac{1}{2} u \varphi^2}{(2\pi k/L)^2} \right) + (N-1) \ln \left(1 + \frac{\tau + \frac{1}{6} u \varphi^2}{(2\pi k/L)^2} \right) \right].\end{aligned}\quad (32.40)$$

After renormalization, from dimensional analysis one infers that, in the expansion of $\Sigma_{\text{1 loop}}$, the coefficient of $\tau^m (u \varphi^2)^l$ is proportional to L^{2m+2l} . After the change of variables (32.32), taking into account that τ is of order $u^{1/2} L^{-d/2}$ (equation (32.33)), one obtains a contribution proportional to $\tau^m (\varphi^2)^l \propto (u L^{4-d})^{(l+m)/2}$ which, for $d > 4$, goes to 0 for ΛL large.

The latter argument can be generalized to higher orders in the loop expansion. Each loop in the perturbative expansion yields a factor u , which is the loop expansion parameter, and thus a factor $u L^{4-d}$ for dimensional reasons. It follows that a term of ℓ -loop order proportional to $\varphi^{2l} \tau^m$ is of order $(u L^{4-d})^\kappa$, with $\kappa = \ell - 1 + \frac{1}{2}(l+m)$.

The conclusion is that, for $d > 4$, the effective action (32.31) can be simply derived from mean-field theory, as in the infinite volume limit, the only modification coming from the last integration over the average field (32.26). In particular, the result (32.37) indeed is universal.

At two-loop order, the leading term has $\kappa = 3/2$ and, thus, the only meaningful corrections at one-loop order correspond to $m+l \leq 2$. A short calculation shows that, for what concerns the dimensionless ratios (32.36), for $d < 8$, the leading corrections can be reproduced by replacing τ by a quantity τ_L , which has the form

$$\tau_L = (1 + A_1 u L^{4-d}) \tau + A_2 u L^{2-d}, \quad \text{with } A_1, A_2 \text{ constants.}$$

The specific heat receives an additional contribution of order L^{4-d} .

32.3.3 Dimension $d = 4 - \varepsilon$

We now use the RG arguments of Section 32.1: in the critical domain, instead of calculating physical quantities as function of $\{\tau, M, u, L\}$, we can set $\Lambda = L = 1$, $u = u^*$ (the IR fixed point value), then replace τ by $\tau L^{1/\nu}$, M (in the presence of a magnetic field) by $ML^{\beta/\nu}$, and thus φ by $\varphi L^{\beta/\nu}$ in $\Sigma(\varphi)$.

At leading order, the function $\Sigma(\varphi)$ is

$$\Sigma(\varphi, L = 1, \tau, u^*) = \frac{1}{2} \tau \varphi^2 + \frac{1}{4!} u^* (\varphi^2)^2.$$

At the same order, u^* can be replaced by its value at order ε :

$$u^* = \frac{48\pi^2 \varepsilon}{N+8} + O(\varepsilon^2). \quad (32.41)$$

Then, replacing τ by $\tau L^{1/\nu}$, and φ by $\varphi L^{\beta/\nu}$, and integrating over φ , one finds the moments m_σ at leading order,

$$m_\sigma(L, \tau) = \frac{L^{-\sigma(d-2+\eta)/2}}{(u^*)^{\sigma/4}} \mu_\sigma(\tau L^{1/\nu} (u^*)^{-1/2}), \quad (32.42)$$

where μ_σ has been defined in equation (32.34). The equation shows that the ε -expansion is not uniform. The method used here, in which the zero mode is treated separately, gives the correct leading order only if τ is formally assumed to be of order $\varepsilon^{1/2}$ (this condition is realized in particular at $\tau \propto T - T_c = 0$).

Note the appearance of powers of $(u^*)^{1/2}$ which, for ε small, is equivalent to $\varepsilon^{1/2}$. This indicates that physical quantities have an expansion in powers of $\varepsilon^{1/2}$, instead of ε . The analysis of higher order corrections confirms this observation, as a one-loop calculation illustrates.

One-loop calculation. The one-loop contribution $\Sigma_{1\text{loop}}$ to Σ is given by equation (32.40). The divergent part of the one-loop term has first to be subtracted. We can then set $L = 1$, $u = u^*$. As we have already pointed out in Section 32.3.2 (remark (ii)), at L fixed, all terms in perturbation theory can then be expanded in powers of φ^2 and τ . After the change of variables (32.32), φ^2 has a coefficient proportional to $(u^*)^{1/2} \sim \varepsilon^{1/2}$. In the same way, τ is of order $\varepsilon^{1/2}$. A term contributing to the ℓ -loop order and proportional to $\varphi^{2l}\tau^m$ is of order $\varepsilon^{\ell-1+(l+m)/2}$. The leading two-loop correction comes from the term proportional to φ^2 , and thus is of order $\varepsilon^{3/2}$. Therefore, at the one-loop order, only the terms proportional to φ^2 , $\varphi^2\tau$, $(\varphi^2)^2$, τ , and τ^2 have to be taken into account. The form of $\Sigma(\varphi)$ then is

$$\Sigma(\varphi, L = 1, \tau, u^*) - \Sigma(0, 1, \tau, u^*) = \frac{1}{2} (\tau(1 + a_1 u^*) + a_2 u^*) \varphi^2 + u^*(1 + a_3 u^*) \frac{(\varphi^2)^2}{4!}.$$

Note that the correction to the coefficient of $(\varphi^2)^2$ can be eliminated by a rescaling of φ . The only relevant effect is to change a_1 into $a_1 - a_3/2$. But the complete coefficient of τ can be absorbed into a finite change of normalization of τ . The conclusion is that, at one-loop order, the only relevant correction is related to the coefficient a_2 , which we now calculate.

The φ^2 contribution. The coefficient \tilde{a}_2 of $u^* \varphi^2 / 2$ in the expansion of expression (32.40) is

$$\tilde{a}_2 = \frac{N+2}{6} \sum_{k \in \mathbb{Z}^d \neq 0} \frac{1}{(2\pi k)^2}. \quad (32.43)$$

The coefficient a_2 is obtained from \tilde{a}_2 by subtracting the infinite-size contribution, which is a shift of the critical temperature.

As shown in Section A32.1, in terms of the function (related to Jacobi elliptic functions),

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} e^{-\pi s n^2}, \quad (32.44)$$

we can formally express the sum in equation (32.43) as

$$\sum_{k \neq 0} \frac{1}{(2\pi k)^2} = \frac{1}{4\pi} \int_0^{+\infty} ds (\vartheta_0^d(s) - 1). \quad (32.45)$$

The integral in the right-hand side converges exponentially for s large. For $s \rightarrow 0$, one finds (Section A32.1)

$$\vartheta_0(s) - (1/s)^{1/2} \underset{s \rightarrow 0}{\sim} 2s^{-1/2} e^{-\pi/s}, \quad (32.46)$$

where $s^{-1/2}$ corresponds to the infinite-size limit. The integral thus diverges at $s = 0$. Subtracting the infinite-size limit, for $d > 2$, one obtains the finite result,

$$\sum_{k \neq 0} \frac{1}{(2\pi k)^2} - \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2} = \frac{1}{4\pi} \int_0^{+\infty} ds \left(\vartheta_0^d(s) - 1 - s^{-d/2} \right). \quad (32.47)$$

The coefficient a_2 follows:

$$a_2 = \frac{N+2}{24\pi} \int_0^{+\infty} ds \left(\vartheta_0^d(s) - 1 - s^{-d/2} \right). \quad (32.48)$$

Introducing rescaled temperature $\tau' \propto \tau$ and field ϕ , we can rewrite the moments at one-loop order as

$$m_\sigma(L, \tau') = L^{-\sigma(d-2+\eta)/2} \mu_\sigma(\tau' L^{1/\nu} + b), \quad (32.49)$$

where the constant b is given by equations (32.41) and (32.48):

$$b = a_2 u^{*1/2} = \frac{N+2}{\sqrt{N+8}} \frac{(3\varepsilon)^{1/2}}{6} \int_0^{+\infty} ds \left[\vartheta_0^4(s) - 1 - 1/s^2 \right] + O\left(\varepsilon^{3/2}\right). \quad (32.50)$$

The ratio $\mathcal{R}_4(0)$ ($T = T_c$). From expression (32.49), setting $t = 0$, one immediately derives the universal dimensionless ratio $\mathcal{R}_4(0)$ at order $\varepsilon^{1/2}$:

$$\mathcal{R}_4(0) = \frac{g_{4+N}(b)g_N(b)}{\left[g_{2+N}(b)\right]^2}, \quad (32.51)$$

g_σ being defined by equation (32.35).

Using the value of the integral

$$\int_0^{+\infty} ds \left[\vartheta_0^4(s) - 1 - 1/s^2 \right] = -0.561843942 \dots,$$

in three dimensions, for $N = 1$, one obtains

$$\mathcal{R}_4(0) = 1.800 \dots. \quad (32.52)$$

This result can be compared to the mean-field value 2.188 and a Monte Carlo numerical estimate 1.6. The agreement is comparable to other results at order ε .

The specific heat. The τ^2 terms, both divergent and convergent in Σ have to be taken into account. A short calculation yields

$$C(L, \tau) = L^{\alpha/\nu} \left[\frac{1}{4} \left(\mu_4(\tau' L^{1/\nu} + b) - \mu_2^2(\tau' L^{1/\nu} + b) \right) + \frac{3N}{4-N} \right] + \text{const.},$$

where the constant comes from the non-universal regular contribution.

32.4 The ϕ^4 field theory: Cylindrical geometry

In what follows, we consider a system infinite in one dimension hereafter called (Euclidean) time, and of finite-size L with periodic boundary conditions in the remaining $(d - 1)$ space dimensions. To isolate the zero modes, we expand the fields in Fourier components in the $(d - 1)$ space dimensions:

$$\begin{aligned}\phi(z, x) &= \varphi(z) + \chi(z, x), \quad x \in \mathbb{Z}^{d-1}, \\ \chi(z, x) &= (2\pi/L)^{d-1} \sum_{k \in \mathbb{Z}^{d-1} \neq 0} e^{i2\pi k \cdot x / L} \phi_k(z).\end{aligned}\quad (32.53)$$

Again, we consider only the example of correlation functions of the space integral

$$\varphi(z) = L^{1-d} \int d^{d-1}x \phi(x, z).$$

These can be calculated using only the effective action $\mathcal{S}_L(\varphi)$ obtained by integrating over χ ,

$$\exp[-\mathcal{S}_L(\varphi)] = \mathcal{N}^{-1} \int [d\chi] \exp[-\mathcal{S}(\varphi + \chi)],$$

where the normalization \mathcal{N} is now chosen in such a way that $\mathcal{S}_L(\varphi = 0, t = 0) = 0$. The partition function \mathcal{Z} and φ -field correlation functions have simple path integral representations of quantum mechanics (QM) type corresponding to the effective action \mathcal{S}_L .

We illustrate the method with the calculation of the finite-size correlation length ξ_L .

32.4.1 Correlation length: Dimensions $d > 4$

For $d > 4$, in the infinite volume limit, the correlation length can be derived from mean-field theory. The finite-size correlation length ξ_L can be derived from a path integral corresponding to an action, which at leading order is obtained by neglecting all corrections due the integration over the $k \neq 0$ components of the field ($\dot{\varphi} \equiv d\varphi/dz$):

$$\mathcal{S}_L(\varphi) = \mathcal{S}(\varphi) = L^{d-1} \int dz \left[\frac{1}{2} (\dot{\varphi}(z))^2 + \frac{1}{2} \tau \varphi^2(z) + \frac{1}{4!} u (\varphi^2(z))^2 \right]. \quad (32.54)$$

To simplify the action, we rescale φ and time z ,

$$\varphi \mapsto u^{-1/6} L^{(1-d)/3} \varphi, \quad z \mapsto u^{-1/3} L^{(d-1)/3} z. \quad (32.55)$$

For the Fourier (energy) variable E associated with z , the rescaling corresponds to

$$E \mapsto L^{-(d-1)/3} u^{1/3} E. \quad (32.56)$$

The action becomes,

$$\mathcal{S}'_L(\varphi) = \int dz' \left[\frac{1}{2} (\dot{\varphi}(z))^2 + \frac{1}{2} u^{-2/3} L^{2(d-1)/3} \tau \varphi^2(z) + \frac{1}{4!} (\varphi^2(z))^2 \right]. \quad (32.57)$$

The action corresponds to the quantum Hamiltonian (see Chapter 2)

$$\mathbf{H} = \frac{1}{2} \mathbf{p}^2 + \frac{1}{2} v \mathbf{q}^2 + \frac{1}{4!} (\mathbf{q}^2)^2, \quad (32.58)$$

where \mathbf{q} and \mathbf{p} are the position and momentum operators, and $v = u^{-2/3} L^{2(d-1)/3} \tau$.

The finite-size correlation length ξ_L is related to the two lowest eigenvalues E_0 and E_1 of \hat{H} rescaled according to the transformation (32.56):

$$\xi_L = u^{-1/3} L^{(d-1)/3} X(\tau L^{2(d-1)/3} u^{-2/3}), \quad (32.59)$$

with

$$X(v) = (E_1(v) - E_0(v))^{-1}. \quad (32.60)$$

Introducing the infinite-size correlation length ξ_∞ , one can rewrite the expression as

$$\xi_L(t) = \xi_\infty(t) Y(L^{(d-4)/3} L/\xi_\infty(t)), \quad Y(v) = \sqrt{v} X(v). \quad (32.61)$$

The expression exhibits a violation of the naive extension of the scaling form (32.11), proven for $d < 4$. Here again, the reason is that, in case of the Gaussian fixed point, the amplitude of expected leading terms vanishes, and the first non-trivial contributions correspond to what, for $d < 4$, are only corrections to scaling (see Section 17.1).

Finally, we observe that, in the limit of interest, t has to be taken of order $u^{2/3} L^{2(1-d)/3}$.

Loop corrections. We now show that loop corrections due to the integration over the non-zero modes do not modify the leading order of the scaling form (32.61). The corrections are of two types:

(i) Corrections corresponding to contributions divergent with the cut-off, and already present in the infinite volume limit. They can be cancelled by adding, at any finite order, local polynomials to the initial action $\mathcal{S}(\phi)$. This yields a correction of the form

$$\delta\mathcal{S}_{L\text{div.}}(\varphi) = L^{d-1} \int dz \mathcal{L}(\varphi, z),$$

where $\mathcal{L}(\varphi, z)$ is a sum of monomials in $\varphi(z)$ and its derivatives.

After the changes of variables (32.55), a monomial with $2k$ powers of φ and $2n$ derivatives is of order L^κ , with $\kappa = 2(d-1)(2-k-n)/3$. The terms with $k+n \leq 2$ simply renormalize the coefficients of the leading order action. All other terms, which are absent from the leading order action, have coefficients suppressed for L large. The simplest example is ϕ^6 , which induces a correction proportional to $L^{-2(d-1)/3} \hat{q}^6$ in \hat{H} .

(ii) Corrections that are finite in the large cut-off limit. The one-loop correction $\mathcal{S}_{L\text{1loop}}(\varphi)$ to the effective action, analogous to expression (32.40), in the initial φ, z variables, is given by ($\partial_z \equiv \partial/\partial z$)

$$\begin{aligned} \mathcal{S}_{L\text{1loop}}(\varphi) = & \frac{1}{2} \sum_{k \neq 0} \left\{ \text{tr} \ln \left[1 + (-\partial_z^2 + (2\pi k/L)^2)^{-1} (\tau + u\varphi^2(z)/2) \right] \right. \\ & \left. + (N-1) \text{tr} \ln \left[1 + (-\partial_z^2 + (2\pi k/L)^2)^{-1} (\tau + u\varphi^2(z)/6) \right] \right\}. \end{aligned} \quad (32.62)$$

After renormalization, it yields a UV-finite non-local contribution. Therefore, at first sight, it would seem that we can no longer use a Hamiltonian formalism to evaluate corrections to the mean-field approximation.

However, we first note that φ appears only in the combination $uL^2\varphi^2$. After the change (32.55), it becomes proportional to $u^{2/3} L^{2(4-d)/3} \varphi^2$, which goes to zero for L large. In the same way, τ appears only in the combination τL^2 which, according to the form (32.59), is of order $u^{2/3} L^{2(4-d)/3}$, and thus also goes to zero for L large. Therefore, because L provides an IR cut-off we can expand in powers of φ and τ .

Finally, we face the problem of the non-local operator $[-L^2\partial_z^2 + (2\pi k)^2]^{-1}$. After the change of variables (32.55), the differential operator $L^2\partial_z^2$ becomes also of order $u^{2/3}L^{2(4-d)/3}$, and thus is small compared to k^2 . We then know from general arguments that the action (32.62) has a local expansion. Actually, this problem is related to the classical problem of the high-energy expansion of the resolvent of the Schrödinger operator:

$$\text{tr} [-(d/dq)^2 + U(q) - E]^{-1} = \sum_{n=0} E^{-1/2-n} \Pi_n(U),$$

where one derives that the $\Pi_n(U)$ are local polynomials in the potential U .

This implies that, at one-loop order, for L large, the effective action \mathcal{S}_L has a local expansion. This argument can be generalized to higher orders in the loop expansion and, therefore, $\mathcal{S}_L(\varphi)$ has a local expansion to any finite order.

Moreover, we have found a kind of power counting property,

$$\varphi^2(z) \sim \tau \sim -\partial_z^2 \sim (uL^{4-d})^{2/3}. \quad (32.63)$$

In addition, each new loop carries a factor uL^{4-d} . Therefore, at loop-order l , a term with $2k$ powers of φ , m powers of τ and $2n$ derivatives is of order $(uL^{4-d})^\kappa$, with $\kappa = l + 2(k + m + n - 2)/3$. For $d > 4$, all loop corrections have a negative power of L . The dominant correction for large L comes from the term proportional to $\int dz \varphi^2(z)$, which vanishes like $L^{-2(d-4)/3}$.

Remark. Here again, although the result (32.59) has been proven only in the scaling regime $\tau L^{2(d-1)/3}$ finite, it has the correct large argument behaviour. For $\tau > 0$,

$$X(v) \underset{v \rightarrow +\infty}{\sim} v^{-1/2} \Rightarrow \xi_L \sim \tau^{-1/2}.$$

For $\tau < 0$, the behaviour changes drastically, depending on whether the symmetry is continuous or discrete, as we discuss in detail in Sections 32.5.2 and A32.2.2. For $N > 2$ and $v \rightarrow -\infty$, we note that the lowest eigenvalues of the Hamiltonian (32.58) can be obtained by approximating \hat{H} by the angular moment part, fixing the radial coordinate at $|q| = \sqrt{-6v}$. The corresponding eigenvalues are $\ell(\ell + N - 2)/(-12v)$. It follows that

$$X(v) \underset{v \rightarrow -\infty}{\sim} -\frac{12v}{N-1} \Rightarrow \frac{\xi_L}{L} \sim -\frac{12}{u} \frac{\tau}{N-1} L^{d-2} \propto \frac{L^{d-2}}{\xi_\infty^2(\tau)},$$

a result that can be compared with the result (32.94).

By contrast, for $N = 1$, instantons are responsible for the splitting between the two lowest-lying states (Chapter 39). A Wentzel–Kramers–Brillouin (WKB) analysis yields

$$\ln X(v) \sim 2(-2v)^{3/2} \Rightarrow \ln \xi_L \sim 2(-2\tau)^{3/2} \frac{L^{d-1}}{u} = \frac{2}{u} \frac{L^{d-1}}{\xi_\infty^3(\tau)}, \quad (32.64)$$

a result that can be compared with the result (A32.15).

32.4.2 Dimensions $d = 4 - \varepsilon$

For $d = 4 - \varepsilon$, at leading order, we can replace u by its IR fixed point value u^* , which is of order ε . Using the result (32.59), and the RG scaling (32.11), one finds

$$\xi_L/L = (u^*)^{-1/3} X[\tau L^{1/\nu}/(u^*)^{2/3}]. \quad (32.65)$$

We then use the preceding considerations to analyse the leading corrections for ε small. We have seen that the expansion parameter is actually uL^{4-d} , which, for $d > 4$ is small because L is large, while here it is small, because u is of order ε . Therefore, the power counting argument given previously transforms into an argument about the powers of ε . An interaction term generated at loop order l with m powers of τ , $2k$ fields ϕ , and $2n$ time derivatives is multiplied by ε^κ , with $\kappa = l + 2(k + m + n - 2)/3$. The leading one-loop term is thus proportional to $\int dz \varphi^2(z)$ and of order $\varepsilon^{1/3}$. Note that the leading two-loop correction, which we neglect in this calculation is of order $\varepsilon^{4/3}$. Since the one-loop contribution to the two-point function is a constant, no term proportional to $\dot{\varphi}^2(z)$ is generated.

The one-loop terms proportional to $\int dz (\varphi^2(z))^2$ and $t \int dz \varphi^2(z)$ are of order ε . The $\int dz (\varphi^2(z))^3$ term and the term with two derivatives, coming from the four-point function, are of order $\varepsilon^{5/3}$, and can be neglected.

From now on, the discussion closely follows the lines of Section 32.3. We call $a_1 u^*$ the coefficients of $t \int dz \varphi^2(z)$, $a_2 u^*$ of $\int dz \varphi^2(z)$ and $a_3 u^{*2}$ of $\int dz (\varphi^2(z))^2$ in the expansion of the expression (32.62). An addition of these contributions to the effective action simply amounts to the substitutions

$$\tau \mapsto \tau(1 + a_1 u^*) + a_2 u^*, \quad u^* \mapsto u^*(1 + a_3 u^*).$$

After these substitutions into equation (32.65), one finds the finite-size correlation length at one-loop order

$$\xi_L(\tau') = L (u^* + a_3 u^{*2})^{-1/3} X(\tau' L^{1/\nu} + b), \quad (32.66)$$

where τ' is a renormalized temperature, and

$$b = a_2 u^* (u^* + a_3 u^{*2})^{-2/3} = a_2 \left(\frac{48\pi^2 \varepsilon}{N+8} \right)^{1/3} + O(\varepsilon^{4/3}), \quad t' = t [1 + (a_1 - \frac{2}{3} a_3) u^*] / u^{*2/3}.$$

As in the case of the hypercubic geometry, we observe that the contribution coming from a_3 , which is of the same order as the two-loop contribution ($\varepsilon^{4/3}$), is negligible at this order. Thus, we need only the coefficient \tilde{a}_2 of $u^* \int dz \varphi^2(z)/2$.

The coefficient of φ^2 . The coefficient \tilde{a}_2 of $u^* \int dz \varphi^2(z)/2$ in the expansion of the expression (32.62) is

$$\tilde{a}_2 = \frac{N+2}{6} \int \frac{d\omega}{2\pi} \sum_{k \in \mathbb{Z}^{d-1} \neq 0} \frac{1}{\omega^2 + (2\pi k)^2}. \quad (32.67)$$

We integrate over ω . We then use the formal identity

$$\frac{1}{2} \sum_{k \neq 0} \frac{1}{|2\pi k|} = \frac{1}{4\pi} \int_0^{+\infty} \frac{ds}{\sqrt{s}} (\vartheta_0^{d-1}(s) - 1), \quad (32.68)$$

in which the function $\vartheta_0(s)$ is defined by equation (32.44).

The integral has a divergence for s small, which, as we have seen in Section 32.3.3 is cancelled by the critical temperature shift. We then find

$$a_2 = \frac{N+2}{24\pi} \int_0^{+\infty} \frac{ds}{\sqrt{s}} \left[\vartheta_0^3(s) - 1 - s^{-3/2} \right] + O(\varepsilon). \quad (32.69)$$

The finite-size correlation length at T_c . In particular, substituting the value of a_2 into equation (32.66), we obtain the finite-size correlation length at T_c at one-loop order:

$$\frac{\xi_L}{L} = \left(\frac{48\pi^2\varepsilon}{N+8} \right)^{-1/3} X \left[K\pi^{-1/3} \frac{N+2}{12} \left(\frac{6\varepsilon}{N+8} \right)^{1/3} \right] (1 + O(\varepsilon)), \quad (32.70)$$

with

$$K = \int_0^{+\infty} \frac{ds}{\sqrt{s}} \left[\vartheta_0^3(s) - 1 - s^{-3/2} \right] = -2.8372974 \dots . \quad (32.71)$$

32.5 Finite size effects in the non-linear σ -model

In models with continuous symmetries, in the ordered phase, the propagator corresponding to the Goldstone modes has a pole at zero momentum. In a finite volume, or in a cylindrical geometry with periodic boundary conditions, the zero mode leads to IR divergences. Therefore, it is necessary to separate the zero mode and treat it non perturbatively. Taking the example of the $O(N)$ -symmetric non-linear σ -model (see Chapter 19 for details), we examine again the two examples of the hypercubic and cylindrical geometries. The partition function reads (equations (19.3, 19.4))

$$\mathcal{Z} = \int [d\phi(x)] \prod_x \delta(\phi^2(x) - 1) \exp[-\mathcal{S}(\phi)/t], \text{ with } \mathcal{S}(\phi) = \frac{1}{2} \int d^d x (\nabla \phi(x))^2, \quad (32.72)$$

where $t > 0$ is the coupling constant, and ϕ an N -component scalar field, which, when needed, we parametrize as

$$\phi(x) = \{ \sqrt{1 - \pi^2(x)}, \pi(x) \}. \quad (32.73)$$

32.5.1 The hypercubic geometry

Here, the zero mode can be associated with the set of collective coordinates which parametrize the degenerate classical minima of the potential. The integration over the zero mode is then equivalent to an average over all directions, which restores the $O(N)$ symmetry broken by the choice of a classical minimum of the action, around which perturbation theory is expanded. In the case of $O(N)$ -invariant observables, the averaging is trivial, and the result is simply that the zero mode has to be omitted in the perturbative expansion. We recall that, for $d \geq 2$, these observables then have a finite limit when the size L becomes infinite. For illustration purpose, we first consider the second moment m_2 of the field distribution (equation (32.26)):

$$m_2 = L^{-d} \int d^d x \langle \phi(x) \cdot \phi(0) \rangle. \quad (32.74)$$

In terms of the functions (19.105) and (19.106), m_2 has the scaling form

$$m_2 = M_0^2(t) \mu_2(L/\xi(t)). \quad (32.75)$$

For $d > 2$, m_2 is finite at the critical temperature t_c , and thus

$$\mu_2(v) \underset{v \rightarrow 0}{\propto} v^{-(d-2+\eta)}.$$

It is convenient to set

$$L^{-2d} \int d^d x d^d y \phi(x) \cdot \phi(y) = 1 - 2\mathcal{O}. \quad (32.76)$$

Expanding in powers of the field π (equation (32.73)), one obtains

$$\mathcal{O} = \frac{1}{4} L^{-2d} \int d^d x d^d y [\pi(x) - \pi(y)]^2 + O(\pi^4). \quad (32.77)$$

The π -propagator is

$$\Delta_{ij}(x) = \delta_{ij} \Delta(x), \quad \text{with } \Delta(x) \equiv t \Lambda^{2-d} L^{-d} \sum_{k \neq 0} \frac{e^{i2\pi k \cdot x}}{(2\pi k)^2}. \quad (32.78)$$

The moment m_2 at one-loop order follows:

$$m_2 = 1 - 2 \langle \mathcal{O} \rangle = 1 + t(N-1)(\Lambda L)^{2-d} \int d^d x \sum_{k \neq 0} \frac{(e^{i2\pi k \cdot x/L} - 1)}{(2\pi k)^2} + O(t^2). \quad (32.79)$$

After integration over x , the expression becomes

$$m_2 = 1 - t(N-1)(\Lambda L)^{2-d} \sum_{k \neq 0} \frac{1}{(2\pi k)^2} + O(t^2), \quad (32.80)$$

a large momentum cut-off being implied. At this order, for $d > 2$, the cut-off dependence is removed by dividing m_2 by the square of the spontaneous magnetization M_0 :

$$M_0 = 1 - \frac{t \Lambda^{2-d}}{2(2\pi)^d} \int \frac{d^d p}{p^2} + O(t^2) = 1 - \frac{t \Lambda^{2-d}}{8\pi} \int_0^\infty ds s^{-d/2}. \quad (32.81)$$

As usual, we transform the sum over k into an integral involving the function (32.44). Taking into account equation (32.81), one finds for the scaling function (32.75)

$$\mu_2 = 1 - \frac{t(N-1)}{4\pi} (\Lambda L)^{2-d} \int_0^\infty ds \left(\vartheta_0^d(s) - 1 - s^{-d/2} \right) + O(t^2). \quad (32.82)$$

RG results imply that one can set $\Lambda L = 1$, and replace t by the effective coupling t_L at scale L , which at leading order at $d > 2$ is given by (equation (32.22))

$$t_L = t(\Lambda L)^{2-d} + O(t^2).$$

One then obtains

$$\mu_2 = 1 - \frac{t_L(N-1)}{4\pi} \int_0^\infty ds \left(\vartheta_0^d(s) - 1 - s^{-d/2} \right) + \dots,$$

an expression valid for $2 < d < 4$ and $t < t_c$ fixed. Since then t_L goes to 0 for L large, this expression exhibits how μ_2 approaches 1.

Using the estimate (32.20), one finds

$$\mu_2 - 1 \sim -\frac{N-1}{4\pi} \int_0^\infty ds \left(\vartheta_0^d(s) - 1 - s^{-d/2} \right) \left(\frac{\xi(t)}{L} \right)^{d-2}. \quad (32.83)$$

To generate universal quantities, we again consider the ratios (32.36) $R p_\sigma = m_\sigma / (m_2)^{\sigma/2}$. With the definition (32.76),

$$\left(L^{-d} \left| \int d^d x \phi(x) \right| \right)^\sigma = 1 - \sigma \mathcal{O} + \frac{\sigma(\sigma-2)}{2} \mathcal{O}^2 + \dots.$$

Therefore,

$$\mathcal{R}_\sigma - 1 \sim \frac{1}{2} \sigma(\sigma-2) \left[\langle \mathcal{O}^2 \rangle - (\langle \mathcal{O} \rangle)^2 \right].$$

Using equations (32.77) and (32.78), one infers

$$\begin{aligned} \mathcal{R}_\sigma &= 1 + \frac{1}{8} \sigma(\sigma-2)(N-1) L^{-2d} \int d^d x d^d y [\Delta(x) - \Delta(y)]^2, \\ &= 1 + \frac{1}{4} \sigma(\sigma-2)(N-1) \frac{t^2 (\Lambda L)^{4-2d}}{16\pi^2} \int_0^\infty ds s [\vartheta_0^d(s) - 1] + O(t^3). \end{aligned}$$

Using RG arguments, one can then set $\Lambda L = 1$ and replace t by t_L (equation (32.22)):

$$\mathcal{R}_\sigma = 1 + \sigma(\sigma-2) \frac{N-1}{4} \frac{t_L^2}{16\pi^2} \int_0^\infty ds s [\vartheta_0^d(s) - 1]. \quad (32.84)$$

For $d > 2$ and $t < t_c$ fixed, t_L goes to 0 for L large. Using the estimate (32.20), one thus finds how \mathcal{R}_σ goes to 1:

$$\mathcal{R}_\sigma = 1 + \sigma(\sigma-2) \frac{N-1}{4} \frac{1}{16\pi^2} \int_0^\infty ds s [\vartheta_0^d(s) - 1] \left(\frac{\xi(t)}{L} \right)^{2d-4}. \quad (32.85)$$

The neighbourhood of the critical temperature can only be studied within the $\varepsilon = d-2$ expansion. Setting $t_L = t_c$, one obtains the universal ratio \mathcal{R}_σ at order ε^2 :

$$\mathcal{R}_\sigma = 1 + \varepsilon^2 \sigma(\sigma-2) \frac{N-1}{16(N-2)^2} \int_0^\infty ds s [\vartheta_0^2(s) - 1]. \quad (32.86)$$

Induced magnetization in a small field (equation (19.19)). In a small magnetic field h , that is, such that hL^d is small, one can calculate the magnetization M at leading order. Since $\langle \phi(x) \rangle$ is not an $O(N)$ -invariant observable, the average over the zero mode is not trivial. RG equations predict (see also equation (19.117))

$$M(h, t, L) = M_0(t) m \left[\frac{h M_0(t)}{t} \xi^d(t), \frac{L}{\xi(t)} \right]. \quad (32.87)$$

At leading order, the partition function in a field is given by

$$\begin{aligned} \mathcal{Z}(h, t, L) &= \int d^N \boldsymbol{\varphi} \delta(\boldsymbol{\varphi}^2 - 1) \exp(L^d \mathbf{h} \cdot \boldsymbol{\varphi}/t) \\ &= \frac{1}{\pi} \int_0^\pi d\theta (\sin \theta)^{N-2} \exp(L^d h \cos \theta/t). \end{aligned} \quad (32.88)$$

The integral is a modified Bessel function. The magnetization is the logarithmic derivative of \mathcal{Z} . At this order, M depends only on the scaling variable

$$v = h M_0(t) L^d / t, \quad (32.89)$$

and is given by

$$M(h, t, L) = M_0(t) \frac{d}{dv} \ln \int_0^\pi d\theta (\sin \theta)^{N-2} e^{v \cos \theta}. \quad (32.90)$$

32.5.2 The cylindrical geometry

As an illustration, we now calculate the finite-size correlation length in a cylindrical geometry. At leading order, the action of the zero mode is the action of the $O(N)$ rigid rotator, which we have already discussed in Section 3.5:

$$\mathcal{S}_L(\varphi) = \frac{\Lambda^{d-2} L^{d-1}}{2t} \int dz \dot{\varphi}^2(z). \quad (32.91)$$

From equation (3.46), we infer the correlation length at leading order

$$\xi_L(t) = \frac{2}{N-1} \frac{\Lambda^{d-2} L^{d-1}}{t}. \quad (32.92)$$

Combining this result with the RG arguments of Section 32.1.2, we can rewrite this equation in terms of the effective coupling at scale L (equations (32.17) and (32.22)) as

$$\frac{\xi_L(t)}{L} \sim \frac{2}{(N-1)t_L}. \quad (32.93)$$

For $d > 2$, in the ordered phase, t_L goes to 0, and we obtain the scaling form

$$\frac{\xi_L(t)}{L} \underset{\text{for } 2 < d < 4}{\sim} \frac{2}{N-1} \left(\frac{L}{\xi_\infty(t)} \right)^{d-2}. \quad (32.94)$$

Immediately, we learn that, in a system in which a continuous symmetry is broken, at any temperature below t_c , the finite-size correlation length grows like L^{d-1} . This behaviour has to be contrasted with the behaviour of the correlation length in the case of discrete symmetry (see Section A32.2.2).

The critical temperature is an RG fixed point, and thus $t_L = t_c$. We then obtain the universal ratio ξ_L/L at leading order in $\varepsilon = d - 2$:

$$\frac{\xi_L(t_c)}{L} \underset{d \rightarrow 2}{\sim} \frac{N-2}{(N-1)\pi\varepsilon}.$$

Finally, in two dimensions, for $L/\xi_\infty(t)$ small, we find

$$\frac{\xi_L(t)}{L} \sim \frac{N-2}{(N-1)\pi} \ln(\xi_\infty(t)/L).$$

One-loop corrections. To calculate the one-loop corrections to expression (32.93), we evaluate the effective action for $\varphi(z)$ obtained by integrating over non-zero modes, at one-loop order. We define $\varphi(z)$ as the unit vector along the direction of the average of $\int d^{d-1}x \phi(z, x)$. We then parametrize $\phi(z, x)$ by a rotation $R(z)$ acting on a field $\phi'(z, x)$, whose average $\int d^{d-1}x \phi'(z, x)$ points towards a fixed direction. To simplify, instead of calculating for generic fields $\varphi(z)$, it is convenient to restrict φ to lie in a two-dimensional plane:

$$\varphi(z) = \{\cos(\alpha(z)), \sin(\alpha(z)), \mathbf{0}\}.$$

The action (32.91) then reduces to

$$\mathcal{S}_L(\alpha) = \frac{\Lambda^{d-2} L^{d-1}}{2t} \int dz \dot{\alpha}^2(z). \quad (32.95)$$

At order $\dot{\alpha}^2$, the mapping between the actions (32.91) and (32.95) is unambiguous.

We parametrize the field $\phi(z, x)$ as

$$\phi(z, x) = \begin{cases} \cos(\alpha(z))\sigma_1(z, x) - \sin(\alpha(z))\sigma_2(z, x), \\ \sin(\alpha(z))\sigma_1(z, x) + \cos(\alpha(z))\sigma_2(z, x), \\ \boldsymbol{\pi}(z, x), \end{cases} \quad (32.96)$$

in which the field $\boldsymbol{\pi}(z, x)$ here has only $(N - 2)$ components. The fields $\boldsymbol{\pi}$ and σ_2 have no zero mode. Since the transformation (32.96) is a rotation, the three fields σ_1 , σ_2 , and $\boldsymbol{\pi}$ still satisfy the constraint:

$$\sigma_1^2(x, z) + \sigma_2^2(x, z) + \boldsymbol{\pi}^2(x, z) = 1, \quad (32.97)$$

and the integration measure in the field integral is left invariant.

The action $\mathcal{S}(\phi)$ in the new fields reads

$$\begin{aligned} \mathcal{S}(\sigma_1, \sigma_2, \boldsymbol{\pi}) = & \frac{\Lambda^\varepsilon}{2t} \int dz d^{d-1}x \left[\dot{\alpha}^2 (\sigma_1^2 + \sigma_2^2) + \dot{\sigma}_1^2 + \dot{\sigma}_2^2 + \dot{\boldsymbol{\pi}}^2 \right. \\ & \left. + 2\dot{\alpha}(\dot{\sigma}_2\sigma_1 - \dot{\sigma}_1\sigma_2) + \sum_i ((\partial_i\sigma_1)^2 + (\partial_i\sigma_2)^2 + (\partial_i\boldsymbol{\pi})^2) \right]. \end{aligned} \quad (32.98)$$

To evaluate the effects of non-zero modes, we eliminate the field σ_1 using the constraint (32.97),

$$\sigma_1 = (1 - \sigma_2^2 - \boldsymbol{\pi}^2)^{1/2},$$

and expand the action in powers of σ_2 and $\boldsymbol{\pi}$. The quadratic part of the action $\mathcal{S}_2(\sigma_2, \boldsymbol{\pi})$, needed for the one-loop calculation, is then ($\partial_i \equiv \partial/\partial x_i$, $\dot{\alpha} \equiv d\alpha/dz\dots$)

$$\begin{aligned} \mathcal{S}_2(\sigma_2, \boldsymbol{\pi}) = & \frac{\Lambda^\varepsilon}{2t} L^{d-1} \int dz \dot{\alpha}^2(z) + \frac{\Lambda^\varepsilon}{2t} \int dz d^{d-1}x \left[-\dot{\alpha}^2(z)\boldsymbol{\pi}^2(x, z) + \dot{\sigma}_2^2(x, z) \right. \\ & \left. + \dot{\boldsymbol{\pi}}^2(x, z) + 2\dot{\alpha}(z)\dot{\sigma}_2(x, z) + (\partial_i\sigma_2(x, z))^2 + (\partial_i\boldsymbol{\pi}(x, z))^2 \right]. \end{aligned} \quad (32.99)$$

The term proportional to $\dot{\alpha}\dot{\sigma}_2$ which, after integration by parts, is equivalent to $\sigma_2\ddot{\alpha}$, gives, as leading contribution, a term of order $(\ddot{\alpha})^2$, which has four time derivatives, and can be neglected. At this order, the integration over σ_2 thus gives a factor independent of α , which can be absorbed into the normalization of the field integral. The integral over $\boldsymbol{\pi}$ yields a determinant to the power $(N - 2)/2$. Hence, the result at one-loop order is

$$\mathcal{S}_L(\alpha) = \frac{\Lambda^\varepsilon}{2t} L^{d-1} \int dz \dot{\alpha}^2(z) + \frac{1}{2}(N-2) \text{tr} \ln \left[(-\dot{\alpha}^2 - \partial_z^2 - \partial_i^2) (-\partial_z^2 - \partial_i^2)^{-1} \right]. \quad (32.100)$$

Neglecting terms with more than two derivatives, we obtain the one-loop contribution to the effective action by expanding to first order in $\dot{\alpha}^2$. This yields a renormalization of the coefficient of the leading term

$$\begin{aligned} \mathcal{S}_L(\alpha) &= \mathfrak{S}(t, \Lambda, L) \int dz \dot{\alpha}^2(z), \quad \text{with} \\ \mathfrak{S}(t, \Lambda, L) &= \frac{\Lambda^\varepsilon}{2t} L^{d-1} - \frac{N-2}{2} \sum_{k \in \mathbb{Z}^{d-1} \neq 0} \frac{1}{2\pi} \int \frac{d\omega}{\omega^2 + (2\pi k/L)^2}. \end{aligned} \quad (32.101)$$

The sum has to be understood with a large momentum cut-off. The cut-off dependence can be eliminated by subtracting the infinite-size limit. We use the identity ($d > 2$)

$$\begin{aligned} & \sum_{k \neq 0} \frac{1}{2\pi} \int \frac{d\omega}{\omega^2 + (2\pi k/L)^2} - \frac{L^{d-1}}{(2\pi)^d} \int^\Lambda \frac{d^d p}{p^2} \\ &= \int_0^\infty ds \frac{e^{-s\omega^2}}{2\pi} \left[\sum_{k \neq 0} e^{-4\pi^2 k^2 s/L^2} - \frac{L^{d-1}}{(2\pi)^{d-1}} \int d^{d-1} p e^{-sp^2} \right] \\ &= \frac{L}{4\pi} \int_0^\infty \frac{ds}{\sqrt{s}} \left[\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2} \right], \end{aligned}$$

where the change of variables $4\pi s/L^2 \mapsto s$, and the function (32.44) have been used. For dimensional reasons,

$$\int^\Lambda \frac{d^d p}{p^2} = -\frac{\tilde{\beta}_2(d)}{d-2} \Lambda^{d-2},$$

where $\tilde{\beta}_2$ is finite for $d \rightarrow 2$, but depends on the specific regularization.

Comparing with equation (32.92), we conclude

$$\begin{aligned} \frac{\xi_L(t)}{L} &= \frac{2}{N-1} \left[(L\Lambda)^{d-2} \left(\frac{1}{t} + \frac{\tilde{\beta}_2(d)}{d-2} \right) - \frac{N-2}{4\pi} \int_0^\infty \frac{ds}{\sqrt{s}} \left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2} \right) \right] \\ &\quad + O(t). \end{aligned}$$

Comparing with expression (32.22), we verify that $\tilde{\beta}_2 = \beta_2$ (β_2 is the coefficient of order t^2 of the RG β function), and express ξ_L/L in terms of the size-dependent coupling constant t_L ,

$$\begin{aligned} \frac{\xi_L(t)}{L} &= \frac{2}{N-1} \left[\frac{1}{t_L} + \frac{\beta_2(d)}{d-2} - \frac{N-2}{4\pi} \int_0^\infty \frac{ds}{\sqrt{s}} \left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2} \right) \right] \\ &\quad + O(t_L). \end{aligned} \tag{32.102}$$

We now specialize the expression to several situations. For $d > 2$ and $t < t_c$ fixed, or $d = 2$, $t \rightarrow 0$, we obtain the correction to the leading behaviour when t_L goes to zero.

To obtain the result at t_c , we have to expand in powers of $\varepsilon = d-2$. The subtracted integral diverges for s large for $d \rightarrow 2$:

$$\int_0^\infty \frac{ds}{\sqrt{s}} \left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2} \right) = -\frac{2}{\varepsilon} + \gamma - \ln 4\pi,$$

where γ is Euler's constant. The expression in the right-hand side of equation (32.102) has, as expected, a limit for $d \rightarrow 2$. At the fixed point t_c we have $t_L = t_c$. We need t_c at order ε^2 . Setting

$$\beta(t) = (d-2)t + \beta_2(d)t^2 + \beta_3(d)t^3,$$

we find

$$\frac{1}{t_c} + \frac{\beta_2(d)}{d-2} = \frac{\beta_3(2)}{\beta_2(2)} + O(d-2),$$

which is a scheme-independent quantity. Using the expression (19.129), coming from the minimal subtraction (MS) scheme, we find $\beta_3/\beta_2 = 1/2\pi$, and thus the universal ratio

$$\frac{\xi_L(t_c)}{L} = \frac{N-2}{N-1} \frac{1}{\pi\varepsilon} \left[1 + \frac{\varepsilon}{N-2} - \frac{\varepsilon}{2} (\gamma - \ln 4\pi) + O(\varepsilon^2) \right]. \tag{32.103}$$

A32 Additional remarks

In the Appendix, we first add a few remarks concerning the calculation of Feynman diagrams in a finite volume. We then briefly discuss finite-size effects in the case of discrete symmetries.

A32.1 Perturbation theory in a finite volume

To calculate Feynman diagrams in a finite volume, it is convenient to use Schwinger's representation and write the momentum space propagator as

$$\tilde{\Delta}(p) = \frac{1}{p^2 + \mu^2} = \int_0^\infty ds e^{-s(p^2 + \mu^2)},$$

a method already used in the infinite volume limit. After this transformation, Gaussian integrals over momenta are replaced by infinite sum over integers, which can no longer be calculated exactly. However, dimensional continuation can be defined, and the infinite-size limit studied. In the chapter, we have considered only simple one-loop diagrams D_γ , which can be written as

$$D_\gamma \equiv L^{-d} \sum_{p=2\pi k/L} (p^2 + \mu^2)^{-\sigma} = \frac{L^{-d}}{\Gamma(\sigma)} \int_0^\infty ds s^{\sigma-1} \sum_{p=2\pi k/L} e^{-s(p^2 + \mu^2)}.$$

In terms of the function $\vartheta_0(s)$ defined by equation (32.44),

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} e^{-\pi s n^2} \equiv \theta_3(0, e^{-\pi s}), \quad (A32.1)$$

where θ_3 is Jacobi's elliptic function, the sums can be written as

$$D_\gamma = \frac{L^{-d}}{\Gamma(\sigma)} \int_0^\infty ds s^{\sigma-1} e^{-s\mu^2} \vartheta_0^d(4s\pi/L^2) = \frac{L^{2\sigma-d}}{(4\pi)^\sigma \Gamma(\sigma)} \int_0^\infty ds s^{\sigma-1} e^{-sL^2\mu^2/(4\pi)} \vartheta_0^d(s).$$

Poisson's formula is useful in this context. Let $f(x)$ be a function that has a Fourier transform

$$\tilde{f}(k) = \int dx f(x) e^{i2\pi kx}.$$

Then from

$$\sum_{k=-\infty}^{+\infty} e^{i2\pi kx} = \sum_{l=-\infty}^{+\infty} \delta(x-l),$$

follows Poisson's formula,

$$\sum_{k=-\infty}^{+\infty} \tilde{f}(k) = \sum_{l=-\infty}^{+\infty} f(l). \quad (A32.2)$$

Applying this relation to the function $e^{-\pi s x^2}$, one finds the identity

$$\vartheta_0(s) = s^{-1/2} \vartheta_0(1/s). \quad (A32.3)$$

In particular, this identity shows that the infinite-size limit is approached exponentially when the mass μ is finite:

$$\vartheta_0(s) - s^{-1/2} \sim 2s^{-1/2} e^{-\pi/s} \Rightarrow D_\gamma(L) - D_\gamma(L=\infty) \underset{L \rightarrow \infty}{\propto} \mu^{2d-\sigma} (\mu L)^{\sigma-(d+1)/2} e^{-\mu L}.$$

A32.2 Discrete symmetries and finite-size effects

We have already characterized finite-size effects in the case of second order phase transitions, in the critical domain. For completeness, we consider a few examples of finite-size effects in the absence of critical fluctuations: first-order phase transitions, general phase transitions below the critical temperature. We recall that in the case of discrete symmetries, the infinite-size correlation length is finite in the ordered phase.

A32.2.1 Finite volume

We consider first order phase transitions, where at the transition the order parameter jumps from one constant value to another. We calculate here only homogeneous quantities, like the magnetization. The basic function we need is the free energy $-\mathfrak{S}(\varphi, L)$ at fixed field average,

$$\exp[-\mathfrak{S}(\varphi, L)] = \int [d\phi(x)] \delta \left(\varphi - L^{-d} \int d^d x \phi(x) \right) \exp[-\mathcal{S}(\phi)], \quad (A32.4)$$

a quantity whose physical meaning has already been discussed in Section 7.11. In the case of first order transitions, fluctuations are not critical, because the correlation length remains finite at the transition. Therefore, no zero-mode problem arises, and the integration over the zero-momentum component only restores the symmetry.

The free energy, in the presence of a constant magnetic field h , is given by

$$e^{W(h)} = \int d\varphi \exp[-\mathfrak{S}(\varphi, L) + L^d \beta h \varphi]. \quad (A32.5)$$

In a translation-invariant finite system, \mathfrak{S} at large size L behaves like

$$\mathfrak{S}(\varphi, L) \sim L^d \Sigma(\varphi), \quad (A32.6)$$

where $\Sigma(\varphi)$ is independent of L .

For L large, the integral can thus be calculated by the steepest descent method. In the case of a unique saddle point, one finds

$$W(h)/L^d = -\Sigma(\varphi) + \beta h \varphi, \quad (A32.7)$$

with

$$\Sigma'(\varphi) = \beta h.$$

Note that $\mathfrak{S}(\varphi, L)$ is such that the corrections for L large to (A32.7) are exponentially small in L , because the infinite-size correlation length remains finite.

Degenerate minima. By contrast, when several saddle points are found, $W(h)$ is the sum of saddle point contributions. As an example, we discuss an Ising-like system ($\mathfrak{S}(\varphi) = \mathfrak{S}(-\varphi)$), in the ordered phase, in an infinitesimal magnetic field h . The minimum is then almost degenerate. We denote by $\pm M_0$ the two minima of $\Sigma(\varphi)$ (the generalization to any discrete set of minima is simple):

$$\Sigma(\varphi) = \Sigma(M_0) + \frac{1}{2} \Sigma''(M_0) (\varphi - M_0)^2 + O(\varphi - M_0)^3. \quad (A32.8)$$

The free energy is the sum of the two saddle point contributions:

$$W(h) - W(0) = \ln \cosh(\beta h L^d M_0) + \frac{1}{2} \beta^2 h^2 L^d / \Sigma''(M_0). \quad (A32.9)$$

The magnetization M and the zero-field susceptibility χ are then given by

$$M = \langle \varphi \rangle = (\beta L^d)^{-1} W'(h) = M_0 \tanh(\beta h L^d M_0) + \beta h / \Sigma''(M_0), \quad (A32.10)$$

$$\chi = \beta / \Sigma''(M_0), \quad (A32.11)$$

where χ is here defined as $\chi = \langle \varphi^2 \rangle - (\langle |\varphi| \rangle)^2$.

In equation (A32.10), for $h L^d$ finite, the second term is negligible, and the finite-volume magnetization takes the universal form,

$$M = M_0 \tanh(\beta h L^d M_0). \quad (A32.12)$$

In zero field, the situation is more subtle, as we also discuss in the next section. For $|\varphi| < M_0$, almost uniform configurations compete with configurations in which in one fraction ρ ($0 < \rho < 1$) of the total volume $\varphi = -M_0$, and in the remaining fraction $1 - \rho$ $\varphi = M_0$. The average field expectation value is $\varphi = M_0(1 - 2\rho)$. The cost in energy is then proportional to a surface tension $\sigma(T)$ multiplied by the minimal area separating the two phases, which is of the order $L^{d-1} \rho^{(d-1)/d}$. When ρ increases, eventually non-uniform configurations are favoured when

$$\frac{1}{2} L^d \Sigma''(M_0) (\varphi - M_0)^2 = 2 \Sigma''(M_0) M_0^2 \rho^2 L^d > \sigma(T) (\rho^{1/d} L)^{d-1}.$$

Therefore, $\rho > \text{const. } (\xi(T)/L)^{d/(d+1)}$, where the correlation length $\xi(T)$ is implied by dimensional considerations. For larger values of ρ , at leading order $\mathfrak{S}(\varphi, L)$ is of the form

$$\mathfrak{S}(\varphi, L) = L^d \Sigma(M_0) + \text{const. } \sigma(T) L^{d-1} (1 - (\varphi/M_0)^2)^{(d-1)/d}.$$

Finally, for ρ large enough, in a finite volume with periodic boundary conditions, the dominant configuration consists in two phases separated by two flat interfaces and then the cost in energy becomes constant and equal for an hypercube to $2\sigma(T)L^{d-1}$.

A32.2.2 Finite size correlation length in Ising-like systems below T_c

In Chapter 39, we relate the restoration of discrete symmetries in one dimension to the existence of instantons. Assuming the results of Chapter 39, we indicate how these arguments extend to scalar field theories in higher dimensions, in order to demonstrate the complete parallelism between the low-temperature analysis of Section 14.2 for the Ising model, and the instanton analysis. The form of the finite-size correlation length below T_c follows.

We consider the example of an effective Euclidean action in d dimensions of the form,

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + V(\phi(x)) \right], \quad (A32.13)$$

where $V(\phi)$ is a potential, invariant in the reflection $\phi(x) \mapsto -\phi(x)$ and which has degenerate minima. For instance, it could have the form of the double-well potential

$$V(\phi) \sim (\phi^2 - M_0^2)^2. \quad (A32.14)$$

We want to evaluate the correlation ξ_L in the cylindrical geometry, with periodic boundary conditions in the transverse direction. We then follow the strategy of Section 32.4, and expand the field in Fourier components.

At leading order, we keep only the zero mode and thus

$$\mathcal{S}(\varphi) = L^{d-1} \int dz \left[\frac{1}{2} (\dot{\varphi}(z))^2 + V(\varphi(z)) \right].$$

The correlation length in the time direction is related to the difference in energy between the two lowest eigenvalues of the corresponding Hamiltonian. The limit $L \rightarrow \infty$ corresponds to the classical limit $\hbar \rightarrow 0$ of QM. In the classical limit, the energy difference is related to instanton configurations, which interpolate between the minima of the potential $V(\varphi)$. We can identify the classical action of the instanton with a surface tension, which is traditionally denoted by $\sigma(T)$. Instantons play, in continuous systems, the role of the walls of lattice models. We thus obtain for the correlation length

$$\ln \xi_L(T) \sim \sigma(T) L^{d-1}. \quad (A32.15)$$

This behaviour of the finite-size correlation length, characteristic of the breaking of discrete symmetries, has to be contrasted with the power-law behaviour (32.94) found in the case of continuous symmetries. Equation (A32.15) is valid for temperatures $T < T_c$, and the surface tension vanishes at T_c . The behaviour near T_c is given by equation (32.64) for $d > 4$, and by RG arguments for $2 < d < 4$:

$$\begin{aligned} \sigma(T) &\propto (T_c - T)^{3/2} \propto [\xi_\infty(T)]^3, \text{ for } d > 4, \\ &\propto (T_c - T)^{\nu(d-1)} \propto [\xi_\infty(T)]^{d-1}, \text{ for } d < 4. \end{aligned}$$

Remarks. For $T < T_c$, we have found that the finite-size correlation length is much larger than the size of the system. This justifies a mode expansion. Integrating over non-zero modes, one obtains the effective action for the remaining almost zero mode.

From the point of view of path integrals, the splitting between the two lowest eigenvalues of the Hamiltonian is related to instantons. In the case of one-dimensional instantons, to integrate out the fluctuations around the saddle point, one has to introduce collective coordinates corresponding to the position of the saddle point. In d dimensions, the position of the wall is defined by a function $\theta(x)$, which has also to be considered as a set of collective coordinates. Translation invariance implies that the action can only depend on the derivatives of $\theta(x)$. It has thus the form

$$\mathcal{S}_{\text{eff.}}(\theta) \sim \int d^{d-1}x \left[(\nabla \theta(x))^2 + O(|\nabla \theta(x)|^4) \right]. \quad (A32.16)$$

The term with only two derivatives gives the leading contribution, as long as the surface tension is large. When one approaches the critical point, terms with more derivatives become important. General Euclidean invariance implies that the effective action begins with a term proportional to the area of the wall, and has thus the form

$$\mathcal{S}_{\text{eff.}}(\theta) \sim \int d^{d-1}x \left[1 + (\nabla \theta(x))^2 \right]^{1/2}. \quad (A32.17)$$

It has been conjectured [338] that this model, the almost planar interface model, could describe the critical properties of the Ising model in $d = 1 + \varepsilon$ dimensions. It is easy to verify that the model has an UV fixed point of order ε , and that the correlation exponent ν can be calculated as a series in ε . These properties are reminiscent of the non-linear σ -model.

33 Quantum field theory (QFT) at finite temperature: Equilibrium properties

In this chapter, we review some *equilibrium properties* in statistical QFT, that is, relativistic QFT at finite temperature, a relativistic extension of the statistical quantum theories discussed in Sections 4.4.2 and 4.7 [339]. Study of QFT at finite temperature was initially motivated by cosmological problems [340], and then has gained additional attention in connection with high-energy heavy ion collisions, and speculations about possible phase transitions [341], also searched for in numerical simulations. Since we are interested here only in equilibrium physics, the Euclidean (or imaginary) time formalism is used throughout the chapter. Non-equilibrium phenomena can be described either by the same formalism after analytic continuation in the time variable or, alternatively by Schwinger's closed time path formalism in the convenient field integral formulation [342].

In particular, we discuss the limit of high temperature, or the situation of finite temperature phase transitions. Here, high temperature refers to an ultra-relativistic limit where the temperature, in energy unit, is much larger than the physical masses of particles. The concept of *dimensional reduction* then emerges [343]: in many cases, statistical properties of finite temperature QFT in $(1, d - 1)$ dimensions can be described by a classical statistical effective field theory (EFT) in $(d - 1)$ dimensions. Dimensional reduction generalizes a property already observed in the non-relativistic example of the Bose gas in Section 15.10, and indicates that quantum effects are less important at high temperature. (This property remains true, beyond dimensional reduction, in non-equilibrium processes.) The corresponding technical tools are a mode-expansion of fields in the Euclidean time variable, singling out the zero modes of boson fields, followed by a local expansion of the resulting $(d - 1)$ -dimensional effective field theory.

We especially emphasize that additional physical intuition about QFT at finite temperature in $(1, d - 1)$ dimensions can be gained by realizing that it can also be considered as a classical statistical field theory in d dimensions, with a finite size in one dimension (and, conversely, d -dimensional classical theory may shed some light on zero-temperature quantum transitions [344]). In particular, this identification makes an analysis of finite temperature QFT in terms of the renormalization group (RG), and the theory of finite-size effects of the classical theory, possible [345].

We illustrate these ideas with several simple examples, the ϕ^4 QFT, the non-linear σ model, the Gross–Neveu (GN) model, and some gauge theories. We construct the corresponding effective reduced theories at one-loop order. In models where the field is a N -component vector, the large N expansion provides an especially convenient tool to study the complete crossover between low and high temperature and, therefore, dimensional reduction.

33.1 Finite- (and high-) temperature field theory

In this section, we describe some general properties of QFT at thermal equilibrium in $(1, d - 1)$ dimensions, discuss the role of the mode expansion of fields in the Euclidean time variable, study the conditions under which statistical properties of finite temperature QFT in $(1, d - 1)$ dimension can be derived from an effective local classical statistical field theory in $(d - 1)$ dimensions, and indicate how to construct it explicitly.

33.1.1 Finite temperature QFT

The equilibrium properties of QFT at finite temperature can be derived from the partition function $\mathcal{Z} = \text{tr } e^{-H/T}$, where H is the quantum Hamiltonian, and T the temperature. For a simple theory with scalar boson fields ϕ and Euclidean action $\mathcal{S}(\phi)$, the partition function is given by the field integral

$$\mathcal{Z} = \int [d\phi] \exp [-\mathcal{S}(\phi)], \quad (33.1)$$

where $\mathcal{S}(\phi)$ is the integral of the Euclidean Lagrangian density $\mathcal{L}(\phi)$,

$$\mathcal{S}(\phi) = \int_0^{1/T} dt \int d^{d-1}x \mathcal{L}(\phi; t, x),$$

and the field ϕ satisfies periodic boundary conditions in the (Euclidean or imaginary) time direction (Section 2.4.1),

$$\phi(t=0, x) = \phi(t=1/T, x).$$

The QFT may also involve fermions. By contrast, fermion fields $\psi(t, x)$ satisfy anti-periodic boundary conditions (Section 4.6):

$$\psi(t=0, x) = -\psi(t=1/T, x).$$

Mode expansion. As a consequence of periodicity, fields have a Fourier series expansion in the Euclidean time direction with quantized frequencies ω_n (also called Matsubara frequencies). For boson fields,

$$\phi(t, x) = \sum_{n \in \mathbb{Z}} e^{i\omega_n t} \phi_n(x), \quad \text{with } \omega_n = 2n\pi T. \quad (33.2)$$

In the case of fermions, anti-periodic boundary conditions lead to the expansion

$$\psi(t, x) = \sum_{n \in \mathbb{Z}} e^{i\omega_n t} \psi_n(x), \quad \text{with } \omega_n = (2n+1)\pi T. \quad (33.3)$$

Remark. The mode expansion (33.2) is well-suited to simple situations where the field belongs to a linear space. In the case of non-linear σ -models, or non-Abelian gauge theories, the separation of zero modes is a more complicated problem.

Classical statistical field theory and finite temperature correlation length. The quantum partition function (33.1) has also the interpretation of the partition function of a classical statistical field theory in d dimensions. In this interpretation, finite temperature for the quantum partition function (33.1) corresponds, for the classical partition function, to a finite size $L = 1/T$ in one direction. The zero temperature limit of the QFT corresponds to the usual infinite volume limit of the classical theory.

An important parameter is the correlation length $\xi_L = 1/m_T$, which characterizes the decay of correlations in space directions. A crossover is expected between a d -dimensional behaviour, when ξ_L is small compared to L , that is, the thermal mass m_T is large compared to the temperature T , to a $(d-1)$ -dimensional behaviour when m_T is small compared to T .

For momenta much smaller than the temperature T , or distances much larger than L , this regime can be described by an effective $(d - 1)$ -dimensional local field theory. In the reduced theory, the temperature plays the role of a large momentum cut-off. The ratio m_T/T can be expected to be small in two situations, at high temperature and near a finite temperature phase transition. Actually, it is also small at low temperature in a third peculiar situation, when a symmetry is broken at zero temperature and no phase transition is possible at finite temperature.

Note that, in QFT the initial microscopic scale Λ^{-1} , where Λ is the QFT cut-off, is always present in the bare theory. Even at high temperature, the ratio Λ/T is assumed to remain large.

RG. General results obtained in the study of finite-size effects also apply here. For example, RG equations are only sensitive to short-distance singularities and, therefore, finite sizes do not modify RG equations, as we discuss in Chapter 32. Correlation functions satisfy the RG equations of the corresponding d -dimensional theory. Finite size only affects the solutions of RG equations, because a new dimensionless, RG-invariant variable appears, which can be chosen to be the ratio m_T/T .

In the tree approximation, the thermal mass m_T coincide with the physical mass m . At high temperature, the ratio m/T goes to 0. If m_T remains of order m , beyond leading order, the same applies to the relevant ratio m_T/T .

RG equations can also be solved by introducing the effective coupling at the temperature scale T . If the effective coupling goes to 0 at high temperature, then really the ratio m_T/T becomes small. Two examples will be met: the first one corresponds to theories where the free-field theory is an infrared (IR) fixed point, like the $\phi_{d=4}^4$ scalar field theory or quantum electrodynamics (QED), the second to ultraviolet (UV) asymptotically free-field theories like quantum chromodynamics (QCD). Conversely, when a non-trivial IR fixed point is present the ratio m_T/T goes to a constant. At high temperature, one then has to rearrange the initial perturbation theory by adding and subtracting a mass term to suppress fictitious perturbative large IR contributions.

33.1.2 The role of the zero mode, dimensional reduction

First, we consider a free-scalar field theory with the action,

$$\mathcal{S}(\phi) = \frac{1}{2} \int_0^{1/T} dt \int d^{d-1}x \left[(\partial_t \phi(t, x))^2 + (\nabla_x \phi(t, x))^2 + m^2 \phi^2(t, x) \right].$$

If we introduce the mode expansion (33.2) in the action, and integrate over time, we obtain a Euclidean $(d - 1)$ -dimensional field theory with an infinite number of fields, the modes $\phi_n(x)$:

$$\mathcal{S}(\phi) = \frac{1}{2T} \int d^{d-1}x \sum_{n \in \mathbb{Z}} [|\nabla_x \phi_n(x)|^2 + (m^2 + 4\pi^2 n^2 T^2) |\phi_n(x)|^2].$$

At high temperature, the thermal masses (the masses which govern the decay of correlations in space directions) of all modes, except the zero mode, become very large. The thermal mass $m_T = m$ of the zero mode remains finite and dominates the $\phi(t, x)$ -field correlation functions at large distance. Large-distance, low-momentum physics can entirely be described by an effective $(d - 1)$ -dimensional field theory only involving the zero mode.

Note that, even in a free theory, due to the anti-periodic boundary conditions, the thermal masses of the fermion modes (33.3) all become very large, and fermion modes decouple at high temperature.

The interacting theory. In an interacting theory, we expect all scalar non-zero modes and fermion modes to have thermal masses at least of order T . If the thermal mass m_T of the zero mode remains much smaller than the temperature, then the low-momentum physics can be described in terms of an effective $(d - 1)$ -dimensional Euclidean local field theory involving only the zero mode. This theory can be constructed by integrating out perturbatively all non-zero modes, and performing a local expansion of the resulting effective action. In the reduced $(d - 1)$ -dimensional theory, the temperature T acts as a large momentum cut-off.

Note that, in the following, when we mention masses in the reduced theory without qualification, we refer to these thermal masses, and not to the physical masses of particles, which are defined at zero temperature.

33.1.3 The EFT

We outline the construction of the $(d - 1)$ -dimensional EFT in the example of a general scalar field theory, assuming that the mass m_T of the zero mode is indeed much smaller than T . We first set

$$\phi(t, x) = \varphi(x) + \chi(t, x), \quad (33.4)$$

where φ is the zero mode, and χ the sum of all other modes (equation (33.2)):

$$\chi(t, x) = \sum_{n \neq 0} e^{i\omega_n t} \phi_n(x), \quad \omega_n = 2n\pi T. \quad (33.5)$$

The action $\mathcal{S}_T(\varphi)$ of the reduced theory is then defined by

$$e^{-\mathcal{S}_T(\varphi)} = \int [d\chi] \exp[-\mathcal{S}(\varphi + \chi)]. \quad (33.6)$$

At leading order in perturbation theory, one can set $\chi = 0$, and finds the local action

$$\mathcal{S}_T(\varphi) = \frac{1}{T} \int d^{d-1}x \mathcal{L}(\varphi; x). \quad (33.7)$$

In this leading approximation, T plays the formal role of \hbar , and the small T expansion reduces to a loop expansion. If T/Λ , which is small, is the relevant expansion parameter, which means that the perturbative expansion is dominated by large momentum (UV) contributions, then the effective $(d - 1)$ -dimensional theory can still be studied by perturbative methods. This is expected when the number $(d - 1)$ of space dimensions is large, and field theories are not renormalizable. However, m/T is another dimensionless combination, which at high temperature is small. This may be the relevant expansion parameter for theories that are dominated by small momentum IR contributions, a problem which arises in low dimensions. Then, perturbation theory is no longer possible or useful. Actually, the relevant parameter in the full effective theory is m_T/T . Therefore, the contributions to the mass of the zero mode due to quantum and thermal fluctuations have to be investigated.

Loop corrections to the effective action. After an integration over the non-zero modes, the effective action contains non-local interactions. To study long-wavelength phenomena, one can perform a *local expansion* of the effective action, an expansion that breaks down at momenta of order T . One expects, but this has to be checked carefully, that, in general, higher order corrections coming from the integration over non-zero modes generate terms which renormalize the terms already present at leading order, and additional interactions suppressed by powers of $1/T$. Exceptions are provided by gauge theories, where new low-dimensional interactions are generated by the breaking of the $O(1, d - 1)$ invariance.

Renormalization. If the initial $(1, d - 1)$ -dimensional theory has been renormalized at $T = 0$, the complete theory is finite in the formal infinite cut-off limit. As a consequence of the zero-mode subtraction, cut-off dependent terms may appear in the reduced $(d - 1)$ -dimensional action. These terms provide the necessary counter-terms that render the perturbative expansion of the EFT finite. The effective action can thus be written as

$$\mathcal{S}_T(\varphi) = \mathcal{S}_T^{\text{finite}}(\varphi) + \text{counter-terms}.$$

Correlation functions have finite expressions in terms of the parameters of the effective action, in which the counter-terms have been omitted. The first part $\mathcal{S}_T^{\text{finite}}(\varphi)$ thus satisfies the RG equations of the d -dimensional theory.

However, new apparent divergences can be generated by the local expansion. Determining the finite parts may involve some non-trivial calculations.

33.2 The example of the $\phi_{1,d-1}^4$ field theory

We first study the example of the $O(N)$ -symmetric $(\phi^2)^2$ scalar field theory, where the field ϕ is an N -component vector, and the Hamiltonian has the form,

$$\mathcal{H}(\boldsymbol{\Pi}, \boldsymbol{\phi}) = \frac{1}{2} \int d^{d-1}x \boldsymbol{\Pi}^2(x) + \Sigma(\boldsymbol{\phi}), \quad (33.8)$$

where $\boldsymbol{\Pi}$ is the conjugate momentum, and

$$\Sigma(\boldsymbol{\phi}) = \int d^{d-1}x \left\{ \frac{1}{2} [\nabla \boldsymbol{\phi}(x)]^2 + \frac{1}{2} (r_c + \tau) \boldsymbol{\phi}^2(x) + \frac{1}{4!} u (\boldsymbol{\phi}^2(x))^2 \right\}. \quad (33.9)$$

As usual, a *large momentum cut-off* Λ , which renders the field theory UV finite, is implied everywhere. The quantity r_c , which has the form of a mass renormalization is determined by the condition that, at $T = 0$ (zero temperature), when τ vanishes the physical mass m of the field ϕ vanishes. At $\tau = 0$, a transition occurs between a symmetric phase, $\tau > 0$, and a broken phase, $\tau < 0$. We recall that the field theory is meaningful only if the physical mass m is much smaller than the cut-off Λ . This implies either (the famous *fine-tuning* problem) $|\tau| \ll \Lambda^2$ or, for $N \neq 1$, $\tau < 0$ which corresponds to a spontaneously broken symmetry with massless Goldstone modes. The latter situation will be examined in Section 33.4, within the more suitable formalism of the non-linear σ -model.

We define a dimensionless coupling λ , setting

$$u = \Lambda^{4-d} \lambda, \quad (33.10)$$

which later will be assumed to take generic (*i.e.*, not very small) values.

The quantum partition function, at finite temperature T , is given by the field integral (33.1) with the action

$$\mathcal{S}(\boldsymbol{\phi}) = \int_0^{1/T} dt \left[\int d^{d-1}x \frac{1}{2} (\partial_t \boldsymbol{\phi}(t, x))^2 + \Sigma(\boldsymbol{\phi}) \right], \quad (33.11)$$

where, in Σ , the field $\boldsymbol{\phi}(x)$ is replaced by $\boldsymbol{\phi}(t, x)$. The field satisfies periodic boundary conditions in the Euclidean-time direction:

$$\boldsymbol{\phi}(0, x) = \boldsymbol{\phi}(1/T, x).$$

33.2.1 RG at finite temperature

A useful information can be derived from an RG analysis. Correlation functions at finite temperature satisfy the RG equations of the zero-temperature QFT, or the d -dimensional classical field theory in infinite volume. The dimension $d = 4$ is special, because the $\phi_{d=4}^4$ QFT is just renormalizable. One important quantity is the ratio m_T/T , where the thermal mass m_T governs the decay of correlations in space directions and is also the mass of the zero mode in the effective reduced theory.

Dimensions $d > 4$. For $d > 4$, the theory is not renormalizable (the interactions are irrelevant), and the Gaussian fixed point $u = 0$ is stable. The coupling constant $u = \lambda\Lambda^{4-d}$ is small in the physical domain, and perturbation theory is applicable. At zero temperature, the physical mass in the symmetric phase has the scaling behaviour of a free or Gaussian theory, $m \propto \tau^{1/2}$. The leading corrections to the two-point function due to finite temperature effects are of order u . Therefore, in the symmetric phase, for dimensional reasons,

$$m_T \propto (\tau + \text{const. } \lambda\Lambda^{4-d}T^{d-2})^{1/2}.$$

This general form has several consequences.

If the $O(N)$ symmetry is broken ($\tau < 0$) at zero temperature, a phase transition occurs at a temperature T_c which scales like

$$T_c \propto \Lambda (-\tau/\Lambda^2)^{1/(d-2)} \gg (-\tau)^{1/2}.$$

This means that, for $N > 1$, the critical temperature is large compared to the crossover mass scale (see Section 19.12.3) of the massless phase at $T = 0$ and, for $N = 1$, compared to the physical mass $m \propto (-\tau)^{1/2}$.

At high temperature, or in the massless theory ($\tau = 0$), the effective mass m_T behaves like

$$m_T/T \propto (T/\Lambda)^{(d-4)/2} \ll 1.$$

The property $m_T/T \ll 1$ implies the validity of dimensional reduction.

Dimension $d = 4$. The theory is just renormalizable and logarithmic deviations from naive scaling appear. RG equations for vertex functions (in the Fourier representation) take the form

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(\lambda) \frac{\partial}{\partial \lambda} - \frac{n}{2} \eta(\lambda) - \eta_2(\lambda) \tau \frac{\partial}{\partial \tau} \right] \tilde{\Gamma}^{(n)}(p_i; \Lambda, \lambda, \tau, T) = 0. \quad (33.12)$$

The ratio $m_T/T = F(\Lambda, \lambda, \tau, T)$ is dimensionless and RG-invariant and, thus, satisfies

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(\lambda) \frac{\partial}{\partial \lambda} - \eta_2(\lambda) \tau \frac{\partial}{\partial \tau} \right] F(\Lambda, \lambda, \tau, T) = 0.$$

The solution can be written as

$$m_T/T = F(\Lambda/T, \lambda, \tau/T^2, 1) = F(\ell\Lambda/T, \lambda(\ell), \tau(\ell)/T^2, 1), \quad (33.13)$$

where ℓ is a scale parameter and $\lambda(\ell), \tau(\ell)$ the corresponding running parameters (or effective parameters at scale ℓ):

$$\ell \frac{d\lambda(\ell)}{d\ell} = \beta(\lambda(\ell)), \quad \ell \frac{d\tau(\ell)}{d\ell} = -\tau(\ell)\eta_2(\lambda(\ell)).$$

The form of the RG β -function,

$$\beta(\lambda) = \frac{(N+8)}{48\pi^2} \lambda^2 + O(\lambda^3), \quad (33.14)$$

implies that the theory is IR-free, that is, that $\lambda(\ell) \rightarrow 0$ for $\ell \rightarrow 0$. The effective coupling constant at the physical scale is logarithmically small. For example, to describe physics at the scale T , we have to choose $\ell = T/\Lambda \ll 1$ and, thus,

$$\lambda(T/\Lambda) \sim \frac{48\pi^2}{(N+8)\ln(\Lambda/T)}. \quad (33.15)$$

From

$$\eta_2(\lambda) = -\frac{N+2}{48\pi^2} \lambda + O(\lambda^2),$$

one also infers

$$\tau(T/\Lambda) \propto \frac{\tau}{(\ln \Lambda/T)^{(N+2)/(N+8)}}. \quad (33.16)$$

Therefore, RG improved perturbation theory can be used to derive the effective action of the reduced theory. The behaviour of m_T/T and T_c is discussed in Section 33.3.2.

Dimension $d = 3$. The three-dimensional classical theory has an IR fixed point $\lambda^* \neq 0$. Then, in the symmetric phase, finite-size scaling (equation (33.13)) predicts the scaling form,

$$m_T/T = f(\tau/T^{1/\nu}),$$

where ν is the correlation exponent of the three-dimensional system. Therefore, in general, m_T remains of order T at high temperature.

The zero mode is special only if the function $f(x)$ is small (compared to 1). This happens near a phase transition but, in an effective two-dimensional theory, a phase transition is possible only for $N = 1$. Moreover, a finite temperature phase transition can occur only if the symmetry is broken at zero temperature ($\tau < 0$) (temperature increases disorder). If $f(x)$ vanishes at a value $x = x_0 < 0$, the critical temperature T_c has a scaling behaviour of the form

$$T_c = (-x_0)^{-\nu}(-\tau)^\nu \propto m,$$

where m is the physical mass in the low-temperature phase. Near T_c , the IR properties are described by an effective two-dimensional theory.

The situation in which, for $N \neq 1$, the symmetry is broken at zero temperature is examined in the framework of the non-linear σ -model starting with Section 33.4.

33.2.2 One-loop effective action

We now construct the effective $(d-1)$ -dimensional theory and discuss its validity. However, this construction is useful mainly if the IR contributions are large enough to invalidate perturbation theory, that is, in dimensions $d \leq 4$. By contrast, for dimensions $d > 4$ the reduced $(d-1)$ -dimensional theory has a finite perturbation expansion, even in the massless limit. This property we can be verified with the example of the dimension $d = 5$.

Mode expansion and effective action at leading order. To construct the effective field theory in $(d - 1)$ dimensions, one expands the field in eigenmodes in the Euclidean-time direction (equation (33.2)). One then determines the effective action (33.6) by integrating perturbatively over all non-zero modes (in the decomposition (33.4)). In the notation (33.9), at leading order, the result is

$$\mathcal{S}_T(\varphi) = \frac{1}{T} \Sigma(\varphi) = \frac{1}{T} \int d^{d-1}x \left\{ \frac{1}{2} [\nabla \varphi(x)]^2 + \frac{1}{2} \tau \varphi^2(x) + \frac{1}{4!} u(\varphi^2(x))^2 \right\}, \quad (33.17)$$

because r_c vanishes at this order.

Note that, if we rescale φ into $\varphi T^{1/2}$, the coupling constant is changed into uT , or in terms of the dimensionless coupling constant λ into $\lambda T/\Lambda^{d-4}$. A meaningful expansion parameter is dimensionless and, therefore, $\lambda T/\Lambda^{d-4}$ has to be multiplied by a mass function of T and τ to the power $(d - 5)$.

For $d \geq 5$, the expansion parameter is always small because Λ/T is large.

For $d = 4$, the situation is more subtle. If the symmetry is unbroken at zero temperature, at high temperature $T \gg m = \sqrt{\tau}$, one expects the corrections to the mass m_T of the zero mode to be dominated by the one-loop contribution to the two-point function, which is of order $T^2 \lambda(T/\Lambda)$. This yields a mass $T\sqrt{\lambda}$, which is small compared to T , but may be large compared to m . A renormalization of the mass in order to introduce the mass parameter m_T , which involves a summation, leads to an expansion in powers of $T/m_T = O(\sqrt{\lambda})$, a parameter that is small. In such a situation, mode and local expansions are justified, but the effective three-dimensional theory is still perturbative.

By contrast, if τ is sufficiently negative, m_T may be much smaller near a phase transition, and perturbation theory in the reduced theory then fails.

For $d < 4$, IR singularities are always present both in the initial and the reduced theory, and the small coupling regime can never be reached for interesting situations. The $\varepsilon = 4 - d$ expansion can be useful in some limits, otherwise the problem has to be studied by non-perturbative methods.

One-loop calculation. The one-loop contribution to the effective action generated by integrating over the non-zero modes is given by (using $\ln \det = \text{tr} \ln$),

$$\mathcal{S}_T^{(1)}(\varphi) = \frac{1}{2} \text{tr} \ln [(-\partial_t^2 - \nabla_{d-1}^2 + \tau + \frac{1}{6} u \varphi^2(x)) \delta_{ij} + \frac{1}{3} u \varphi_i(x) \varphi_j(x)] - (\varphi = 0).$$

The situation of interest here is when the physical mass m (the inverse of the correlation length $\xi = 1/m$ of the infinite volume d -dimensional system) is smaller than the temperature T . In this situation, we expect to be able to make a local expansion in φ . The leading order in the derivative expansion is obtained by treating $\varphi(x)$ as a constant.

As in the general finite-size calculations, to evaluate the one-loop contribution $\mathcal{S}_T^{(1)}$ to the reduced action, one can introduce Schwinger's representation and Jacobi's related ϑ_0 function (see Section 32.3.3), or infer it directly from the partition function of the harmonic oscillator:

$$\begin{aligned} \text{tr} \ln(-\partial_t^2 - \nabla_{d-1}^2 + M^2) &= V_{d-1} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \sum_{n \neq 0} \ln(\omega_n^2 + k^2 + M^2) \\ &= 2V_{d-1} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln [2 \sinh(\omega(k)/2T)/\omega(k)], \end{aligned} \quad (33.18)$$

where V_{d-1} is the $(d - 1)$ -dimensional volume and $\omega(k) = \sqrt{k^2 + M^2}$.

Applied to the present example one finds,

$$\begin{aligned} \mathcal{S}_T^{(1)}(\varphi) = & \int d^{d-1}x \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \left\{ (N-1) \ln [2 \sinh(\omega_T(k)/2T)/\omega_T(k)] \right. \\ & \left. + \ln [2 \sinh(\omega_L(k)/2T)/\omega_L(k)] \right\} - (\varphi = 0), \end{aligned} \quad (33.19)$$

with

$$\omega_T(k) = \sqrt{k^2 + \tau + \frac{1}{6}u\varphi^2(x)}, \quad \omega_L(k) = \sqrt{k^2 + \tau + \frac{1}{2}u\varphi^2(x)}. \quad (33.20)$$

We now expand in powers of φ . At this order, the expansion only makes sense if $-\tau/T^2 < 4\pi^2$, a condition that, more generally, involves the dimensionless RG-invariant ratio m^2/T^2 , where m is the mass parameter of the ordered phase. The condition implies that the expansion around $\varphi = 0$ is only useful if the field expectation value is small enough.

Order φ^2 . For the quadratic term, the local approximation is not needed, because the corresponding one-loop diagram is a constant:

$$[\mathcal{S}_T^{(1)}]_2 = \frac{1}{12}(N+2)\bar{G}_2(\tau, T) \frac{u}{T} \int d^{d-1}x \varphi^2(x), \quad (33.21)$$

where the constant \bar{G}_2 is given by

$$\bar{G}_2(\tau, T) = \int \frac{d^{d-1}k}{(2\pi)^{d-1}\omega(k)} \left(\frac{1}{2} + \frac{1}{e^{\omega(k)/T}-1} - \frac{T}{\omega(k)} \right), \quad (33.22)$$

with now $\omega(k) = \sqrt{k^2 + \tau}$. One recognizes the sum of the zero-temperature result, the thermal fluctuations, and the subtracted zero-mode contribution.

Introducing the zero-temperature diagram $\Omega_d(m)$ (see equations (18.16) and (18.28)),

$$\Omega_d(m) = \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^dk}{k^2 + m^2}, \quad (33.23)$$

and the UV finite function

$$\begin{aligned} f_d(s) &= N_{d-1} \int_0^\infty \frac{x^{d-2}dx}{\sqrt{x^2+s}} \frac{1}{\exp[\sqrt{x^2+s}] - 1} \\ &= N_{d-1} \int_{\sqrt{s}}^\infty (y^2-s)^{(d-3)/2} \frac{dy}{e^y - 1}, \end{aligned} \quad (33.24)$$

where N_d is the loop factor, (equation (10.12))

$$N_d = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}, \quad (33.25)$$

we can rewrite \bar{G}_2 as

$$\bar{G}_2(\tau, T) = \Omega_d(\sqrt{\tau}) - T\Omega_{d-1}(\sqrt{\tau}) + T^{d-2}f_d(\tau/T^2). \quad (33.26)$$

In particular,

$$f_d(0) = N_{d-1}\Gamma(d-2)\zeta(d-2), \quad f'_d(0) = -\frac{1}{2}(d-3)N_{d-1}\Gamma(d-4)\zeta(d-4), \quad (33.27)$$

where $\zeta(s)$ is Riemann's function (defined by equation (A33.5)).

Order $(\varphi^2)^2$. The quartic term,

$$\left[\mathcal{S}_T^{(1)} \right]_4 = -\frac{1}{144}(N+8)\bar{G}_4(\tau, T) \frac{u^2}{T} \int d^{d-1}x (\varphi^2(x))^2, \quad (33.28)$$

is proportional to the initial interaction. One verifies that

$$\bar{G}_4(\tau, T) = -\frac{\partial}{\partial \tau} \bar{G}_2(\tau, T). \quad (33.29)$$

The one-loop reduced action. We first only keep the contributions proportional to terms already present in the tree approximation. The value of r_c corresponds to the mass renormalization, which renders the $T=0$ theory massless at one-loop order. Thus, \bar{G}_2 has to be replaced by $[\bar{G}_2]_r$. One obtains (for $d > 2$),

$$\begin{aligned} [\bar{G}_2]_r(\tau, T) &= \bar{G}_2(\tau, T) - \Omega_d(0) \\ &= -\tau D_d(\sqrt{\tau}) - T\Omega_{d-1}(\sqrt{\tau}) + T^{d-2}f_d(\tau/T^2), \end{aligned}$$

with the definition,

$$D_d(m) = \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d k}{k^2(k^2 + m^2)} = \frac{1}{m^2} [\Omega_d(0) - \Omega_d(m)]. \quad (33.30)$$

After the rescaling $\varphi \mapsto \varphi T^{1/2}$, the effective action can be written as

$$\mathcal{S}_T(\varphi) = \int d^{d-1}x \left\{ \frac{1}{2} [\nabla \varphi(x)]^2 + \frac{1}{2}\sigma_2 \varphi^2(x) + \frac{1}{4!}\sigma_4 (\varphi^2(x))^2 \right\}, \quad (33.31)$$

with

$$\sigma_2 = r + \frac{1}{6}(N+2)u[\bar{G}_2]_r, \quad \sigma_4/T = u - \frac{1}{6}(N+8)u^2\bar{G}_4.$$

Other interactions. For space dimensions $d < 6$, the coefficients of the other interaction terms are no longer UV divergent. Since the zero-mode contribution has been subtracted, no IR divergence is generated even in the massless limit. In this limit, the coefficients are thus proportional to powers of $1/T$ obtained by dimensional analysis (in the normalization (33.31)):

$$[\mathcal{S}_T]_{(2n)} \propto \lambda^n (T/\Lambda)^{n(d-4)} T^{d-1-n(d-3)} \int d^{d-1}x (\varphi^2(x))^n, \quad (33.32)$$

and, therefore, increasingly negligible at high temperature, at least for $d \geq 4$.

The local expansion of the one-loop determinant also generates monomials with derivatives. No term proportional to $(\nabla_x \varphi)^2$ is generated at one-loop order. All other terms with derivatives are finite for $d < 6$, and thus the structure of the coefficients again is given by dimensional analysis. To $2k$ derivatives corresponds an additional factor $1/T^{2k}$.

Finally, for $\tau \neq 0$ but $\tau/T^2 \ll 1$, we can expand in powers of τ , and the preceding arguments immediately generalize.

33.3 High temperature and critical limits

We now examine two interesting situations. First, we discuss the limit $\tau \rightarrow 0$, which corresponds to a massless theory at zero temperature in the QFT context (and to the critical temperature of the d -dimensional statistical field theory). This gives the leading contributions in the high-temperature limit. It will prove useful to also keep terms linear in $\tau\varphi^2$.

The constants $[\bar{G}_2]_r$ and \bar{G}_4 , for $\tau = 0$, become

$$[\bar{G}_2]_r = -T\Omega_{d-1}(0) + T^{d-2}N_{d-1}\Gamma(d-2)\zeta(d-2), \quad (33.33)$$

$$\bar{G}_4 = D_d(0) - TD_{d-1}(0) + \frac{1}{2}N_{d-1}(d-3)\Gamma(d-4)\zeta(d-4)T^{d-4}, \quad (33.34)$$

where the values (33.27) have been used. The expression for $[\bar{G}_2]_r$ is the sum of two terms, a renormalized mass term for the zero mode, and the one-loop counter-term, which renders the two-point function one-loop finite in the reduced theory. The expression for \bar{G}_4 contains a finite temperature contribution, a zero-temperature renormalization for $d \geq 4$, and a counter-term for the reduced theory for $d \geq 5$.

Finally, from the relation (33.29) and the value of \bar{G}_4 , we obtain the term linear in τ in \bar{G}_2 :

$$\bar{G}_2(\tau, T) = \bar{G}_2(0, T) - \tau\bar{G}_4(0, T) + O(\tau^2). \quad (33.35)$$

Then, we can consider the case of a phase transition at finite temperature and calculate the critical temperature T_c . As we verify, a phase transition is possible only if the symmetry is broken at zero temperature (as expected, since increasing the temperature increases the disorder). Note that, from the point of view of the classical d -dimensional theory, it is more natural to determine the parameter τ , which is related to the classical temperature, as a function of the size $1/T$. At T_c , the correlation length diverges and the mass m_T of the zero mode vanishes, justifying dimensional reduction and local expansion.

33.3.1 Dimension $d = 5$

For $d > 4$, the coupling constant u , which is of order Λ^{4-d} , is very small. At high temperature, the ratio m_T/T is of order $(\Lambda/T)^{(4-d)/2}$ and thus small, justifying a mode expansion.

We now examine more precisely the $d = 5$ case, keeping the contribution of order τ in \bar{G}_2 . Then,

$$[\bar{G}_2]_r = \frac{T^3}{4\pi^2}\zeta(3) - T\Omega_4(0) - \tau \left[D_5(0) - \frac{T}{8\pi^2}(\ln(\Lambda/T) + \kappa_5) \right] + O(\tau^2),$$

$$\bar{G}_4 = D_5(0) - \frac{T}{8\pi^2}(\ln(\Lambda/T) + \kappa_5),$$

where κ_5 is a renormalization scheme dependent constant, and $\Omega_4(0) \propto \Lambda^2$, $D_5(0) \propto \Lambda$. The infinite volume terms proportional to $D_5(0)$ induce finite renormalizations $\lambda \mapsto \lambda_r$, of the dimensionless φ^4 coupling constant λ , and $\tau \mapsto \tau_r$, of the φ^2 coefficient,

$$u = \lambda/\Lambda, \quad \lambda_r = \lambda - \frac{D_5(0)}{\Lambda} \frac{N+8}{6} \lambda^2, \quad \tau_r/\tau = 1 - \frac{D_5(0)}{\Lambda} \frac{N+2}{6} \lambda.$$

The remaining cut-off dependent terms in $[\bar{G}_2]_r$ and \bar{G}_4 render the effective four-dimensional theory one-loop finite.

Using the expression (33.31), and introducing the small dimensionless (effective) coupling constant

$$\lambda_T = \lambda_r T / \Lambda,$$

we can write the effective action at one-loop order (after the rescaling $\varphi \mapsto \varphi T^{1/2}$) as

$$\mathcal{S}_T(\varphi) = \int d^4x \left\{ \frac{1}{2} [\nabla \varphi(x)]^2 + \frac{1}{2} \tau_T \varphi^2(x) + \frac{1}{4!} \lambda_T (\varphi^2(x))^2 \right\} + \delta \mathcal{S}_{T,\Lambda}(\varphi), \quad (33.36)$$

with

$$\tau_T = \tau_r + \frac{1}{6}(N+2) \frac{\zeta(3)}{4\pi^2} T^2 \lambda_T, \quad (33.37)$$

and where $\delta \mathcal{S}_{T,\Lambda}$ is the sum of one-loop counter-terms:

$$\delta \mathcal{S}_{T,\Lambda}(\varphi) = \int d^4x \left[\frac{1}{2} \delta \tau_T \varphi^2(x) + \frac{1}{4!} \delta \lambda_T (\varphi^2(x))^2 \right],$$

with

$$\begin{aligned} \delta \tau_T &= -\frac{N+2}{6} \left\{ \Omega_4(0) + \frac{\tau_r}{8\pi^2} [\ln(\Lambda/T) + \kappa_5] \right\} \lambda_T, \\ \delta \lambda_T &= \frac{N+8}{48\pi^2} [\ln(\Lambda/T) + \kappa_5] \lambda_T^2. \end{aligned}$$

Note that, for dimensions $d \geq 5$, we can study the effective theory by perturbation theory and RG. The temperature T plays the role of the cut-off in the reduced theory. For $d = 5$, the perturbative expansion of the reduced four-dimensional theory contains large logarithms of the form $\ln(\tau_r/T^2)$, which can be summed by RG techniques. However, the initial coupling constant λ_T , because it is very small, has no time to run.

Other local interactions. In the same normalization, an interaction term with $2n$ fields and $2k$ derivatives is proportional to $\lambda_T^n / T^{2n-4+2k}$, and thus negligible in the situations under study for $n > 2$ or $n = 2, k > 0$.

The massless theory. At leading order in the massless theory $\tau = 0$, the ratio m_T/T is

$$m_T/T = [(N+2)\zeta(3)\lambda_T/24\pi^2]^{1/2} \ll 1.$$

Although the induced mass remains very small, because the effective four-dimensional theory has at most logarithmic IR singularities, and because the effective coupling constant λ_T is of order T/Λ , the reduced theory can still be expanded in perturbation theory.

The critical temperature. Using the expression (33.37), one finds in the tree approximation

$$T_c^2 = -\frac{24\pi^2}{(N+2)\zeta(3)} \frac{\tau_r}{\lambda_T},$$

an expression that makes sense only if $T_c \ll \Lambda$, and thus if $|\tau_r| \ll T\Lambda$, is very small. This condition justifies a small τ expansion and implies that the critical temperature is large in the zero-temperature mass scale $\sqrt{-\tau_r}$ (the mass scale corresponding to the crossover between critical and Goldstone mode behaviours). The equation also confirms that a phase transition is only possible if, at $T = 0$ (zero QFT temperature), the system is in the ordered phase (as expected, since increasing the temperature increases the disorder).

33.3.2 Dimension $d \leq 4$

Four dimensions. For $d = 4$ and at order τ , we now obtain

$$[\bar{G}_2]_r = \frac{T^2}{12} - T\Omega_3(0) - \frac{1}{8\pi^2} [\ln(\Lambda/T) + \kappa_4] \tau, \quad (33.38)$$

$$\bar{G}_4 = \frac{1}{8\pi^2} [\ln(\Lambda/T) + \kappa_4], \quad (33.39)$$

where $\Omega_3(0)$ is proportional to Λ , and κ_4 is a renormalization scheme-dependent constant.

The coupling constant $u \equiv \lambda$ is dimensionless. Then, \bar{G}_4 just yields the one-loop contribution to the perturbative expansion of the running (or effective) coupling constant:

$$\lambda(T/\Lambda) = \lambda - \frac{N+8}{48\pi^2} [\ln(\Lambda/T) + \kappa_4] \lambda^2 + O(\lambda^3).$$

In fact, we know from RG arguments that all quantities can be expressed entirely in terms of the running coupling constant.

Similarly, \bar{G}_2 contains a one-loop contribution to the perturbative expansion of $\tau(T/\Lambda)$, the running coefficient of φ^2 ,

$$\tau(T/\Lambda)/\tau = 1 - \frac{N+2}{48\pi^2} [\ln(\Lambda/T) + \kappa_4] \lambda.$$

The three-dimensional effective theory is super-renormalizable and, thus, requires only a mass renormalization. In \bar{G}_2 , we find two terms, $-T\Omega_3(0)$, which is cut-off dependent and yields the one-loop counter-term, and $T^2/12$, which contributes to the mass of the zero mode. The one-loop effective action takes the form

$$\begin{aligned} \mathcal{S}_T(\varphi\sqrt{T}) &= \int d^3x \left\{ \frac{1}{2} [\nabla\varphi(x)]^2 + \frac{1}{2}\tau_T\varphi^2(x) + \frac{1}{4!}\lambda_T(\varphi^2(x))^2 \right\} \\ &\quad + \text{one-loop counter-terms}, \end{aligned} \quad (33.40)$$

with

$$\tau_T = \tau(T/\Lambda) + \frac{N+2}{72} T^2 \lambda(T/\Lambda), \quad \lambda_T = T \lambda(T/\Lambda).$$

The massless theory. For $\tau = 0$ (the massless zero-temperature theory), in the tree approximation of the reduced theory, one finds

$$(m_T/T)^2 = \frac{N+2}{72} \lambda(T/\Lambda).$$

From the solution of the RG equation, we know that $\lambda(T/\Lambda)$ goes to zero as $1/\ln(\Lambda/T)$ for $T/\Lambda \ll 1$ (equation (33.15)). Therefore,

$$(m_T/T)^2 \sim \frac{2\pi^2(N+2)}{3(N+8)} \frac{1}{\ln(\Lambda/T)}. \quad (33.41)$$

The mass of the zero mode is smaller, although only logarithmically smaller, than the other modes, justifying the mode expansion. Moreover, the perturbative expansion of the three-dimensional effective theory is, for small momenta, an expansion in $T\lambda(T/\Lambda)$ divided by the mass which is of order $T\sqrt{\lambda(\Lambda/T)}$. The expansion parameter thus is of order $\sqrt{\lambda(\Lambda/T)}$ which is small, due to the IR freedom of the four-dimensional theory.

Higher order calculations have been performed. However, the convergence is expected to be extremely slow and, therefore, summation techniques have been proposed. General summation methods, which have been used in the calculation of three-dimensional critical exponents, could also be useful here.

The critical temperature. If, at zero temperature, the system is in an ordered phase ($\tau < 0$), at higher temperatures a phase transition may occur at a critical temperature T_c , which, at leading order, is solution of the equation

$$\tau_T = \tau(T/\Lambda) + \frac{N+2}{72} T^2 \lambda(T/\Lambda) = 0.$$

This relation can be rewritten in various forms, for example, as

$$\sqrt{(N+2)/12} T_c \propto m \sqrt{\ln(\Lambda/m)} \propto (-\tau)^{1/2} (\ln(-\tau))^{3/(N+8)}, \quad (33.42)$$

where m is the low-temperature physical (crossover) mass scale. The critical temperature is large compared to the mass scale m , and thus belongs to the high-temperature regime. The critical theory, which can no longer be studied by perturbative methods, is the theory relevant to a large class of phase transitions in statistical physics, to which we devote Chapters 15–17.

Other local interactions. In the same normalization, an interaction term with $2n$ fields and $2k$ derivatives is proportional to λ^n/T^{n-3+2k} , and thus negligible in the situations under study for $n > 2$ or $n = 2$, $k > 0$, because even the zero-mode mass is large.

Dimension $d < 4$. RG arguments imply that the finite-temperature mass scale is proportional to T for $d < 4$. Since m_T/T is of order unity, the separation of the zero mode is no longer justified. To calculate correlation functions at momenta small compared to the temperature, and for small field expectation value, a local expansion of the type of chiral perturbation theory still makes physical sense, but it is necessary to modify the perturbative expansion. For example, in the d -dimensional field theory, one can add and subtract a mass term of order T for the zero mode. This temperature dependent mass renormalization modifies the propagator and introduces an IR cut-off. One then determines the mass term by demanding cancellation of the higher-order corrections to the mass.

An alternative strategy is to work in $d = 4 - \varepsilon$ dimension and use the ε -expansion. Then, the zero-mode effective mass is formally small of order $T\sqrt{\varepsilon}$, and the expansion parameter is $\sqrt{\varepsilon}$.

The critical temperature. For $d = 3$, a phase transition can occur at finite temperature only for $N = 1$ (Ising universality class). RG equations then lead to the scaling relation $T_c \propto m$, where m is the low-temperature physical mass.

33.4 The non-linear σ -model in the large N limit

We now discuss another related example, the non-linear σ -model because, when the symmetry is continuous, it is better adapted to a study of the ordered phase and the influence of Goldstone modes. Moreover, due the non-linear character of the group representation, one is confronted with difficulties that are also present in non-Abelian gauge theories. Actually, the non-linear σ -model and non-Abelian theories share another property: both are asymptotically free in the dimensions in which they are renormalizable.

Before discussing the finite temperature properties of the non-linear σ -model by perturbative methods, we discuss them in the large N limit. Large N methods provide interesting information about finite-temperature QFT, because they solve the problem of crossover between different dimensions. Moreover, in Section 18.6.1, it has been proved, within the framework of the $1/N$ expansion, that the non-linear σ -model is equivalent to the $(\phi^2)^2$ field theory (at least for generic $(\phi^2)^2$ coupling), both QFTs corresponding to two perturbative expansions, in different parameters, of the same model (*e.g.* the lattice model in the continuum limit).

The non-linear σ -model. The model has been studied at zero temperature in Chapter 19. It is an $O(N)$ -symmetric QFT, with an N -component scalar field $\mathbf{S}(t, x)$ which belongs to the sphere $\mathbf{S}^2(t, x) = 1$. When required, a lattice regularization with spacing $1/\Lambda$ is assumed.

The partition function of the model can be written as

$$\mathcal{Z} = \int [d\mathbf{S}(t, x)d\lambda(t, x)] \exp [-\mathcal{S}(\mathbf{S}, \lambda)/g], \quad (33.43)$$

where g is the coupling constant of the quantum model, as well as the temperature of the corresponding classical theory in d dimensions, and

$$\mathcal{S} = \frac{1}{2} \int_0^{1/T} dt d^{d-1}x \left[(\partial_t \mathbf{S}(t, x))^2 + (\nabla_x \mathbf{S}(t, x))^2 + \lambda(t, x) (\mathbf{S}^2(t, x) - 1) \right]. \quad (33.44)$$

The λ integration runs along the imaginary axis and enforces the constraint $\mathbf{S}^2(t, x) = 1$. We recall that finite temperature T corresponds to one finite size $L = 1/T$ with periodic boundary conditions in the corresponding d -dimensional classical theory.

33.4.1 The large N limit: Finite temperature saddle point equations

In Section 18.6.2, the non-linear σ -model has been discussed at zero temperature, in the large N limit, with a slightly different notation ($T \rightarrow g$). At finite temperature, the saddle point equation (18.75a) remains unchanged: $m_T^2 \langle \mathbf{S}_T \rangle = 0$, where $\langle \lambda_T \rangle = m_T^2$. The saddle point equation (18.75b) is modified, because the frequencies in the time direction are quantized. In the symmetric phase $\langle \mathbf{S}_T \rangle = 0$, it becomes

$$1 = (N - 1) g G_2(m_T, T), \quad (33.45)$$

where $\xi_T = m_T^{-1}$ has the meaning of a correlation length in space directions, and

$$\begin{aligned} G_2(m_T, T) &= \frac{T}{(2\pi)^{d-1}} \sum_{n \in \mathbb{Z}} \int^\Lambda \frac{d^{d-1}k}{(2\pi n T)^2 + k^2 + m_T^2}, \\ &= \int^\Lambda \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\omega(k)} \left(\frac{1}{2} + \frac{1}{e^{\omega(k)/T} - 1} \right). \end{aligned} \quad (33.46)$$

Introducing the functions (33.24) and (33.23), we can write equation (33.45) (for N large) as

$$1/Ng = \Omega_d(m_T) + T^{d-2} f_d(m_T^2/T^2). \quad (33.47)$$

A phase transition is possible only if $f_d(0)$ is finite for $m_T = 0$. Equation (33.27) then implies $d > 3$. The result has a simple interpretation: for $d = 3$, IR divergences come from the contribution of the zero mode, and are those of a two-dimensional theory, where no phase transition is possible. This property illustrates the effect of *dimensional reduction* $d \mapsto (d - 1)$.

We have seen that, from the point of view of perturbation theory, a crossover between different dimensions is a source of technical difficulties, because IR divergences are more severe in lower dimensions. By contrast, the large N expansion makes it possible to study the problem, because the expansion exists in all dimensions.

Dimension $d = 2$. The case $d = 2$ is a special example of a situation discussed in Section 32.5.2, because, even at zero temperature, the symmetry is unbroken. In the zero-temperature QFT, or the infinite volume classical statistical system, the continuum limit corresponds to $g \ll 1$ and the physical mass m then is given by equation (18.85):

$$1/Ng = \Omega_2(m) \Rightarrow m = \xi_0^{-1} \propto \Lambda e^{-2\pi/Ng}.$$

By subtracting this equation from equation (33.47) (the finite temperature gap equation), one finds

$$\ln(m_T/m) = \ln(\xi_0/\xi_T) = 2\pi f_2(m_T^2/T^2).$$

High temperature corresponds to $T/m \gg 1$, and thus we also expect $m_T \gg m$. The integral (33.24) then is dominated by the contribution of the zero mode and, therefore,

$$\frac{T}{m_T} = \frac{1}{\pi} \ln(m_T/m) \sim \frac{1}{\pi} \ln(T/m). \quad (33.48)$$

The logarithmic decrease of the ratio m_T/T at high temperature corresponds to the UV asymptotic freedom of the classical non-linear σ -model in two dimensions.

Dimensions $d > 2$. In higher dimensions, the system has a phase transition for $T = 0$ at a value g_c of the coupling constant. We can then write the gap equation as

$$\frac{1}{Ng} - \frac{1}{Ng_c} = \Omega_d(m_T) - \Omega_d(0) + T^{d-2} f_d(m_T^2/T^2). \quad (33.49)$$

For $g > g_c$, the equation can be also be written in terms of the physical mass m (equation (18.79)) as

$$[\Omega_d(m) - \Omega_d(m_T)]/T^{d-2} = f_d(m_T^2/T^2). \quad (33.50)$$

The behaviour of the system then depends on the ratio T/m . To obtain more detailed results, we have to specify dimensions.

33.4.2 Dimension $d = 3$

No phase transition can occur at finite temperature, because no phase transition is possible in two dimensions. The gap equation has a scaling form, as predicted by finite-size RG arguments. A short calculation yields

$$f_3(s) = -\frac{1}{2\pi} \ln \left(1 - e^{-\sqrt{s}} \right),$$

and

$$\Omega_3(m_T) - \Omega_3(0) = -\frac{m_T}{4\pi}.$$

For $g > g_c$, after some simple algebra, the gap equation can be written as

$$2 \sinh(m_T/2T) = e^{m/2T}.$$

One verifies that, for m/T large (low temperature), $m_T \rightarrow m$, and that, at high temperature $T \gg m$, m_T becomes proportional to T :

$$m_T/T \sim 2 \ln((1 + \sqrt{5})/2).$$

For $g < g_c$, when the symmetry is broken at zero temperature, one has to return to the general form

$$2 \sinh(m_T/2T) = \exp \left[-\frac{2\pi}{NT} \left(\frac{1}{g} - \frac{1}{g_c} \right) \right]. \quad (33.51)$$

One can also introduce the mass scale $m_{\text{cr}}(g)$ (see equations (18.83) and (33.59)):

$$m_{\text{cr}}(g) = \frac{1}{g} - \frac{1}{g_c},$$

which, at zero temperature, characterizes the crossover between critical and Goldstone mode behaviours. Then,

$$2 \sinh(m_T/2T) = e^{-2\pi m_{\text{cr}}/NT}.$$

For $g < g_c$, the zero mode dominates if the ratio m_T/T is small, and thus if m_{cr}/T is large. This condition is realized for all temperatures if $|g - g_c|$ is not small, because then $m_{\text{cr}} = O(\Lambda) \gg T$: this is the situation of chiral perturbation theory and corresponds to the deep IR (perturbative) region where only Goldstone particles propagate. It is also realized in the critical domain $|g - g_c|$ small, if $T \ll m_{\text{cr}}$, that is, at low (but non-vanishing) temperature. Then,

$$m_T \sim T e^{-2\pi m_{\text{cr}}/NT} = T \exp \left[-\frac{2\pi}{NT} \left(\frac{1}{g} - \frac{1}{g_c} \right) \right]. \quad (33.52)$$

Note that, when the coupling constant g or the temperature T go to 0, the mass m_T has an exponential behaviour characteristic of the dimension 2.

The property that dimensional reduction makes sense at low temperature is somewhat surprising. This peculiar phenomenon is in fact a precursor of the existence of a broken phase at zero temperature.

33.4.3 Dimensions $d > 3$: Critical and high temperatures

For $d > 3$, the quantity $f_d(0)$ is finite and, therefore, a phase transition at finite temperature is possible, in agreement with dimensional reduction, and the property that a phase transition is possible in dimensions larger than 2 (in the case of continuous symmetries). From equation (33.49), one infers

$$T_c^{d-2} = \frac{1}{N f_d(0)} \left(\frac{1}{g} - \frac{1}{g_c} \right). \quad (33.53)$$

Since $f_d(0)$ is positive, this result confirms that a transition is possible only for $g < g_c$, that is, if at zero temperature the symmetry is broken.

However, this result is meaningful only if $T \ll \Lambda$, and thus only for $|g - g_c|$ small. Then, T_c can be compared with the crossover scale $m_{\text{cr}}(g)$ between critical and Goldstone behaviours, which, in all dimensions near g_c , has the same scaling property as the physical mass above g_c .

Dimension d = 4. For $d = 4$, since $f_4(0) = 1/12$, one finds

$$T_c^2 = \frac{12}{N} \left(\frac{1}{g} - \frac{1}{g_c} \right) \propto m_{\text{cr}}^2 \ln(\Lambda/m_{\text{cr}}) \gg m_{\text{cr}}^2. \quad (33.54)$$

Another limit of interest is the high temperature, or massless limit. For $m_T \neq 0$, an additional cut-off dependence appears:

$$\Omega_4(m_T) - \Omega_4(0) \sim -\frac{m_T^2}{8\pi^2} \ln(\Lambda/m_T).$$

At $g = g_c$, the ratio m_T/T decreases logarithmically with the cut-off. At leading order, using $f_4(0) = 1/12$, one finds

$$(m_T/T)^2 = \frac{2\pi^2}{3 \ln(\Lambda/T)},$$

in agreement with equation (33.41).

Dimension d = 5. From $f_5(0) = \zeta(3)/4\pi^2$, one infers the critical temperature given by

$$T_c^3 \sim \frac{4\pi^2}{N\zeta(3)} \left(\frac{1}{g} - \frac{1}{g_c} \right) \propto \Lambda m_{\text{cr}}^2 \gg m_{\text{cr}}^3.$$

Thus for $d \geq 4$, the critical temperature T_c is large in the relevant physical scale.

In the massless limit $g = g_c$,

$$(m_T/T)^2 \sim \frac{\zeta(3)}{4\pi^2} \frac{T}{D_5(0)}, \quad (33.55)$$

$(D_5(0) \propto \Lambda)$ in agreement with the behaviour found in Section 33.3.1.

The $(\phi^2)^2$ field theory for N large. To compare with the situation in the $(\phi^2)^2$ theory of Section 33.3.2, it is interesting to also consider the corresponding gap equation for $d = 5$ in the large N limit. One finds

$$m_T^2 = \frac{N\lambda}{6\Lambda} (G_2(m_T, T) - \Omega_5(0)).$$

For m_T/T small, one infers

$$(m_T/T)^2 \sim \frac{\zeta(3)}{4\pi^2} \frac{T}{(6\Lambda/N\lambda) + D_5(0)} = \frac{N}{6} \frac{\zeta(3)}{4\pi^2} \lambda_T,$$

with

$$\lambda_T = \lambda \frac{T}{\Lambda} \frac{1}{1 + N\lambda D_5(0)/6\Lambda}.$$

The result is consistent with the behaviour found in Section 33.3.1, and the behaviour in equation (33.55).

33.5 The perturbative non-linear σ -model at finite temperature

We now consider the non-linear σ -model in the context of perturbation theory and RG. When the temperature-dependent mass is much smaller than the temperature, dimensional reduction is useful. We thus derive the reduced effective action for the non-linear σ -model. Because the field lives on a sphere, a simple mode expansion somewhat destroys the geometry of the model. For example, one can solve the constraint $\mathbf{S}^2(t, x) = 1$ by parametrizing the field $\mathbf{S}(t, x)$ as

$$\mathbf{S}(t, x) = \{\sigma(t, x), \boldsymbol{\pi}(t, x)\},$$

and eliminate locally the field $\sigma(t, x)$, using $\sigma(t, x) = (1 - \boldsymbol{\pi}^2(t, x))^{1/2}$.

One then performs a mode expansion of $\boldsymbol{\pi}(t, x)$, and integrates perturbatively over the non-zero modes. After mode expansion, the $O(N)$ symmetry is no longer explicit, and this is a source of a difficulty. Otherwise, provided one uses dimensional regularization (or lattice regularization, but the calculations are much more difficult) to deal with the functional measure, this strategy is possible. We will not discuss it here.

Several other strategies are available in which the geometric properties remain obvious. Here we explore two of them and mention a third one.

One convenient method involves parametrizing the zero mode in terms of a time-dependent rotation matrix, which rotates the field zero mode to a standard direction, in the spirit of Section 32.5.2. By contrast, here we describe a method based on the introduction of an auxiliary field. This makes it possible to use a more physical momentum cut-off regularization of Pauli–Villars’s type.

33.5.1 RG equations at finite temperature

We again start from the action in the form (18.72), but rescale the coupling constant $g \mapsto g\Lambda^{2-d}$ in such a way that g becomes dimensionless. The vertex functions of the \mathbf{S} field satisfy the RG equations (Sections 19.11 and 32.1.2),

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \zeta(g) \right] \tilde{\Gamma}^{(n)}(p_i; \Lambda, g, T) = 0. \quad (33.56)$$

For $d > 2$ and $g < g_c$, the solution can be written as

$$\tilde{\Gamma}^{(n)}(p_i; \Lambda, g, T) = m^d(g) M_0^{-n}(g) F^{(n)}(p_i/m(g), T/m(g)), \quad (33.57)$$

with (using $\beta(g) = (d-2)g + O(g^2)$)

$$M_0(g) = \exp \left[-\frac{1}{2} \int_0^g \frac{\zeta(g')}{\beta(g')} dg' \right], \quad (33.58)$$

$$m(g) = \frac{1}{\xi(g)} = \Lambda g^{-1/(d-2)} \exp \left[- \int_0^g \left(\frac{1}{\beta(g')} - \frac{1}{(d-2)g'} \right) dg' \right]. \quad (33.59)$$

The functions $m(g)$ and $M_0(g)$ characterize properties of the zero-temperature theory. The function $M_0(g)$ is proportional to the field expectation value. For $g < g_c$ fixed, the mass scale $m(g)$ is of the size of the cut-off but, for g close to g_c , $m(g)$ becomes much smaller than the cut-off, and becomes a crossover scale between the large momentum critical behaviour and the low-momentum perturbative behaviour.

In the symmetric phase, the definitions of the two functions have to be slightly modified. For example, for $d = 2$,

$$\beta(g) = -\frac{N-2}{2\pi}g^2 + O(g^3), \quad \zeta(g) = \frac{N-1}{2\pi}g + O(g^2), \quad (33.60)$$

and $m(g)$, which now is proportional to the physical mass, is defined by

$$m(g) \propto \Lambda \exp \left[- \int_{g_T}^g \frac{dg'}{\beta(g')} \right] \Rightarrow \ln(m/\Lambda) = -\frac{2\pi}{(N-2)g} + O(\ln g). \quad (33.61)$$

At finite temperature, RG equations can also be solved in another way, by introducing the effective coupling constant g_T at mass scale T :

$$\ln(\Lambda/T) = \int_{g_T}^g \frac{dg'}{\beta(g')}, \quad (33.62)$$

where g_T is a function of g and T only through the combination $m(g)/T$:

$$\ln(m(g)/T) = - \int^{g_T} \frac{dg'}{\beta(g')}. \quad (33.63)$$

For $d > 2$ and $g < g_c$ fixed (which implies $m(g) = O(\Lambda)$), equation (33.62) implies that g_T approaches the IR fixed point $g = 0$ at fixed temperature,

$$g_T \propto (T/m(g))^{d-2}. \quad (33.64)$$

This is a low-temperature regime, where finite temperature effects can be calculated from perturbation theory and RG.

At g_c , and more generally in the critical domain, one can find a high-temperature regime where $m(g) \ll T \ll \Lambda$ and then g_T is close to the UV fixed point g_c . Techniques based on an $\varepsilon = d - 2$ expansion can be used to study this regime.

In two dimensions ($d = 2$), equation (33.63) implies that g_T goes to 0 for $m(g)/T$ small, that is, at high temperature, because $g = 0$ then is a UV fixed point,

$$g_T \sim \frac{2\pi}{(N-2) \ln(m(g)/T)}, \quad (33.65)$$

and this is the limit in which two-dimensional perturbation theory is useful.

33.5.2 Dimensional reduction at one-loop order

We expand the fields in eigenmodes in the time variable and keep the tree and one-loop contributions. We denote by φ, ρ the zero momentum modes, and by $\mathcal{S}_T(\varphi, \rho)$ the reduced $(d-1)$ -dimensional action. At leading order, we find

$$\mathcal{S}_T(\varphi, \rho) = \mathcal{S}(\varphi, \rho)/T. \quad (33.66)$$

The one-loop contribution is

$$\begin{aligned} \mathcal{S}_T^{(1)}(\varphi, \rho) &= \frac{1}{2} N \text{tr} \ln(-\nabla^2 + \rho) + \frac{1}{2} \text{tr} \ln [\varphi(-\nabla^2 + \rho)^{-1} \varphi] \\ &= \frac{1}{2}(N-1) \text{tr} \ln(-\nabla^2 + \rho) + \frac{1}{2} \text{tr} \ln [\varphi(-\nabla^2 + \rho)^{-1} \varphi(-\nabla^2 + \rho)]. \end{aligned} \quad (33.67)$$

The form of the last term may surprise, until one notes that the perturbative expansion is performed around a non-vanishing value of φ .

We use the identity, obtained after a few commutations,

$$\begin{aligned}\varphi(-\nabla^2 + \rho)^{-1}\varphi(-\nabla^2 + \rho) &= \varphi \cdot \varphi + \varphi(-\nabla^2 + \rho)^{-1}[\nabla^2, \varphi] \\ &= \varphi \cdot \varphi + \varphi(-\nabla^2 + \rho)^{-1}[(\nabla^2\varphi) + 2\nabla\varphi \cdot \nabla].\end{aligned}$$

At this order $\varphi \cdot \varphi = 1$, and we expect that ρ can be neglected, because it yields interactions of higher dimensions, and is thus irrelevant at large distance. The expansion of the tr ln then generates a first term with two derivatives. Higher orders yield terms with additional derivatives, which also are subleading at large distance. The first term yields

$$\text{tr } \varphi(-\nabla^2 + \rho)^{-1}[(\nabla^2\varphi) + 2\nabla\varphi \cdot \nabla] \sim \text{tr}(\nabla\varphi)^2(-\nabla^2)^{-1},$$

where the relations

$$\varphi \cdot \nabla\varphi = 0, \quad (\nabla\varphi)^2 + \varphi \cdot \nabla^2\varphi = 0,$$

valid at leading order, have been used.

In the same way, we expand the first term in equation (33.67) in powers of the field ρ . At leading order, only one term is relevant, and we thus obtain

$$\mathcal{S}_T^{(1)} = \frac{1}{2T} \bar{G}_2(0, T) \int d^{d-1}x [(\nabla_x\varphi(x))^2 + (N-1)\rho(x)], \quad (33.68)$$

where the constant \bar{G}_2 defined by equation (33.22), is given by (equation (33.26))

$$\bar{G}_2(0, T) = f_d(0)T^{d-2} - T\Omega_{d-1}(0) + \Omega_d(0). \quad (33.69)$$

We conclude that at one-loop order

$$\mathcal{S}_T(\varphi, \rho) = \frac{\Lambda^{d-2}}{2gT} \int d^{d-1}x [(Z_\varphi/Z_g)(\nabla_x\varphi(x))^2 + \rho(x)(\varphi^2(x) - Z_\varphi^{-1})], \quad (33.70)$$

with

$$Z_g = 1 + (N-2)\Lambda^{2-d}\bar{G}_2 g + O(g^2), \quad (33.71a)$$

$$Z_\varphi = 1 + (N-1)\Lambda^{2-d}\bar{G}_2 g + O(g^2). \quad (33.71b)$$

Dimension d = 3. For $d = 3$, the constant \bar{G}_2 in equation (33.69) has a UV contribution that is three dimensional, $\Omega_3(0) \propto \Lambda$, and a two-dimensional contribution of order $\ln(\Lambda/T)$, corresponding to the omitted zero mode:

$$\bar{G}_2 = -\frac{T}{2\pi}(\ln(\Lambda/T) + \kappa_3) + \Omega_3(0),$$

where κ_3 is a constant. The term $\Omega_3(0)$ generates finite renormalizations of g ,

$$g_r = g + (N-2)(\Omega_3(0)/\Lambda)g^2,$$

and of the field φ :

$$\varphi = [1 - \frac{1}{2}(N-1)(\Omega_3(0)/\Lambda)g]\varphi_r.$$

We introduce the effective coupling constant

$$g_T = g_r T / \Lambda.$$

The effective action then becomes

$$\mathcal{S}_T(\varphi_r, \rho_r) = \frac{1}{2g_T} \int d^2x \left[(\tilde{Z}_\varphi / \tilde{Z}_g) (\nabla_x \varphi_r(x))^2 + \rho_r(x) \left(\varphi_r^2(x) - \tilde{Z}_\varphi^{-1} \right) \right]. \quad (33.72)$$

We verify that the remaining factors $\tilde{Z}_g, \tilde{Z}_\varphi$ render the reduced theory one-loop finite:

$$\begin{aligned} \tilde{Z}_g &= 1 - \frac{N-2}{2\pi} (\ln(\Lambda/T) + \kappa_3) g_T + O(g_T^2), \\ \tilde{Z}_\varphi &= 1 - \frac{N-1}{2\pi} (\ln(\Lambda/T) + \kappa_3) g_T + O(g_T^2). \end{aligned}$$

The solution of the two-dimensional non-linear σ -model then requires non-perturbative techniques, but the two-dimensional RG implies

$$\ln(m_T/T) \propto -\frac{2\pi}{(N-2)g_T} = -\frac{2\pi\Lambda}{(N-2)gT} = -\frac{2\pi}{N-2} \frac{m(g)}{T}, \quad (33.73)$$

where the last equation involves the three-dimensional RG. The result is consistent with equation (33.52).

Dimension d = 2. Then,

$$\bar{G}_2 = \frac{1}{2\pi} [\ln(\Lambda/T) + \kappa_2].$$

The reduced one-dimensional theory is finite. Therefore, Z_φ and Z_g are the renormalization factors which are associated with the change from the scale Λ to the temperature scale T . We set

$$\varphi_r = Z_\varphi^{1/2} \varphi = [1 + (N-1)\bar{G}_2 g/2] \varphi, \quad (33.74a)$$

$$\frac{1}{g_T} = \frac{1}{g Z_g} = \frac{1}{g} - (N-2)\bar{G}_2 + O(g). \quad (33.74b)$$

Both quantities Z_φ and g_T satisfy the RG equations of the zero-temperature field theory:

$$\Lambda \frac{\partial Z_\varphi}{\partial \Lambda} + \beta(g) \frac{\partial Z_\varphi}{\partial g} - \zeta(g) Z_\varphi = 0, \quad \Lambda \frac{\partial g_T}{\partial \Lambda} + \beta(g) \frac{\partial g_T}{\partial g} = 0,$$

where the RG functions at this order are given in equation (33.60).

The one-dimensional non-linear σ -model cannot be expanded in perturbation theory, but can be solved exactly. The difference between the energies of the ground state energy and first excited state is

$$m_T = \frac{1}{2}(N-1)Tg_T.$$

Expressing g_T by equation (33.65) in terms of the mass scale (33.61), which is proportional to the physical mass, we obtain

$$\frac{T}{m_T} \sim -\frac{1}{\pi} \frac{N-2}{N-1} \ln(m/T), \quad (33.75)$$

a result consistent with equation (33.48). The result reflects the UV asymptotic freedom of the non-linear σ -model in two dimensions; the effective coupling constant decreases at high temperature where $m/T \rightarrow 0$.

33.5.3 Matching conditions

If the explicit form of the reduced theory can be guessed, another strategy is available, based on matching conditions [346]. The idea is to calculate some physical observables in d dimensions and to expand them for high temperature. One then calculates the same quantities in the guessed reduced theory in $(d-1)$ dimensions. Identifying the two set of results, one relates the parameters of the initial and reduced actions. One advantage of the method is the possibility to check the ansatz of dimensional reduction by calculating more quantities than needed, and requiring consistency. In addition, one has a better control of the correspondence in the case of large momentum effects. The main drawback is that one is often led to calculate explicit expressions, here the two-point correlation function in an external field, where the main part is not useful (related to IR properties). Contributions of the zero mode have to be separated for each diagram.

To guess the reduced theory, the main guiding principles are power counting and symmetries, as usual for effective low-energy field theories.

In the absence of a Lagrange multiplier, a Pauli–Villars cut-off does not regularize the $O(N)$ -invariant functional measure. The more ‘physical’ lattice regularization is available, but explicit calculations are tedious. For explicit calculations, we thus use the non-physical dimensional regularization, in which the divergent contributions of the functional measure cancel. Moreover, in the dimensions of interest, it is necessary to add an explicit symmetry breaking linear in the field, to avoid IR divergences. Once the correspondence between the parameters of the finite temperature and the reduced theories has been determined, one can take the symmetric limit.

We consider the finite temperature d -dimensional theory corresponding to the action

$$\mathcal{S}(\mathbf{S}) = \frac{\Lambda^{d-2}}{g} \int dt d^{d-1}x \left\{ \frac{Z_S}{2Z_g} \left[(\partial_t \mathbf{S}(t, x))^2 + (\nabla_x \mathbf{S}(t, x))^2 \right] - \mathbf{h} \cdot \mathbf{S}(t, x) \right\}, \quad (33.76)$$

where the MS scheme (see Section 10.4) is used to define renormalization constants, and

$$\mathbf{S}^2(t, x) = Z_S^{-1}. \quad (33.77)$$

Therefore, g is the effective coupling constant at scale Λ .

RG equations for vertex functions in an external field \mathbf{h} take the form,

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \zeta(g) + \rho(g) h \frac{\partial}{\partial h} \right] \tilde{\Gamma}^{(n)}(p_i; \Lambda, g, h, T) = 0, \quad (33.78)$$

where $h = |\mathbf{h}|$. The new RG function ρ is related to β and ζ :

$$\rho(g) = 1 - d + \frac{1}{2} \zeta(g) + \beta(g)/g. \quad (33.79)$$

Dimensional reduction. We compare the finite temperature vertex functions with those of a zero-temperature $(d-1)$ -dimensional Euclidean theory, with the action

$$\mathcal{S}(\varphi) = \frac{T^{d-3} Z_\varphi}{2g_T \tilde{Z}_g} \int d^{d-1}x (\nabla_x \varphi(x))^2 - \frac{T^{d-3}}{g_T} \int d^{d-1}x \mathbf{h} \cdot \varphi(x), \quad (33.80)$$

where the MS scheme again is used to define renormalization constants, and

$$\varphi^2(x) = Z_\varphi^{-1}. \quad (33.81)$$

The coupling constant g_T is the effective coupling constant at the temperature scale T .

We expect that between the two fields \mathbf{S} and φ a finite renormalization is required.

The one-loop diagrams are listed in Fig. 19.1. In the reduced model, at one-loop order, the two-point vertex function is (Section 19.11)

$$\tilde{\Gamma}_{d-1}^{(2)}(p) = \frac{T^{d-3}}{g_T} \left(p^2 Z_\varphi / \tilde{Z}_g + h Z_\varphi^{1/2} \right) + [p^2 + \frac{1}{2}(N-1)h] \Omega_{d-1}(\sqrt{h}) + O(g_T).$$

By contrast, at finite temperature, in the d -dimensional theory, one finds

$$\tilde{\Gamma}_d^{(2)}(p_0 = 0, p) = \frac{\Lambda^{d-2}}{g} \left(p^2 Z_S / Z_g + h Z_S^{1/2} \right) + [p^2 + \frac{1}{2}(N-1)h] G_2(\sqrt{h}, T) + O(g),$$

where the function G_2 is defined in equation (33.46). In the limit $h = 0$,

$$G_2(0, T) = f_d(0) T^{d-2}. \quad (33.82)$$

Dimension $d = 3$. For $d \rightarrow 3$, the renormalization constants at one-loop order in the MS scheme are

$$\tilde{Z}_g = 1 + (N-2) \frac{N_d}{d-3} g_T, \quad Z_\varphi = 1 + (N-1) \frac{N_d}{d-3} g_T. \quad (33.83)$$

In particular,

$$Z_\varphi / \tilde{Z}_g = 1 + \frac{N_d}{d-3} g_T.$$

Therefore, the renormalized two-dimensional two-point function reads

$$\tilde{\Gamma}_2^{(2)}(p) = \frac{1}{g_T} (p^2 + h) + [p^2 + \frac{1}{2}(N-1)h] I_r(h),$$

with

$$I_r(h) = \lim_{d \rightarrow 3} \Omega_{d-1}(\sqrt{h}) + \frac{N_d}{d-3} T^{d-3} = -\frac{1}{4\pi} \ln(h/T^2). \quad (33.84)$$

In the finite temperature theory, no renormalization is required because the theory is not renormalizable, and dimensional regularization cancels all power divergences. Thus,

$$\tilde{\Gamma}_3^{(2)}(p) \underset{d \rightarrow 3}{=} \frac{\Lambda}{g} (p^2 + h) + [p^2 + \frac{1}{2}(N-1)h] G_2(\sqrt{h}, T) + O(g), \quad (33.85)$$

with, for $d = 3$ and $T^2 \gg h$,

$$G_2(\sqrt{h}, T) = -\frac{T}{4\pi} \ln(h/T^2) + O(\sqrt{h}).$$

We note that, at this order, no field renormalization beyond the trivial rescaling $\mathbf{S} = \varphi \sqrt{T}$ is required to compare the two functions and then

$$\frac{1}{g_T} = \frac{\Lambda}{Tg} + O(g).$$

Dimension $d = 2$. In $d = 2$ dimensions, the reduced theory has no divergences, and the one-loop expression reads

$$\tilde{\Gamma}_1^{(2)}(p) = \frac{1}{Tg_T}(p^2 + h) + [p^2 + \frac{1}{2}(N-1)h] \frac{1}{2\sqrt{h}} + O(g_T).$$

We compare this expression with the finite temperature two-point function, calculated in the MS scheme (with renormalization scale Λ). For this purpose, we have to subtract to expression (33.82) the MS counter-term. For $d \rightarrow 2$, we find (γ is Euler's constant),

$$[G_2]_r(0, T) = \frac{1}{2\pi} (\ln(\Lambda/T) + \gamma - \ln(4\pi)) \quad (33.86)$$

and, therefore,

$$\begin{aligned} \tilde{\Gamma}_2^{(2)}(p) &= \frac{1}{g} (p^2 + h) \\ &+ [p^2 + \frac{1}{2}(N-1)h] \left[\frac{1}{2\sqrt{h}} + \frac{1}{2\pi T} (\ln(\Lambda/T) + \gamma - \ln(4\pi)) \right] + O(g). \end{aligned} \quad (33.87)$$

In this case, a field renormalization is required. We set

$$\sqrt{T}\varphi(x) = \mathbf{S}(x)\sqrt{Z_\varphi}\mathbf{s}, \quad Z_{\mathbf{S}\varphi} = 1 + (N-1)(\ln(\Lambda/T) + \gamma - \ln(4\pi))\frac{g}{2\pi},$$

and

$$\frac{1}{g} = \frac{1}{g_T Z_{gt}}, \quad \text{with} \quad Z_{gt} = 1 - (\ln(\Lambda/T) + \gamma - \ln(4\pi))\frac{g}{2\pi},$$

or inverting the relation

$$\frac{1}{g_T} = \frac{1}{g} - \frac{(N-2)}{2\pi} (\ln(\Lambda/T) + \gamma - \ln(4\pi)) + O(g).$$

The results for Z_{gt} and g_T are consistent with equations (33.74).

33.6 The GN model in the large N expansion

To gain some intuition about the role of fermions at finite temperature, we now examine a simple model of self-interacting fermions, the GN model. The GN model is described in terms of a $U(N)$ -symmetric action for a set of N massless Dirac fermions $\{\psi^i, \bar{\psi}^i\}$ (for details, see Section 20.5):

$$\mathcal{S}(\bar{\psi}, \psi) = - \int dt d^{d-1}x \left[\bar{\psi}(t, x) \cdot \not{\partial} \psi(t, x) + \frac{1}{2} G (\bar{\psi}(t, x) \cdot \psi(t, x))^2 \right]. \quad (33.88)$$

The GN model has in all dimensions a discrete symmetry (with the notation $x_0 \equiv t$)

$$\mathbf{x} = \{x_0, \dots, x_\mu, \dots, x_{d-1}\} \mapsto \tilde{\mathbf{x}} = \{x_0, \dots, -x_\mu, \dots, x_{d-1}\}, \quad \begin{cases} \psi(\mathbf{x}) \mapsto \gamma_\mu \bar{\psi}(\tilde{\mathbf{x}}), \\ \bar{\psi}(\mathbf{x}) \mapsto -\bar{\psi}(\tilde{\mathbf{x}}) \gamma_\mu, \end{cases}$$

which prevents the addition of a mass term.

In even dimensions, the discrete symmetry is equivalent to a discrete chiral symmetry, and in odd dimensions, it corresponds to space reflection. In what follows, to simplify we will speak about chiral symmetry, irrespective of dimensions.

The GN model is renormalizable in $d = 2$ dimension, where it is asymptotically free, and its chiral symmetry is always broken at zero temperature.

It has been proved in Section 20.3.1 that, within the $1/N$ expansion, the GN model is equivalent to the GNY (Y for Yukawa) model, at least for generic couplings: the GNY model has the same symmetry, but contains an additional scalar particle coupled to fermions through a Yukawa-like interaction, and is renormalizable in four dimensions. This equivalence provides a simple interpretation to some of the results that follow.

At finite temperature, due to the anti-periodic boundary conditions, fermions have no zero modes. Therefore, limited insight about the physics of the model at high temperature can be gained from perturbation theory; all fermions can simply be integrated out. By contrast, here we study the GN model within the framework of the $1/N$ expansion. After integration over fermions, we obtain the non-local action \mathcal{S}_N for a periodic scalar field σ , equivalent to the large N action of the GNY model that we have already discussed in the zero-temperature limit in Section 20.3:

$$\mathcal{S}_N(\sigma) = \frac{1}{2G} \int_0^{1/T} dt \int d^{d-1}x \sigma^2(t, x) - N \text{tr} \ln (\partial + \sigma(\cdot)), \quad (33.89)$$

where T is the temperature, and the σ field satisfies periodic boundary conditions in the Euclidean time variable. Additional effects due to the addition of a chemical potential are not considered here. As we have seen, a non-trivial perturbation theory is obtained by expanding for N large.

In the situations in which the σ mass is small compared to the temperature, one can perform a mode expansion of the σ field, integrate over the non-zero modes, and obtain a local $(d-1)$ -dimensional action for the zero mode. It is important to realize that, since the resulting reduced action is local and symmetric in $\sigma \mapsto -\sigma$, it describes the physics of an Ising-like transition with short-range interactions (unlike what happens at zero temperature). The question which then arises is the possibility of a breaking of this remaining symmetry of the Ising type. If a transition exists and is continuous, the σ -mass vanishes at the transition and a potential IR problem appears.

If a phase transition occurs at finite temperature, and if it is second order, IR divergences generated by the σ zero mode appear in the $1/N$ perturbation theory at the transition temperature T_c for $d-1 < 4$. Below T_c , as is the case at zero temperature, the decay of σ -correlation functions in space directions is governed by the saddle point value M_T of the field $\sigma(x)$. Above T_c , the correlation length is also finite, in contrast with the zero-temperature situation.

We then find two regimes, which have to be dealt with differently. Near the transition temperature, $1/N$ perturbation theory is not useful for $d < 5$. By contrast, one has to perform a mode expansion of the σ field and a local expansion of the dimensionally reduced action for the σ zero mode. The EFT relevant for large-distance properties is of σ^4 type (as in the case of the Ising model), with coefficients depending on coupling constant and temperature, which has to be studied by the usual RG methods. Note that this implies the absence of phase transition for $d = 2$ at finite temperature. In the other regime, where the σ correlation length is of order $1/T$, all modes are similar and $1/N$ perturbation theory is directly applicable.

33.6.1 The gap equation

We denote by M_T the saddle point value of the field $\sigma(x)$, which, due to the anti-periodic boundary conditions, no longer governs the decay of the fermion two-point function in space directions. We obtain the action density at finite temperature and large N ,

$$\mathcal{E}(M_T) = \frac{M_T^2}{2G} - \frac{NT}{V_{d-1}} \text{tr} \ln(\mathcal{J} + M_T), \quad (33.90)$$

where V_{d-1} is the $(d-1)$ -dimensional volume. Then,

$$N \text{tr} \ln(\mathcal{J} + M_T) = N' V_{d-1} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \sum_{n \in \mathbb{Z}} \ln(\omega_n^2 + k^2 + M_T^2),$$

where $N' = N \text{tr } \mathbf{1}$ is the total number of fermions, and $\omega_n = (2n+1)\pi T$. The sum over frequencies follows from the identity (A33.2) and one obtains

$$N \text{tr} \ln(\mathcal{J} + M_T) = N' V_{d-1} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln [2 \cosh(\omega(k)/2T)], \quad (33.91)$$

with $\omega(k) = \sqrt{k^2 + M_T^2}$. Alternatively, one could use Schwinger's representation of the propagator and another function of elliptic type

$$\vartheta_1(s) = \sum_n e^{-(n+1/2)^2 \pi s}. \quad (33.92)$$

The gap equation at finite temperature, obtained by differentiating \mathcal{E} with respect to M_T , again factorizes into two equations $M_T = 0$, and

$$1/G = N' \mathcal{G}_2(M_T, T), \quad (33.93)$$

with

$$\mathcal{G}_2(M_T, T) = \int^\Lambda \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\omega(k)} \left(\frac{1}{2} - \frac{1}{e^{\omega(k)/T} + 1} \right), \quad (33.94)$$

and $\omega(k) = \sqrt{k^2 + M_T^2}$. This is the fermion analogue of equations (33.45) and (33.46), and again one recognizes the sum of quantum and thermal contributions.

Note that the function $\mathcal{G}_2(M_T, T)$ has a regular expansion in M_T^2 at $M_T = 0$.

Introducing the function

$$\begin{aligned} g_d(s) &= N_{d-1} \int_0^\infty \frac{x^{d-2} dx}{\sqrt{x^2 + s}} \frac{1}{\exp[\sqrt{x^2 + s}] + 1} \\ &= N_{d-1} \int_{\sqrt{s}}^\infty (y^2 - s)^{(d-3)/2} \frac{dy}{e^y + 1}, \end{aligned} \quad (33.95)$$

where N_d is given in (33.25), we write the gap equation as

$$1/N' G = \Omega_d(M_T) - T^{d-2} g_d(M_T^2/T^2). \quad (33.96)$$

For $d > 2$, we can introduce the critical value G_c where $M \equiv M_{T=0}$ vanishes at zero temperature:

$$\frac{1}{N'} \left(\frac{1}{G} - \frac{1}{G_c} \right) = \Omega_d(M_T) - \Omega_d(0) - T^{d-2} g_d(M_T^2/T^2). \quad (33.97)$$

For $G > G_c$, we also introduce the fermion physical mass $M \equiv m_\psi$ solution of

$$\frac{1}{N'} \left(\frac{1}{G} - \frac{1}{G_c} \right) = \Omega_d(M) - \Omega_d(0). \quad (33.98)$$

The gap equation can then be written as

$$\Omega_d(M_T) - \Omega_d(M) = T^{d-2} g_d(M_T^2/T^2). \quad (33.99)$$

Critical temperature. The difference $\Omega_d(0) - \Omega_d(M_T)$ is always positive. Therefore, the gap equation has no solution for $G \leq G_c$, that is, when the chiral symmetry is unbroken at zero temperature. Then, $M_T = 0$ is the minimum, and the $\sigma \mapsto -\sigma$ symmetry is not broken (as expected) in the infinite N limit.

For $d > 2$, $g_d(0)$ is finite:

$$g_d(0) = N_{d-1} (1 - 2^{3-d}) \Gamma(d-2) \zeta(d-2),$$

($\zeta(s)$ is Riemann's function) and thus, if $G > G_c$, by contrast, one always finds a transition temperature

$$T_c = \left[\frac{1}{N' g_d(0)} \left(\frac{1}{G_c} - \frac{1}{G} \right) \right]^{1/(d-2)} = \left[\frac{\Omega_d(0) - \Omega_d(M)}{g_d(0)} \right]^{1/(d-2)}, \quad (33.100)$$

where M_T vanishes, between two Ising-like phases: a low-temperature broken phase and a symmetric high-temperature phase,

The σ two-point function. Since the correlation length of the σ zero mode plays a crucial role, we also calculate the σ two-point function $\tilde{\Delta}_\sigma(p) \equiv \tilde{\Delta}_\sigma(p_0 = 0, p)$ (see equation (20.53)). For $M_T = 0$, we find

$$\tilde{\Delta}_\sigma^{-1}(p) = \frac{1}{G} - N' \mathcal{G}_2(0, T) + \frac{N' T}{2(2\pi)^{d-1}} p^2 \sum_n \int^\Lambda \frac{d^{d-1}k}{(k^2 + \omega_n^2) [(p+k)^2 + \omega_n^2]}. \quad (33.101)$$

For $d > 2$, the inverse propagator can be expressed in terms of the constant $g_d(0)$ as

$$\begin{aligned} \tilde{\Delta}_\sigma^{-1}(p) &= \frac{1}{G} - \frac{1}{G_c} + N' T^{d-2} g_d(0) \\ &+ \frac{N' T}{2(2\pi)^{d-1}} p^2 \sum_n \int^\Lambda \frac{d^{d-1}k}{(k^2 + \omega_n^2) [(p+k)^2 + \omega_n^2]}. \end{aligned} \quad (33.102)$$

For $M_T \neq 0$, using the gap equation, we rewrite the inverse propagator as

$$\tilde{\Delta}_\sigma^{-1}(p) = \frac{N' T (p^2 + 4M_T^2)}{2(2\pi)^{d-1}} \sum_n \int^\Lambda \frac{d^{d-1}k}{(k^2 + \omega_n^2 + M_T^2) [(p+k)^2 + \omega_n^2 + M_T^2]}. \quad (33.103)$$

Therefore, when the symmetry $\sigma \mapsto -\sigma$ is broken, the correlation length is $1/2M_T$ (or the σ mass $2M_T$), generalizing the zero-temperature result. More detailed properties require considering specifying dimensions.

Local expansion. When the σ mass or expectation value are small compared to T , since the fermion mass is still of order T , we can perform a mode expansion of the field σ , retaining only the zero mode, and then a local expansion of the action (33.89), and study it to all orders in the $1/N$ expansion. In the reduced theory, T plays the role of a large momentum cut-off.

The first terms of the effective $(d-1)$ -dimensional action are

$$\mathcal{S}_{d-1}(\sigma) = \int d^{d-1}x \left[\frac{1}{2}Z_\sigma (\nabla_x \sigma(x))^2 + \frac{1}{2}r\sigma^2(x) + \frac{1}{4!}u\sigma^4(x) \right], \quad (33.104)$$

where terms of order σ^6 and $\nabla^2\sigma^4$, and higher orders, have been neglected. The three coefficients in the action are given by

$$Z_\sigma = \frac{1}{2}N'\mathcal{G}_4(0, T)/T, \quad r = [1/G - N'\mathcal{G}_2(0, T)]/T, \quad u = 6N'\mathcal{G}_4(0, T)/T,$$

and $\mathcal{G}_4(M_T, T)$ can be calculated from

$$\mathcal{G}_4(M_T, T) = -\frac{\partial}{\partial M_T^2} \mathcal{G}_2(M_T, T).$$

For $d < 4$, $\mathcal{G}_4(0, T)$ is finite and thus proportional to T^{d-4} . For $d > 2$, after the shift of the coupling constant, one finds

$$rT = \frac{1}{G} - \frac{1}{G_c} - N'T^{d-2}g_d(0) = N'g_d(0) (T_c^{d-2} - T^{d-2}). \quad (33.105)$$

As we have already explained, the properties of this model are those of the critical ϕ^4 theory and, for $d < 5$, have to be studied by non-perturbative techniques.

33.6.2 Scalar mass

Dimension $d > 4$. For $d > 4$, the critical temperature scales like

$$T_c \propto M(\Lambda/M)^{(d-4)/(d-2)} \Rightarrow M \ll T_c \ll \Lambda$$

and, therefore, T_c is physical and large compared to the particle masses. In the symmetric high-temperature phase $T > T_c$, the behaviour of the σ mass m_σ (inverse correlation length) is given by

$$m_\sigma^2 \propto T^2(T/\Lambda)^{d-4} [1 - (T_c/T)^{d-2}],$$

and thus m_σ is small compared to T , justifying dimensional reduction. For $T < T_c$, but $T \gg M$, one finds $|M_T^2 - M^2| \ll T^2$ and the property remains true.

Dimension $d = 4$. In the high-temperature symmetric phase, the σ mass parameter is given by

$$m_\sigma^2 \propto \frac{1}{\ln(\Lambda/T)} \left(\frac{1}{G} - \frac{1}{G_c} + N'T^2g_4(0) \right).$$

The σ particle only propagates for $|1/G - 1/G_c| \ll \Lambda^2$, that is, in the critical domain of the zero-temperature theory. For $G > G_c$, we can introduce the critical temperature (equation (33.100)). Then,

$$m_\sigma^2 \propto \frac{1}{\ln(\Lambda/T)} (T^2 - T_c^2).$$

Eventually, m_σ vanishes as $(T - T_c)^{1/2}$ a typical mean-field behaviour, and a phase transition occurs. Equation (33.100) yields T_c , which expressed in terms of the physical fermion mass M is given by

$$(T_c/M)^2 \sim \frac{3}{\pi^2} \ln(\Lambda/M).$$

The critical temperature is thus large compared to the physical mass M .

In the broken phase, for T/M finite, the mass parameter M_T remains close to M , and one finds

$$\left(\frac{M_T}{M}\right)^2 = 1 - 8\pi^2 g_4(M^2/T^2) \left(\frac{T}{M}\right)^2 \frac{1}{\ln(\Lambda/M)}.$$

Dimension d = 3. In the symmetric phase,

$$m_\sigma^2 \propto \frac{T}{G} - \frac{T}{G_c} + N'T^2 g_3(0).$$

The σ particle propagates if $T(1/G - 1/G_c) \ll \Lambda^2$. At the transition coupling constant G_c , we find, as expected, $m_\sigma \propto T$.

In the broken phase, since at leading order $\Omega_d(M) - \Omega_d(0) = -M/4\pi$, and

$$g_3(s) = \frac{1}{2\pi} \ln \left(1 + e^{-\sqrt{s}} \right),$$

the gap equation can be written as

$$2 \cosh(M_T/2T) = e^{M/2T},$$

an equation that has the scaling form expected, for $d < 4$, from the correspondence between GN and GNY models, and the existence of an IR fixed point in the latter. The critical temperature is proportional to the fermion mass:

$$\frac{T_c}{M} = \frac{1}{2 \ln 2}.$$

Dimension d = 2. The situation $d = 2$ is doubly special, because at zero temperature chiral symmetry is always broken, and at finite temperature the Ising symmetry is never broken. The GN model is renormalizable and UV free. For N large,

$$\beta(G) = -\frac{N'}{2\pi} G^2.$$

All masses are proportional to the RG-invariant mass scale

$$\Lambda(G) \propto \Lambda \exp \left[- \int^G \frac{dG'}{\beta(G')} \right].$$

In particular, the fermion physical mass M has the form

$$M \propto \Lambda e^{-2\pi/N'G}.$$

At finite temperature, all thermal masses have a scaling property. For example, the σ thermal mass has the form

$$m_\sigma/T = f(M/T).$$

For $T > M$, one can also express the scaling properties by introducing a temperature-dependent coupling constant G_T defined by

$$\int_G^{G_T} \frac{dG'}{\beta(G')} = -\ln(\Lambda/T).$$

At high temperature, G_T decreases like

$$G_T \sim \frac{2\pi}{N' \ln(T/M)}.$$

Therefore, we expect a trivial high-temperature physics with weakly interacting fermions.

At high temperature, the mass parameter m_σ is proportional to T and, therefore, the zero mode is similar to other modes. Eventually, it decreases when T approaches $T_c = M/\pi$. At leading order, one finds

$$m_\sigma^2 \propto T^2 \ln(\pi T/M). \quad (33.106)$$

This result does not imply a phase transition, since for $m_\sigma/T \ll 1$, dimensional reduction is justified, and we are led to an essentially one-dimensional statistical system with short-range interactions, which cannot display a phase transition. Due to statistical fluctuations the correlation length $1/m_\sigma$ never diverges.

For $T < T_c$, the gap equation becomes

$$\ln(M/M_T) = 2\pi g_2(M_T^2/T^2).$$

The function $g_2(s)$ is positive, which again implies $M_T < M$, goes to ∞ for $s \rightarrow 0$, and goes to 0 for $s \rightarrow \infty$. At low temperature, M_T/M converges to 1 exponentially in M/T . At high temperature, M_T/T goes to 0 and

$$g_2(s) \sim -\frac{1}{4\pi} \ln s.$$

The equation implies $M \propto T$ and thus has no solution for $T \rightarrow \infty$, but instead solutions at finite temperature, in agreement with equation (33.106), which shows that m_σ vanishes for some value $T_c \propto M$. The existence of non-trivial solutions to the gap equation here only implies that the σ potential has degenerate minima, but as a consequence of fluctuations, the expectation value of σ nevertheless vanishes.

More precisely, we can apply the expansion (33.104) to the $d = 2$ example. We find a simple model in one-dimensional quantum mechanics: the quartic anharmonic oscillator. Straightforward considerations show that the correlation length, inverse of the σ mass parameter, becomes small only when the coefficient of σ^2 is large and negative. This only happens at low temperature, where the two lowest eigenvalues of the corresponding quantum Hamiltonian are almost degenerate. Then, instantons relate the two classical minima of the potential and restore the symmetry (see Section 39.1). Calculating the difference between the two lowest eigenvalues of the Hamiltonian, one obtains a behaviour of the form

$$m_\sigma/T \propto (\ln M/T)^{5/4} e^{-\text{const.}(\ln M/T)^{3/2}}.$$

Again, the property that m_σ/T is small at low temperature is a precursor of the symmetry breaking at zero temperature.

33.7 Abelian gauge theories: The QED framework

First, we discuss perturbative Abelian gauge theories (Chapter 21), which are much simpler, because the mode expansion is consistent with the gauge structure [349]. Additional problems arising in non-Abelian gauge theories are examined in Section 33.8.

Because the number of components of gauge fields depends on the number of space dimensions, the mode expansion has new properties, which affect gauge transformations. The simplest non-trivial example of a gauge theory is QED, a theory which is IR-free in four dimensions and, therefore, from the RG point of view has properties similar to the scalar ϕ^4 field theory. Another example is provided by the Abelian Higgs model but, since it has a weak first-order phase transition, it has a more limited validity [348].

Notation. In this section, Greek indices refer to space-time components, while Roman indices refer to space components only: $A_\mu \equiv (A_0, A_i)$ with $i = 1, \dots, d - 1$ and $x^\mu \equiv (t, x_i)$. In the symbol $\nabla \cdot \mathbf{X}$, ∇ has d or $(d - 1)$ components depending on a number of components of the vector \mathbf{X} .

Summation over repeated Greek indices is always meant.

33.7.1 Mode expansion and gauge transformations

At finite temperature $T > 0$, we decompose a general gauge field $A_\mu(t, x)$ periodic in time into the sum of a zero mode $B_\mu(x)$ and the sum of all non-zero modes $Q_\mu(t, x)$:

$$A_\mu(t, x) = B_\mu(x) + Q_\mu(t, x),$$

where $Q_\mu(t, x)$ thus satisfies

$$\int_0^{1/T} dt Q_\mu(t, x) = 0. \quad (33.107)$$

The variation $\delta_\Omega A_\mu$ of A_μ in a gauge transformation,

$$\delta_\Omega A_\mu(t, x) = \partial_\mu \Omega(t, x), \quad (33.108)$$

must be a periodic function. In particular, $\partial_t \Omega(t, x)$ must be periodic. Therefore, after integration over time, $\Omega(t, x)$ must have the form

$$\Omega(t, x) = \Theta(x)t + \Omega'(t, x),$$

where $\Omega'(t, x)$ is periodic. Expressing then that $\partial_i \Omega(t, x)$ is periodic, we conclude that $\Theta(x)$ must be a constant. We set

$$\Omega(t, x) = \Theta t + \Omega_0(x) + \Omega_1(t, x), \quad (33.109)$$

where Ω_1 is the sum of non-zero modes. Then,

$$\delta_\Omega B_0(x) = \Theta, \quad \delta B_i(x) = \partial_i \Omega_0(x), \quad (33.110)$$

$(\partial_i \equiv \partial/\partial x_i)$ and

$$\delta_\Omega Q_\mu(t, x) = \partial_\mu \Omega_1(t, x). \quad (33.111)$$

The space components B_i transform as the components of a $(d - 1)$ -dimensional gauge field; the time component B_0 is a $(d - 1)$ -dimensional scalar field, which is translated by a constant. Invariance under the translation (33.110) implies, in the absence of matter fields, that the scalar field B_0 is massless.

Matter fields. We couple the gauge field to matter, for instance, charged fermions $\{\psi(t, x), \bar{\psi}(t, x)\}$. At finite temperature, fermion fields satisfy anti-periodic boundary conditions. The gauge transformation (33.111) corresponds for the fermions to

$$\psi(t, x) = e^{i\Omega(t, x)} \psi'(t, x), \quad \bar{\psi}(t, x) = e^{-i\Omega(t, x)} \bar{\psi}'(t, x).$$

Anti-periodicity implies that

$$\Omega(1/T, x) = \Omega(0, x) \pmod{2\pi}.$$

Since Ω_1 is periodic, this condition implies for the constant Θ in equation (33.109),

$$\Theta = 2n\pi T, \quad n \in \mathbb{Z}. \quad (33.112)$$

This restriction on the transformation (33.110) of the component B_0 has important consequences. As a result of quantum corrections generated by the interactions with charged matter, the scalar field B_0 does not necessarily remain massless, and the thermodynamic potential for constant fields becomes a periodic function of B_0 with period $2\pi T$.

33.7.2 Gauge field coupled to fermions: Quantization

The action for a gauge field coupled to an N -component charged fermion reads

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int dt d^{d-1}x \left[\frac{1}{4e^2} F_{\mu\nu}^2(t, x) - \bar{\psi}(t, x) \cdot (\partial + i\mathcal{A}(t, x)) \psi(t, x) \right], \quad (33.113)$$

where e is the electric charge. We have neglected the fermion mass because we are only interested in high-temperature physics. The theory has RG properties which bear similarities with those of the ϕ^4 theory; it is renormalizable for $d = 4$ and IR free (trivial). It can be solved in the large N limit. Finally, in dimension $d = 2$, it reduces to the massless Schwinger model, which can be solved exactly even at finite temperature, because bosonization methods still work (see Section 30.5).

The temporal gauge. To calculate the partition function, we first quantize the gauge theory in the temporal gauge $A_0(t, x) = 0$, because the corresponding Hamiltonian formalism is simpler (Section 21.5.2). The action becomes

$$\begin{aligned} \mathcal{S}(\bar{\psi}, \psi, A) = \int dt d^{d-1}x & \left[\frac{1}{2e^2} \sum_i \dot{A}_i^2(t, x) + \frac{1}{4e^2} \sum_{i,j} F_{ij}^2(t, x) \right] \\ & - \bar{\psi}(t, x) \cdot (\partial + i\mathcal{A}(t, x)) \psi(t, x). \end{aligned} \quad (33.114)$$

In the calculation of the partition function $\text{tr } e^{-H/T}$, we have to take into account Gauss's law. It implies that the trace has to be restricted to the subspace of states invariant under time-independent gauge transformations. To project onto this subspace, we impose periodic conditions in the Euclidean-time direction, up to a gauge transformation:

$$\begin{aligned} A_i(1/T, x) &= A_i(0, x) - \partial_i \Omega(x)/T, \\ \psi(1/T, x) &= e^{i\Omega(x)/T} \psi(0, x), \end{aligned}$$

and integrate over the gauge transformation $\Omega(x)$.

We then set

$$A_i(t, x) = A'_i(t, x) - t\partial_i\Omega(x)$$

and, correspondingly,

$$\psi(t, x) = e^{it\Omega(x)} \psi'(t, x), \quad \bar{\psi}(t, x) = e^{-it\Omega(x)} \bar{\psi}'(t, x),$$

where the fields A'_i , ψ' , $\bar{\psi}'$ are now periodic and anti-periodic, respectively. This generates two modifications of the action,

$$\begin{aligned} \int dt d^{d-1}x (\partial_t A_i(t, x))^2 &\mapsto \int dt d^{d-1}x (\partial_t A_i(t, x))^2 + \frac{1}{T} \int d^{d-1}x (\partial_i \Omega(x))^2, \\ \int dt d^{d-1}x \bar{\psi}(t, x) \gamma_0 \partial_t \psi(t, x) &\mapsto \int dt d^{d-1}x \bar{\psi}(t, x) \gamma_0 (\partial_t + i\Omega(x)) \psi(t, x). \end{aligned}$$

Therefore, $\Omega(x)$ is simply the residual zero mode of the A_0 component:

$$\Omega(x) \equiv B_0(x).$$

Its presence is a direct consequence of Gauss's law.

The field theory has a $(d-1)$ -dimensional gauge invariance with the zero mode $B_i(x)$ of $A_i(t, x)$ as gauge field. In addition, it contains d families of neutral vector fields with masses $2\pi nT$, $n \neq 0$, quantized in a unitary, and thus non-renormalizable gauge.

Note that, from the technical point of view, the usual difficulties that appear in perturbation calculations with the temporal gauge (the gauge field propagator is singular at $k_0 = 0$, see Section 21.5.2) reduce here to the need for quantizing the remaining zero mode, and to the non-explicit renormalizability. The latter problem can be solved with the help of dimensional regularization, for example (for gauge-invariant observables). An alternative possibility is to change to a renormalizable gauge.

Covariant gauge. The change of gauges follows the standard zero-temperature method (Section 22.3). We first introduce a time component A_0 , without zero mode, for the gauge field (periodic in time) and multiply the functional measure by the corresponding δ -function (symbolically)

$$1 = \int [dQ_0(t, x)] \prod_{t,x} \delta(Q_0(t, x)).$$

Combining Q_0 with $B_0 = \Omega(x)$, we can reintroduce A_0 into the action and write it in a gauge-invariant form.

To change to a covariant gauge, we insert a second identity in the field integral:

$$1 = \det(-\nabla^2) \int [d\Omega_1] \delta(\nabla \cdot \mathbf{Q}(t, x) + \nabla^2 \Omega_1(t, x) - n(t, x)), \quad (33.115)$$

where $\Omega_1(t, x)$ and $n(t, x)$ are two periodic functions without zero mode. We perform the gauge transformation, $Q_\mu + \partial_\mu \Omega_1 \mapsto Q_\mu$. The Ω_1 dependence remains only in $\delta(Q_0 - \partial_t \Omega_1)$, and the integration over Ω_1 yields a constant. An integration over $n(t, x)$ with a Gaussian weight yields the standard covariant gauge

$$\begin{aligned} \mathcal{S}_{\text{gauge}} &= \frac{1}{2\xi T} \int d^{d-1}x (\nabla \cdot \mathbf{B}(x))^2 + \frac{1}{2\xi} \int dt d^{d-1}x (\nabla \cdot \mathbf{Q}(t, x))^2 \\ &\equiv \frac{1}{2\xi} \int dt d^{d-1}x (\nabla \cdot \mathbf{A}(t, x))^2, \end{aligned}$$

since

$$\nabla \cdot \mathbf{A}(t, x) = \partial_t Q_0(t, x) + \sum_i \partial_i B_i(x) + \sum_i \partial_i Q_i(t, x).$$

We conclude that the gauge fixing term is just obtained by substituting the mode decomposition into the gauge fixing term of the zero-temperature action. From the point of view of the B gauge field, this corresponds to a quantization in the covariant gauge in $(d - 1)$ dimensions.

Finally, the transformation from the temporal gauge to the covariant gauge yields a determinant (equation (33.115)) which is field independent, but contributes to the free energy.

33.7.3 Dimensional reduction

At finite temperature, to generate the effective action for the gauge-field zero modes, we have to integrate over all fermion modes (anti-periodic boundary conditions), and over the non-zero modes $Q_\mu(t, x)$ of the gauge field. At leading order, one finds a free theory containing a gauge field B_i and a massless scalar B_0 . At one-loop order, only fermion modes contribute. Replacing the gauge field A_μ by its zero mode B_μ , and performing the fermion integration explicitly, one obtains the effective action

$$\mathcal{S}_T(B) = \frac{1}{T} \int d^{d-1}x \left[\frac{1}{2e^2} \sum_i (\partial_i B_0(x))^2 + \frac{1}{4e^2} \sum_{i,j} B_{ij}^2(x) \right] - N \text{tr} \ln(\mathcal{D} + i\mathcal{B}), \quad (33.116)$$

with $B_{ij} = \partial_i B_j - \partial_j B_i$. An important issue is the behaviour of the induced mass of the time-component B_0 of the gauge field.

The action density. We thus first calculate the action density for constant $B_0(x) = \Theta$. It is given by

$$\mathcal{E}(\Theta) = -\frac{1}{2} N' T \sum_n \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln \left[k^2 + (\Theta + (2n+1)\pi T)^2 \right],$$

where $N' = N \text{tr} \mathbf{1}$ is the total number of fermion degrees of freedom.

The sum over n can be performed with the help of the identity (A33.2), and one obtains

$$\mathcal{E}(\Theta) = -\frac{1}{2} N' T \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln(\cosh(k/T) + \cos(\Theta/T)). \quad (33.117)$$

One verifies that the difference $\mathcal{E}(\Theta) - \mathcal{E}(0)$ is UV finite and has the scaling form $T^d f(\Theta/T)$. This is not surprising, since in the zero-temperature limit, no gauge field mass or quartic Θ potential are generated.

The derivative

$$\mathcal{E}'(\Theta) = \frac{1}{2} N' \sinh(\Theta/T) \frac{1}{(2\pi)^{d-1}} \int \frac{d^{d-1}k}{\cosh(k/T) + \cos(\Theta/T)},$$

is negative for $-\pi < \Theta/T < 0$, and positive for $0 < \Theta/T < \pi$. In the interval $-\pi < \Theta/T < \pi$, the action density has a unique minimum at $\Theta = 0$.

A special case is $d = 2$. One finds

$$\mathcal{E}'(\Theta) = \frac{1}{2} N' \Theta / \pi, \quad \text{for } |\Theta| < \pi, \quad \text{and thus } \mathcal{E}(\Theta) = \frac{1}{4} N' \Theta^2 / \pi.$$

Neglecting all B_0 derivatives, one obtains the contribution to the action \mathcal{S}_T ,

$$-N \text{tr} \ln(\mathcal{D} + i\mathcal{B}) \sim -\frac{1}{2} N' \int d^{d-1}x \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln(\cosh(k/T) + \cos(\Theta(x)/T)).$$

The coefficient $K_2(d)$ of $\frac{1}{2} \int d^{d-1}x \Theta^2(x)$ follows:

$$\begin{aligned} K_2(d) &= N' N_{d-1} \Gamma(d-1) (1 - 2^{3-d}) \zeta(d-2) T^{d-3} \\ &= N' \frac{8}{(4\pi)^{d/2}} \Gamma(d/2) (2^{d-3} - 1) \zeta(d-2) T^{d-3}. \end{aligned}$$

Discussion. At leading order, we thus obtain a mass term proportional to $eT^{(d-2)/2}$. If e is generic, that is, of order 1 at the microscopic scale $1/\Lambda$, then $e \propto \Lambda^{(4-d)/2}$, and the scalar mass m_T is proportional to $(\Lambda/T)^{(4-d)/2} T$. It is thus large with respect to the vector masses for $d < 4$, and small for $d > 4$.

If we take into account loop corrections, we find, for $d > 4$, a finite coupling constant renormalization $e \mapsto e_r$, and the conclusion is not changed. The zero mode becomes massive, but with a mass small compared to T justifying mode and local expansions.

For $d = 4$, QED is IR-free,

$$\beta_{e^2} = \frac{N}{6\pi^2} e^4 + O(e^6),$$

e_r has to be replaced by the effective coupling constant $e(T/\Lambda)$, which is logarithmically small:

$$e^2(T/\Lambda) \sim \frac{6\pi^2}{N \ln(\Lambda/T)},$$

and the scalar mass thus is still small, although only logarithmically,

$$m_{B_0}^2 \propto \frac{T^2}{\ln(\Lambda/T)}.$$

The separation between zero and non-zero modes remains marginally justified. High-temperature QED shares several properties of high-temperature ϕ^4 -field theory, and a perturbative expansion for the same reason remains meaningful.

Note that if one is interested only in IR physics, in a second step one can integrate over the massive scalar field Θ .

Finally, for $d < 4$, one finds an IR fixed point and, therefore, one expects that, in massless QED, m_T becomes proportional to T , and comparable to all other modes, in particular, to all gauge field non-zero modes that become massive vector fields.

Quantization. To quantize the theory, one still has to fix the gauge, using for instance a covariant gauge. The massive modes are quantized here in a unitary non-renormalizable gauge. Therefore, for these also, a change of gauge is required for renormalizability purpose. This is not difficult in the Abelian case, because one can make an independent gauge transformation on each vector field. Furthermore, one notes that the vector masses are not renormalized.

For more details and more systematic QED calculations, we refer to the literature.

33.7.4 The Abelian Higgs model

The field theory of an Abelian gauge field interacting with charged scalar fields has also been investigated, as a toy model for studying properties of the electro-weak phase transition at finite temperature. The gauge action reads

$$\mathcal{S}(A_\mu, \phi) = \int dt d^{d-1}x \left[\frac{1}{4e^2} F_{\mu\nu}^2 t, (x) + |\mathbf{D}_\mu \phi(t, x)|^2 + U(|\phi|^2(t, x)) \right], \quad (33.118)$$

where the function

$$U(s) = rs + \frac{1}{6}gs^2, \quad (33.119)$$

is such that the $U(1)$ symmetry is broken at zero temperature.

The model can directly be quantized in the unitary (non-renormalizable) gauge, and calculations of gauge-independent observables can be performed with dimensional regularization. Instead, we will use the temporal gauge, because the unitary gauge becomes singular near the phase transition.

One limitation of the model is that the RG shows that, in $(3+1)$ dimensions, the hypothesis of second order phase transition is inconsistent. Indeed, in a more general model with N charged scalars, for $d = 4$, at one-loop order the RG β -functions are (equation (21.120))

$$\beta_g = \frac{1}{24\pi^2} [(N+4)g^2 - 18ge^2 + 54e^4], \quad \beta_{e^2} = \frac{1}{24\pi^2} Ne^4. \quad (33.120)$$

Studying the RG flow, one verifies that the origin $e^2 = g = 0$ is a stable IR fixed point only for $N \geq 183$. Therefore, for N smaller, most likely, the transition is weak-first order. Still, the continuum model remains useful if initially the coupling constants are small enough, in such a way that, by the time the running coupling constants reach the physical scale they have not yet reached the region of instability.

Presumably, the same result applies, for small values of N to the three-dimensional statistical field theory, which is also the Landau–Ginzburg model of superconductivity.

Dimensional reduction. To construct the reduced action, we quantize in the temporal gauge $A_0(t, x) = B_0(x)$. We then face the problem that the concept of scalar zero mode is not gauge invariant, since time-dependent gauge transformations with quantized frequencies (33.112) shift the modes. Note that the problem is avoided in the $1/N$ expansion.

We thus further specify the gauge by demanding that $B_0(x)$ fluctuates around $B_0(x) \equiv 0$. We then set

$$A_i = B_i + Q_i, \quad \phi = \Omega + \chi.$$

Neglecting all non-zero modes, we obtain the reduced action at leading order:

$$\begin{aligned} \mathcal{S}_T(B_\mu, \Omega) = & \frac{1}{T} \int d^{d-1}x \left[\frac{1}{2e^2} \sum_i (\partial_i B_0(x))^2 + \frac{1}{4e^2} \sum_{i,j} B_{ij}^2(x) \right. \\ & \left. + \sum_i |\nabla_i \Omega(x)|^2 + |\Omega(x)|^2 B_0^2(x) + U(|\Omega(x)|^2) \right], \end{aligned} \quad (33.121)$$

where covariant derivative and curvature now refer to the gauge field B_i :

$$D_i = \partial_i + iB_i, \quad \text{and} \quad B_{ij} = [D_i, D_j].$$

If Ω has a non-zero expectation value, we obtain one massive vector field degenerated in mass with a scalar field, and the Higgs field. We expect the degeneracy between vector and scalar masses to be lifted by the integration over non-zero modes.

At one-loop order, we need the terms quadratic in $\{Q_\mu, \chi\}$. In the Gaussian integration over Q_μ, χ , at high temperature, the leading effects come from shifts of masses. It is thus sufficient to calculate with B_0 and Ω constant.

For the contribution to the B_0 mass, the relevant quadratic action is

$$\mathcal{S}_2(B_0) = \int dt d^{d-1}x |\nabla_\mu \chi(t, x)|^2,$$

where we have omitted the term proportional to $r|\chi|^2$, a high-temperature approximation. In the limit of constant B_μ , the space components B_i can be eliminated by a gauge transformation. The remaining B_0 component cannot be eliminated because χ satisfies periodic boundary conditions in the time direction.

Instead, the mode integration generates a potential for B_0 :

$$\int d^{d-1}x \sum_{n \neq 0} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln \left[k^2 + (2\pi nT + B_0(x))^2 \right] = \int d^{d-1}x \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \\ \times \left\{ \ln [\cosh(k/T) - \cos(B_0(x)/T)] - \ln(k^2 + B_0^2(x)) \right\}.$$

The zero-mode subtraction is a one-loop counter-term, which can be omitted. The contribution $\delta m_{B_0}^2$ to the coefficient of $\frac{1}{2}B_0^2$ then is

$$\delta m_{B_0}^2 = 2N_{d-1} T^{d-3} \Gamma(d-1) \zeta(d-2) \int d^{d-1}x B_0^2(x).$$

In particular, for $d=4$, one finds

$$\delta m_{B_0}^2 = \frac{1}{3} e^2 T^2.$$

The quadratic terms in the action relevant for the Ω mass shift are ($Q_{ij} \equiv \partial_i Q_j - \partial_j Q_i$)

$$\mathcal{S}_2(\Omega) = \int dt d^{d-1}x \left[\frac{1}{2e^2} \sum_i (\partial_t Q_i)^2 + \frac{1}{4e^2} \sum_{i,j} Q_{ij}^2 + |\partial_t \chi|^2 \right. \\ \left. + \sum_i |\partial_i \chi + iQ_i \Omega|^2 + \frac{2}{3} g |\Omega|^2 |\chi|^2 \right].$$

The integration over χ then yields two contributions, one proportional to $g|\Omega|^2$, and another one that adds to the Q action,

$$|\Omega|^2 \sum_{n \neq 0} \int d^{d-1}k \sum_{i,j} Q_i(\omega_n, k) \left(\delta_{ij} - \frac{k_i k_j}{k^2 + \omega_n^2} \right) Q_j(-\omega_n, -k), \quad \text{with } \omega_n = 2\pi nT.$$

Finally, the integration over Q yields a contribution to the coefficient of $|\Omega|^2$ proportional to e^2 . The total contribution is $\delta r = \bar{G}_2(0, T)((d-1)e^2 + 2g/3)$, where \bar{G}_2 is defined by equation (33.22). For $d=4$, the finite part of \bar{G}_2 is $T^2/12$ (Section 33.3.2) and, therefore,

$$\delta r = \frac{1}{36} T^2 (9e^2 + 2g).$$

Remarks.

(i) As we have discussed several times, in four dimensions additional UV contributions transform the parameters e^2, g, r into the one-loop expansion of the running parameters at scale T/Λ .

(ii) For completeness, we point out that the coefficient of $|\Omega|^2 B_0^2$ is renormalized. At one-loop order, one finds

$$1 \mapsto 1 + \frac{e^2 + g}{12\pi^2}.$$

(iii) We note that the coefficient of $|\Omega|^2$ increases with the temperature. If we assume that, at zero temperature, the $U(1)$ symmetry is broken, which implies that the coefficient r in the potential U (equation (33.119)) is sufficiently negative, eventually a critical temperature is reached where the $U(1)$ symmetry is restored. Near the transition, the scalar field B_0 remains massive and, therefore, the effective theory relevant for the phase transition is simply the $U(1)$ Higgs model in $(d-1)$ dimensions,

$$\tilde{\mathcal{S}}_T(B_i, \Omega) = \frac{1}{T} \int d^{d-1}x \left[\frac{1}{4\tilde{e}^2} \sum_{i,j} B_{ij}^2(x) + \sum_i |\partial_i \Omega(x)|^2 + \tilde{U}(|\Omega(x)|^2) \right], \quad (33.122)$$

where the parameters \tilde{e}^2 and in \tilde{U} can be obtained by integrating the reduced action also over the heavy field B_0 .

33.8 Non-Abelian gauge theories

Non-Abelian gauge theories (Chapter 22) with a limited number of fermions like QCD, are UV asymptotically free in four dimensions (Section 23.5.1). From the RG point of view, we expect some similarities with the non-linear σ -model in two dimensions. In particular, the effective coupling constant $g(T)$ decreases at high temperature, $g(T) \propto 1/\ln(m/T)$, where m is the RG-invariant mass scale of the gauge theory.

QCD and Standard Model calculations are reported in Refs. [349, 350].

Notation. We represent gauge fields A_μ^α by anti-Hermitian matrices:

$$\mathbf{A}_\mu = i\tau^\alpha A_\mu^\alpha,$$

where the Hermitian matrices τ^α are the generators of the Lie algebra of a unitary representation $\mathcal{R}(G)$ of a compact group G ,

$$\text{tr } \tau^\alpha \tau^\beta = \delta_{\alpha\beta}, \quad [\tau^\alpha, \tau^\beta] = if_{\alpha\beta\gamma}\tau^\gamma, \quad (33.123)$$

and the structure constants $f_{\alpha\beta\gamma}$ are chosen totally antisymmetric.

Gauge transformations take the form (equation (22.6))

$$\mathbf{A}'_\mu(t, x) = \mathbf{g}(t, x)\mathbf{A}_\mu(t, x)\mathbf{g}^{-1}(t, x) + \mathbf{g}(t, x)\partial_\mu\mathbf{g}^{-1}(t, x), \quad (33.124)$$

where \mathbf{g} is a group element in the matrix representation.

Covariant derivatives \mathbf{D}_μ act on fields φ transforming under the adjoint representation as (equation (22.13))

$$\mathbf{D}_\mu\varphi = \partial_\mu\varphi + [\mathbf{A}_\mu, \varphi]. \quad (33.125)$$

The corresponding curvature tensor is (equation (22.16))

$$\mathbf{F}_{\mu\nu}(t, x) = [\mathbf{D}_\mu, \mathbf{D}_\nu] = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu]. \quad (33.126)$$

For simplicity, in what follows we discuss only the pure gauge action, which reads

$$\mathcal{S}(\mathbf{A}) = -\frac{1}{4g^2} \text{tr} \int dt d^{d-1}x \mathbf{F}_{\mu\nu}^2(t, x). \quad (33.127)$$

33.8.1 Quantization and mode expansion

A new and important complication occurs with respect to the Abelian case: the mode decomposition is not gauge invariant. Thus, we quantize first, choosing the temporal gauge. In this gauge, the space components \mathbf{A}_i again are periodic up to a gauge transformation that enforces Gauss's law:

$$\mathbf{A}_i(1/T, x) = \mathbf{g}(x)\mathbf{A}_i(0, x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_i\mathbf{g}^{-1}(x).$$

We parametrize the group element \mathbf{g} in terms of an element φ of the Lie algebra,

$$\mathbf{g}(x) = e^{\varphi(x)/T},$$

and introduce

$$\mathbf{g}(t, x) = e^{t\varphi(x)}.$$

After the gauge transformation $\mathbf{A} \mapsto \mathbf{A}'$:

$$\mathbf{A}_i(t, x) = \mathbf{g}(t, x)\mathbf{A}'_i(t, x)\mathbf{g}^{-1}(t, x) + \mathbf{g}(t, x)\partial_i\mathbf{g}^{-1}(t, x),$$

the new field \mathbf{A}'_i is periodic. Since the gauge component \mathbf{A}_0 vanishes, the component \mathbf{A}'_0 reduces to

$$\mathbf{A}'_0(t, x) = \mathbf{g}^{-1}(t, x)\partial_t\mathbf{g}(t, x) = \varphi(x).$$

Again the temporal gauge reduces the time-component \mathbf{A}'_0 to its zero mode, the field φ .

In terms of the new fields, the gauge action (33.127) then reads (omitting now the primes)

$$\begin{aligned} \mathcal{S}(\mathbf{A}, \varphi) &= -\frac{1}{2g^2} \text{tr} \int dt d^{d-1}x \left[\sum_i (\partial_i \varphi(x) - \partial_t \mathbf{A}_i + [\mathbf{A}_i, \varphi])^2 + \frac{1}{2} \sum_{i,j} \mathbf{F}_{ij}^2 \right] \\ &= -\frac{1}{2g^2} \text{tr} \int dt d^{d-1}x \left[\sum_i (\mathbf{D}_i \varphi - \partial_t \mathbf{A}_i)^2 + \frac{1}{2} \sum_{i,j} \mathbf{F}_{ij}^2 \right]. \end{aligned}$$

Mode expansion. We expand the gauge field $\mathbf{A}_i(t, x)$ in a Fourier series in the Euclidean time variable, separating the zero modes,

$$\mathbf{A}_i(t, x) = \mathbf{B}_i(x) + \mathbf{Q}_i(t, x),$$

with

$$\mathbf{Q}_i(t, x) = \sum_{n \neq 0} e^{2i\pi n T t} \mathbf{Q}_{n,i}(x).$$

Then,

$$\mathbf{D}_i(\mathbf{A})\varphi = \mathbf{D}_i(\mathbf{B})\varphi + [\mathbf{Q}_i, \varphi].$$

In the same way,

$$\mathbf{F}_{ij}(\mathbf{A}) = \mathbf{F}_{ij}(\mathbf{B}) + \mathbf{D}_i \mathbf{Q}_j - \mathbf{D}_j \mathbf{Q}_i + [\mathbf{Q}_i, \mathbf{Q}_j],$$

where the covariant derivative now refers to the gauge field \mathbf{B} .

The resulting action is gauge invariant with respect to time-independent gauge transformations with gauge field \mathbf{B} . The gauge field is coupled to one massless scalar φ and massive vector fields with masses $4\pi^2 n^2 T^2$, all transforming under the adjoint representation.

The problem of quantization then reduces to the quantization of the field \mathbf{B} , for which we can choose a covariant gauge. One problem remains: massive vector fields lead to non-renormalizable theories. One way to solve the problem is to go over to a covariant gauge for the field \mathbf{A}_μ . We introduce a time component \mathbf{A}_0 for the gauge field (periodic in time) and multiply the functional measure by the corresponding δ -function. The action is a function only of the sum $\mathbf{A}_0(t, x) + \varphi(x)$. We thus temporarily call $\tilde{\mathbf{A}}_\mu$ the field

$$\tilde{\mathbf{A}}_i = \mathbf{A}_i, \quad \tilde{\mathbf{A}}_0 = \mathbf{A}_0(t, x) + \varphi(x).$$

The δ -function becomes $\delta(\tilde{\mathbf{A}}_0 - \varphi)$. The algebraic manipulations to pass to the covariant gauge with gauge function $\partial_\mu \tilde{\mathbf{A}}_\mu$ then are standard (see Section 22.4). They involve the gauge average of the constraint $\delta(\tilde{\mathbf{A}}_0 - \varphi)$. The result is a determinant which depends only on φ , while φ appears nowhere else. The integral over φ factorizes and gives a constant factor. Finally, ghost fields are required that satisfy periodic boundary conditions, unlike physical fermions.

33.8.2 Dimensional reduction

Neglecting all non-zero modes, we obtain the reduced action at leading order

$$\mathcal{S}_T(\mathbf{B}, \boldsymbol{\varphi}) = -\frac{1}{2Tg^2} \text{tr} \int d^{d-1}x \left[\sum_i (\mathbf{D}_i(\mathbf{B})\boldsymbol{\varphi}(x))^2 + \frac{1}{2} \sum_{i,j} \mathbf{B}_{ij}^2(x) \right],$$

that is, the action of a massless scalar field coupled to a gauge field.

One-loop calculation of the effective $\boldsymbol{\varphi}$ potential. Corrections generated by the integration over non-zero modes give a mass to the scalar field $\boldsymbol{\varphi}$, as in the Abelian example. We verify here this property at one-loop order. Since we need only the terms without derivatives we can set $\boldsymbol{\varphi}$ to a constant. Then, omitting the massless gauge field, we find a simplified action.

We then expand the fields, using the generators (33.123),

$$\mathbf{Q}_i = iQ_i^\alpha \tau^\alpha, \quad \boldsymbol{\varphi} = i\varphi^\alpha \tau^\alpha.$$

The relevant Q action can be written as

$$\begin{aligned} \mathcal{S}_2(Q) = & \frac{1}{2g^2} \int dt d^{d-1}x \left\{ \frac{1}{2} \sum_{i,j} [\partial_i Q_j^\alpha(t, x) - \partial_j Q_i^\alpha(t, x)]^2 \right. \\ & \left. + \sum_i [\partial_t Q_i^\alpha(t, x) + f_{\alpha\beta\gamma} Q_i^\beta(t, x) \varphi^\gamma(t, x)]^2 \right\}. \end{aligned}$$

The integration yields a determinant which generates an additive contribution to the effective action. Introducing the antisymmetric matrix Φ with elements (the element $\boldsymbol{\varphi}$ of the Lie algebra in the adjoint representation),

$$\Phi^{\alpha\beta} = f_{\alpha\beta\gamma} \varphi^\gamma,$$

we can write it as

$$\mathcal{S}_T^{(1)}(\boldsymbol{\varphi}) = \frac{1}{2} \sum_{n \neq 0} \text{tr} \ln \left[(k^2 \delta_{ij} - k_i k_j) \mathbf{1} + \delta_{ij} (\omega_n \mathbf{1} + i\Phi)^2 \right] - (\boldsymbol{\varphi} = 0), \quad (33.128)$$

with $\omega_n = 2\pi n T$. The result is the sum of two contributions, along k and transverse to k . The longitudinal contribution $\Sigma(\boldsymbol{\varphi})$ is

$$\Sigma(\boldsymbol{\varphi}) = \sum_x \sum_{n \neq 0} \text{tr} \ln (\mathbf{1} + i\Phi/\omega_n) = \ln \prod_x \det [2T\Phi^{-1} \sinh(\Phi/2T)]. \quad (33.129)$$

This term contributes to the $\boldsymbol{\varphi}$ integration measure and yields a factor at each point x :

$$\prod_x d\boldsymbol{\varphi}(x) \det \frac{\Phi(x)/2T}{\sinh(\Phi(x)/2T)},$$

which cancels the invariant group measure in the $\boldsymbol{\varphi}$ parametrization (see Section A33.2).

The transverse contribution to the action density is

$$\mathcal{E}(\boldsymbol{\varphi}) = \frac{1}{2}(d-2)T \sum_{n \neq 0} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln \det [k^2 \mathbf{1} + (\omega_n \mathbf{1} + i\Phi)^2] - (\boldsymbol{\varphi} = 0).$$

In terms of the imaginary eigenvalues of the antisymmetric matrix Φ , one obtains a sum of terms of the form (33.117), up to a shift in ω_n . The end of the calculation is thus similar to the Abelian case, and we complete it for the $SU(2)$ group. Then, setting $\Omega = |\varphi|$, we obtain

$$\begin{aligned}\mathcal{E}(\Omega) &= (d-2)T \sum_{n \neq 0} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \ln [k^2 + (\omega_n + \Omega)^2] - (\Omega = 0), \\ &= (d-2)T \int \frac{d^{d-1}k}{(2\pi)^{d-1}} [\ln(\cosh(k/T) - \cos(\Omega/T)) - \ln(k^2 + \Omega^2)] - (\Omega = 0).\end{aligned}$$

Thus,

$$\mathcal{E}'(\Omega) = (d-2) \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \left[\frac{\sin(\Omega/T)}{\cosh(k/T) - \cos(\Omega/T)} - \frac{2T\Omega}{k^2 + \Omega^2} \right].$$

The subtracted zero-mode contribution acts as a one-loop counter-term, and can be omitted. Then, as in the QED example, $\Omega = 0$ is the minimum of the potential (which is also periodic in Ω), and the generated mass term is UV finite:

$$m_T^2 = 2(d-2) \frac{\Gamma(d/2)}{\pi^{d/2}} \zeta(d-2) g^2 T^{d-2},$$

because gauge invariance ensures the absence at $T = 0$ of mass terms for gauge fields.

The mass m_T is thus proportional to $gT^{(d-2)/2}$. If the coupling constant is generic for $d > 4$ again the ratio M_T/T is small. For $d = 4$, the situation is different from the QED case, because the theory is UV asymptotically free. We expect a situation similar to the non-linear σ -model in two dimensions: the effective coupling constant at high temperature is logarithmically small, $g^2(T) \propto 1/\ln(T/m)$, m being the RG-invariant mass scale of the gauge theory (related to the β -function). Thus, we can trust the effective reduced field theory. In the same way, at leading order, the mass m_T is of order $Tg(T)$. In a sense, we find a situation similar to QED, but for different reasons. The effective theory, after mass summation, can be expanded in perturbation theory, but the expansion parameter is not very small, and the expansion may not be useful.

Finally, for $d < 4$, we expect m_T to be large, which would mean that the scalar field can be integrated out.

Remarks. Detailed calculations have been performed for models of physical interest like QCD with the problem of the quark-gluon plasma phase and the Higgs sector of the Standard Model (SM) with the question of the $SU(2) \times U(1)$ symmetry restoration. In QCD, a problem of slow convergence arises and various summation schemes have been proposed [351].

A33 Feynman diagrams at finite temperature

In the appendix, we summarize a few definitions and identities useful for general one-loop calculations. A short section is devoted to a few reminders on group measures useful for gauge theories.

A33.1 One-loop calculations

We describe here a few technical details about explicit calculations of one-loop Feynman diagrams.

A33.1.1 General remarks

Feynman diagrams can be calculated by extending the general methods explained in the framework of finite-size scaling, which involve Jacobi's elliptic functions. However, more specific techniques are also available in finite temperature QFT. The idea is the following: in the mixed $(d - 1)$ -momentum, time representation the propagator is the two-point function $\tilde{\Delta}(t, p)$ of the harmonic oscillator with frequency $\omega(p) = \sqrt{p^2 + m^2}$ and time interval $L = 1/T$:

$$\frac{1}{p^2 + \omega^2 + m^2} \mapsto \tilde{\Delta}(t, p) = \frac{1}{2\omega(p)} \frac{\cosh[(1/2T - |t|)\omega(p)]}{\sinh(\omega(p)/2T)}.$$

Summing over all frequencies is equivalent to setting $t = 0$. For the simple one-loop diagram, one finds

$$\frac{1}{2\pi} \int \frac{d\omega}{\omega^2 + p^2 + m^2} \mapsto \frac{1}{2\omega(p)} \frac{\cosh(\omega(p)/2T)}{\sinh(\omega(p)/2T)}.$$

The expression can be written in a way that separates quantum and thermal contributions:

$$\frac{1}{2\omega(p) \tanh(\omega(p)/2T)} = \frac{1}{2\omega(p)} + \frac{1}{\omega(p)(e^{\omega(p)/T} - 1)},$$

where the first term is the zero-temperature result, and the second term, which involves the relativistic Bose statistical factor, decreases exponentially at large momentum.

Finally, in the example of fermions or gauge theories, we can use a more general identity that can be proven by replacing the sum by a contour integral,

$$\sum_n \frac{x}{(n + \nu)^2 + x^2} = \pi \frac{\sinh(2\pi x)}{\cosh(2\pi x) - \cos(2\pi\nu)}, \quad (A33.1)$$

and thus after integration,

$$\begin{aligned} & \ln(\cosh(2\pi x) - \cos(2\pi\nu)) - \ln 2 \\ &= \lim_{N \rightarrow \infty} \sum_{n=-N}^N [\ln((n + \nu)^2 + x^2) - 2\ln(n + 1/2)]. \end{aligned} \quad (A33.2)$$

A33.1.2 Γ , ψ , ζ , θ -functions: A few useful identities

We remind here a few identities useful in calculations about Γ , ψ , ζ -functions. The Γ function satisfies

$$\sqrt{\pi} \Gamma(2z) = 2^{2z-1} \Gamma(z) \Gamma(z + 1/2), \quad \Gamma(z) \Gamma(1-z) \sin(\pi z) = \pi. \quad (\text{A33.3})$$

These identities imply for the $\psi(z)$ function, $\psi(z) = \Gamma'(z)/\Gamma(z)$,

$$2\psi(2z) = 2 \ln 2 + \psi(z) + \psi(z + 1/2), \quad \psi(z) - \psi(1-z) + \pi/\tan(\pi z) = 0. \quad (\text{A33.4})$$

We also need Riemann's ζ -function

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}. \quad (\text{A33.5})$$

It satisfies the reflection formula

$$\zeta(s) \Gamma(s/2) = \pi^{s-1/2} \Gamma((1-s)/2) \zeta(1-s), \quad (\text{A33.6})$$

which can be written in different forms using Γ -function relations. Moreover,

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^s} = (2^{1-s} - 1) \zeta(s). \quad (\text{A33.7})$$

Finally ($\gamma = -\psi(1)$),

$$\zeta(1+\varepsilon) = 1/\varepsilon + \gamma + O(\varepsilon), \quad (\text{A33.8})$$

$$\zeta(\varepsilon) = -\frac{1}{2}(2\pi)^\varepsilon + O(\varepsilon^2). \quad (\text{A33.9})$$

Jacobi's θ -functions. In Chapter 32, we have already defined the function $\vartheta_0(s)$ (equation (32.44)), related to Jacobi's elliptic functions (see the following examples):

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} e^{-\pi s n^2},$$

and proven in Section A32.1, with the help of Poisson's summation formula, the useful relation

$$\vartheta_0(s) = (1/s)^{1/2} \vartheta_0(1/s).$$

When calculations involve fermions or gauge fields, a more general function is needed:

$$\vartheta_2(s; \nu, \lambda) = e^{-\pi s \nu^2} \theta_3(\lambda + i\nu s, e^{-\pi s}) = \sum_n e^{-\pi s(n+\nu)^2 + 2i\pi n \lambda}, \quad (\text{A33.10})$$

where θ_3 is an elliptic Jacobi's function. Applying Poisson's formula, ϑ_2 can be shown to satisfy

$$\vartheta_2(s; \nu, \lambda) = \vartheta_2(s; -\nu, -\lambda) = s^{-1/2} \vartheta_2(1/s; \lambda, -\nu). \quad (\text{A33.11})$$

For fermions, the relevant function is

$$\vartheta_1(s) \equiv \vartheta_2(s; 1/2, 0) = \sum_n e^{-(n+1/2)^2 \pi s}. \quad (\text{A33.12})$$

From (A33.11), we obtain

$$\vartheta_1(s) = \frac{1}{\sqrt{s}} \sum_n (-1)^n e^{-\pi n^2/s}. \quad (\text{A33.13})$$

A33.2 Group measure

For the discussion of non-Abelian gauge theories, we derive the form of the group measure in the representation of group elements as exponentials of elements of the Lie algebra. The notation and conventions are the same as in Section 33.8.

We set

$$\mathbf{g} = e^{\xi},$$

and we want to determine the invariant measure in terms of the components ξ^α :

$$\xi = i\tau^\alpha \xi^\alpha.$$

We thus introduce a time-dependent group element \mathbf{g} as

$$\mathbf{g}(t) = e^{t\xi}, \quad \mathbf{g}(1) = \mathbf{g}.$$

We also need the element of the Lie algebra

$$\mathbf{L}^\alpha(t) = \frac{\partial \mathbf{g}(t)}{\partial \xi^\alpha} \mathbf{g}^{-1}(t).$$

It satisfies the differential equation

$$\frac{d}{dt} \mathbf{L}^\alpha = i\tau^\alpha + [\xi, \mathbf{L}^\alpha], \quad \mathbf{L}^\alpha(0) = 0. \quad (A33.14)$$

This equation can also be expressed in component form, setting

$$\mathbf{L}^\alpha = iL^{\alpha\beta}\tau^\beta.$$

Then,

$$\frac{d}{dt} L^{\alpha\beta} = \delta_{\alpha\beta} + f_{\gamma\beta\delta}\xi^\gamma L^{\alpha\delta}. \quad (A33.15)$$

We call Λ the matrix of elements $L^{\alpha\beta}$ and introduce the antisymmetric matrix X of elements

$$X^{\alpha\beta} = f_{\alpha\beta\gamma}\xi^\gamma.$$

The solution of equation (A33.15) can then be written as

$$\Lambda(t) = \int_0^t dt' e^{X(t'-t)} = X^{-1} (1 - e^{-tX}).$$

The metric tensor corresponding to the group is

$$g_{\alpha\beta} = -\text{tr } \mathbf{L}^\alpha(1)\mathbf{L}^\beta(1) = L^{\alpha\gamma}(1)L^{\beta\gamma}(1) = (\Lambda\Lambda^T)_{\alpha\beta}, \quad (A33.16)$$

and the group invariant measure is

$$d\mathbf{g} \equiv (\det g_{\alpha\beta})^{1/2} \prod_\alpha d\xi^\alpha = (\det \Lambda\Lambda^T)^{1/2} \prod_\alpha d\xi^\alpha. \quad (A33.17)$$

Then,

$$\begin{aligned} \Lambda\Lambda^T &= -X^{-2} (1 - e^{-X}) (1 - e^X) = -4X^{-2} \sinh^2(X/2) \\ &= -\prod_{n \neq 0} (1 + X^2/(2\pi n)^2), \end{aligned}$$

where we recognize an expression that appears in equation (33.129).

34 Stochastic differential equations: Langevin, Fokker–Planck (FP) equations

This chapter is devoted to the study of Langevin equations [352], first-order in time differential equations, which depend on a random noise, and which belong to a class of stochastic differential equations that describe diffusion processes, or random motion. From a Langevin equation, we derive a Fokker–Planck equation [353] for the probability distribution of the solutions, at given time of the Langevin equation. The FP equation has a form analogous to the equation for the statistical operator in a magnetic field, which we have studied in Section 3.3 (but, in general, the corresponding quantum FP Hamiltonian is not Hermitian).

In Chapters 2 and 3, we have noted, when the quantum Hamiltonian \mathbf{H} is positive and, as functions of momenta, even and quadratic, the path integral representation of the matrix elements of the quantum statistical operator $e^{-\beta \mathbf{H}}$ involves an integrand that defines a positive measure and, therefore, has a probabilistic interpretation. With analogous methods, we show here that observables averaged over the noise can also be calculated from path integrals [354], whose integrands define automatically positive measures.

The path integrals involve dynamic actions that have automatically a Becchi–Rouet–Stora–Tyutin (BRST) symmetry [285] and, when the driving force derives from a potential exhibit the simplest form of supersymmetry [295].

In some cases, like Brownian motion on Riemannian manifolds, difficulties appear in the precise definition of stochastic equations, quite similar to the quantization problem encountered in quantum mechanics (QM). Time discretization provides one possible solution to the problem.

This chapter is meant to serve as an introduction to Chapters 35 and 36, in which *stochastic field equations* and *critical dynamics* are discussed.

34.1 The Langevin equation

We call *Langevin equation* a first order stochastic differential equation of the form

$$\dot{\mathbf{q}}(t) = -\frac{1}{2}\mathbf{f}(\mathbf{q}(t), t) + \boldsymbol{\nu}(t), \quad (34.1)$$

($\dot{\mathbf{q}}(t) \equiv d\mathbf{q}(t)/dt$) in which t is the time, $\mathbf{q}(t)$ is a trajectory in \mathbb{R}^d , $\mathbf{f}(\mathbf{q}, t)$ is the drift force, assumed to be a smooth function of \mathbf{q} ($\mathbf{q} \equiv (q_1, \dots, q_d)$), and $\boldsymbol{\nu}(t)$ a stochastic vector function, called hereafter the *noise*.

The noise is characterized by a functional probability measure $[d\rho(\boldsymbol{\nu})]$. In what follows, we specialize to a *Gaussian white noise*, which corresponds to a measure normalized to 1 when restricted to any finite time interval (a property useful when the measure is factorized into a product of measures corresponding to successive time intervals), of the form

$$[d\rho(\boldsymbol{\nu})] = [d\nu] \exp\left[-\frac{1}{2\Omega} \int dt \boldsymbol{\nu}^2(t)\right], \quad \text{with } \int [d\rho(\boldsymbol{\nu})] = 1. \quad (34.2)$$

The positive constant Ω characterizes the width of the noise distribution.

It follows from the results of Sections 2.5 and 2.6, that the Gaussian noise can, equivalently be characterized by its one- and two-point correlation functions. Using the bra–ket notation $\langle \bullet \rangle_\nu$ for averages over the noise (34.2), one finds

$$\langle \nu_i(t) \rangle_\nu = 0, \quad \langle \nu_i(t) \nu_j(t') \rangle_\nu = \Omega \delta_{ij} \delta(t - t'), \quad \text{for } 1 \leq i, j \leq d, \quad (34.3)$$

where $\delta(t)$ is Dirac function. Although already fairly general, equation (34.1) is not the most general stochastic first-order differential equation. In particular, one can define the equation in bounded domains, or in different manifolds. A few more general equations, for which additional definition problems arise are discussed in Section 34.9.

Langevin equation and Markovian process. Because the values of the noise at different times are uncorrelated, and the Langevin equation is local in time, that is, the variation of $\mathbf{q}(t)$ only depends on its value at time t , but not on the values at preceding times, the Langevin defines a *Markov process*, and even a stationary Markov process when \mathbf{f} is not an explicit function of time.

34.1.1 The discretized Langevin equation

To deal with the formal problems associated with averages of the form $\langle \dot{q}(t) q(t) \rangle$ (see equation (34.52)), we integrate the Langevin equation (34.1) in an infinitesimal time interval between time t and $t + \varepsilon$:

$$q_i(t + \varepsilon) = q_i(t) - \frac{1}{2}\varepsilon f_i(\mathbf{q}(t), t) + \int_t^{t+\varepsilon} \nu_i(\tau) d\tau + O(\varepsilon^{3/2}), \quad (34.4)$$

where $O(\varepsilon^{3/2})$ means that the remainder is a stochastic quantity of order $\varepsilon^{3/2}$ on average, an estimate that is motivated by the estimate (34.45).

At order ε , the continuum Langevin equation is equivalent to the discretized Langevin equation (for details, see also Section 34.9),

$$q_i(t + \varepsilon) = q_i(t) - \frac{1}{2}\varepsilon f_i(\mathbf{q}(t), t) + \sqrt{\varepsilon} \bar{\nu}_i(t), \quad (34.5)$$

where the noise $\bar{\nu}(t)$ is defined by

$$\int_t^{t+\varepsilon} \nu(\tau) d\tau = \sqrt{\varepsilon} \bar{\nu}(t). \quad (34.6)$$

From the definition (34.2), it follows that $\bar{\nu}(t)$ also has a Gaussian distribution. Restricting time to discrete values, for instance, by $t/\varepsilon \in \mathbb{Z}$, then the distribution can be characterized by,

$$\langle \bar{\nu}_i(t) \rangle_{\bar{\nu}} = 0, \quad \langle \bar{\nu}_i(t) \bar{\nu}_j(t') \rangle_{\bar{\nu}} = \Omega \delta_{ij} \delta_{tt'}, \quad (34.7)$$

where $\delta_{tt'}$ is now a Kronecker delta.

Conversely, the Langevin equation (34.1) with noise (34.3) can be defined as the continuum limit, when $\varepsilon \rightarrow 0$, of the discretized Langevin equation (34.5), with noise (34.7).

34.2 Time-dependent probability distribution and FP equation

The solution of the Langevin equation (34.1) with the boundary condition $\mathbf{q}(t_0) = \mathbf{q}_0$ at initial time t_0 , is a stochastic quantity with a probability distribution $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ for the *value* \mathbf{q} of the stochastic vector $\mathbf{q}(t)$ at time t , which can be formally written as

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = \langle \delta^d(\mathbf{q}(t) - \mathbf{q}) \rangle_\nu, \quad t \geq t_0. \quad (34.8)$$

The vector \mathbf{q} is the argument of $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ and should not be confused with the function $\mathbf{q}(t)$ itself.

The definition (34.8) is equivalent to: for any continuous function $\mathcal{O}(\mathbf{q})$, for $t \geq t_0$,

$$\int P(\mathbf{q}, t; \mathbf{q}_0, t_0) \mathcal{O}(\mathbf{q}) d^d q = \langle \mathcal{O}(\mathbf{q}(t)) \rangle_\nu. \quad (34.9)$$

34.2.1 Markov property and FP Hamiltonian

Because the evolution generated the Langevin equation (34.1) and noise (34.3) is Markovian, the distribution P also satisfies a Markov property, analogous to the property (2.3) of the solution of equation (2.4):

$$P(\mathbf{q}_3, t_3; \mathbf{q}_1, t_1) = \int d^d q_2 P(\mathbf{q}_3, t_3; \mathbf{q}_2, t_2) P(\mathbf{q}_2, t_2; \mathbf{q}_1, t_1), \quad \text{for } t_1 \leq t_2 \leq t_3. \quad (34.10)$$

Therefore, the distribution P is completely determined from its knowledge for small time intervals. It is convenient to introduce the quantum bra–ket notation (not to be confused with the notation $\langle \bullet \rangle_\nu$, meaning average over the noise ν), and to express $P(\mathbf{q}, t; \mathbf{q}', t')$ as the expectation value of a quantum operator $\mathbf{P}(t, t')$ between two states $\langle \mathbf{q}|$ and $|\mathbf{q}'\rangle$:

$$P(\mathbf{q}, t; \mathbf{q}', t') \equiv \langle \mathbf{q} | \mathbf{P}(t, t') | \mathbf{q}' \rangle. \quad (34.11)$$

The relation (34.10) implies that the operator $\mathbf{P}(t, t')$ satisfies a differential evolution equation (Section 2.1) of the form ($\partial_t \equiv \partial/\partial t$)

$$\partial_t \mathbf{P}(t, t') = -\mathbf{H}(t) \mathbf{P}(t, t'), \quad (34.12)$$

or, as well,

$$\partial_{t'} \mathbf{P}(t, t') = \mathbf{P}(t, t') \mathbf{H}(t'), \quad (34.13)$$

where the operator $\mathbf{H}(t)$ is the FP Hamiltonian, analogous to a quantum Hamiltonian.

Time-translation invariance. When $\mathbf{f}(\mathbf{q}, t)$ is not an explicit function of time ($\mathbf{f}(\mathbf{q}, t) \equiv \mathbf{f}(\mathbf{q})$), the stochastic evolution is *time-translation invariant*. For simplicity, this is the situation we consider now in the chapter, except when stated explicitly otherwise. For a time-translation-invariant Langevin equation (34.1) and noise (34.3), $\mathbf{H}(t) \equiv \mathbf{H}$ is time independent, and the operator $\mathbf{P}(t, t_0)$ can be expressed as

$$\mathbf{P}(t, t_0) = e^{-(t-t_0)\mathbf{H}}. \quad (34.14)$$

34.2.2 The FP equation

Notation. In what follows, we sometimes omit the dependence on the initial conditions, when it is not directly useful, and write $P(\mathbf{q}, t)$ as a simplified notation for $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$.

We now show that equations (34.1, 34.2) imply a differential equation for $P(\mathbf{q}, t)$. For this purpose, we start from the discretized form (34.5) of the Langevin equation. We consider the probability distribution of $\mathbf{q} = \mathbf{q}(t + \varepsilon)$, with initial condition at time t , $\mathbf{q}' = \mathbf{q}(t)$, which is

$$P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) = \left\langle \delta^{(d)}(\mathbf{q} - \mathbf{q}(t + \varepsilon)) \right\rangle_{\bar{\nu}}.$$

The Fourier transform \tilde{P} of P with respect to \mathbf{q} is given by

$$\begin{aligned} \tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) &= \int d^d q e^{-i\mathbf{p}\cdot\mathbf{q}} P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) = \left\langle e^{-i\mathbf{p}\cdot\mathbf{q}(t+\varepsilon)} \right\rangle_{\bar{\nu}}, \\ &= \exp[-i\mathbf{p} \cdot (\mathbf{q}' - \varepsilon \mathbf{f}(\mathbf{q}')/2)] \left\langle \exp[-i\sqrt{\varepsilon} \mathbf{p} \cdot \bar{\nu}(t)] \right\rangle_{\bar{\nu}}. \end{aligned}$$

The integral over $\bar{\nu}(\tau)$ is Gaussian, and can be performed explicitly. It yields

$$\tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) = e^{-\varepsilon[\Omega\mathbf{p}^2 - i\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')]/2} e^{-i\mathbf{p}\cdot\mathbf{q}'} . \quad (34.15)$$

According to the discussion of Section 34.2.1, the coefficient of ε , in the expansion of the expression for $\varepsilon \rightarrow 0$, is the matrix element of the FP Hamiltonian (equation (34.14)) in the mixed \mathbf{p}, \mathbf{q} representation,

$$\begin{aligned} \tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) &= e^{-i\mathbf{p}\cdot\mathbf{q}'} [1 - \frac{1}{2}\varepsilon(\Omega\mathbf{p}^2 - i\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')) + O(\varepsilon^2)] + O(\varepsilon^{3/2}), \\ &= e^{-i\mathbf{p}\cdot\mathbf{q}'} [1 - \varepsilon H(\mathbf{p}, \mathbf{q}')] + O(\varepsilon^{3/2}). \end{aligned}$$

Inverting the Fourier transformation, one finds that \mathbf{H} is a second-order differential operator, with elements (in the bra-ket notation of QM),

$$\begin{aligned} \langle \mathbf{q} | \mathbf{H} | \mathbf{q}' \rangle &= \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} e^{i\mathbf{p}\cdot(\mathbf{q}-\mathbf{q}')} (\Omega\mathbf{p}^2 - i\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')), \\ &= -\frac{1}{2} \nabla_{\mathbf{q}} \cdot \left[(\Omega\nabla_{\mathbf{q}} + \mathbf{f}(\mathbf{q})) \delta^{(d)}(\mathbf{q} - \mathbf{q}') \right] \end{aligned} \quad (34.16)$$

($\nabla_{\mathbf{q}} \equiv \{\partial/\partial_{q_1} \cdots \partial/\partial_{q_d}\}$). It can also be expressed in terms of position $\hat{\mathbf{q}}$ and momentum $\hat{\mathbf{p}}$ operators of QM, with the commutation relation $[\hat{q}_i, \hat{p}_j] = i\delta_{ij}$, (equation (2.7)) as (see Section 2.2, with $\hbar = 1$)

$$\mathbf{H} = \frac{1}{2} (\Omega\hat{\mathbf{p}}^2 - i\hat{\mathbf{p}} \cdot \mathbf{f}(\hat{\mathbf{q}})) , \quad (34.17)$$

(see Section 3.3.1 for a problem due a similar product of non-commuting operators).

Equation (34.12) then implies that the probability distribution $P(\mathbf{q}, t)$ (equation (34.8)) at finite time $t > t_0$, satisfies the evolution equation,

$$\dot{P}(\mathbf{q}, t) = \frac{1}{2} \nabla_{\mathbf{q}} \cdot [\Omega \nabla_{\mathbf{q}} P(\mathbf{q}, t) + \mathbf{f}(\mathbf{q}) P(\mathbf{q}, t)] . \quad (34.18)$$

This is the FP equation associated with the Langevin equation (34.1) (in the time-translation-invariant case) and noise (34.2).

The factorization of the operator $\hat{\mathbf{p}} = -i\nabla_{\mathbf{q}}$ in the left of the FP Hamiltonian (34.17) and, thus, $\nabla_{\mathbf{q}}$ in the right of equation (34.18), is directly related to conservation of probability. Indeed, assuming that P vanishes fast enough at infinity, and integrating equation (34.18) over whole space, one obtains,

$$\frac{\partial}{\partial t} \int d^d q P(\mathbf{q}, t) = 0,$$

the condition of *probability conservation*.

The FP equation (34.18) is analogous to an imaginary-time Schrödinger equation. This establishes a formal correspondence between stochastic differential equations and Euclidean QM. All observables that can be calculated from the Langevin equation by averaging over the noise, can be recovered by methods of QN using the FP Hamiltonian \mathbf{H} .

Finally, after Fourier transformation, equation (34.15) yields

$$P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) \sim (2\pi\Omega\varepsilon)^{-d/2} \exp \left[-\frac{1}{2\varepsilon\Omega} (\mathbf{q} - \mathbf{q}' + \frac{1}{2}\varepsilon\mathbf{f}(\mathbf{q}'))^2 \right]. \quad (34.19)$$

This form of the probability distribution, valid for $\varepsilon \rightarrow 0$, will be used in Section 34.6 to derive a path integral representation for $P(\mathbf{q}, t; \mathbf{q}', t')$.

34.3 Equilibrium distribution. Correlation functions

The methods of QM can now be used to investigate several interesting questions, such as, for example, the existence of an equilibrium distribution. One has to be careful that the FP Hamiltonian is not Hermitian and, therefore, its left and right eigenvectors do not coincide.

Moreover, if one applies the FP equation to a system whose evolution is not ergodic (see Section 13.4, or Chapter 14), the configuration space is decomposed into disconnected components, and the arguments that follow apply to each connected component separately. This is a situation one encounters in the study of the ordered phase in the theory of phase transitions, in quantum field theory, or statistical physics.

34.3.1 Equilibrium distribution

The equilibrium distribution $P_0(\mathbf{q})$ is the large-time limit, if it exists, of $P(\mathbf{q}, t)$:

$$P_0(\mathbf{q}) = \lim_{t \rightarrow +\infty} P(\mathbf{q}, t). \quad (34.20)$$

This implies that $P_0(\mathbf{q})$ is then a time-independent solution of the FP equation (34.18), which must be positive and normalizable:

$$\int P_0(\mathbf{q}) d^d q < \infty, \quad (34.21)$$

and thus a right eigenvector of the FP Hamiltonian (34.17) corresponding to the eigenvalue 0. Introducing the bra–ket notation of QM, one can write

$$\mathbf{H} |\bar{0}\rangle = 0, \quad \text{with} \quad \langle \mathbf{q} | \bar{0}\rangle = P_0(\mathbf{q}).$$

Moreover, the derivation of probability conservation can be expressed by stating that a constant is a left eigenvector with eigenvalue 0 of the Hamiltonian \mathbf{H} and, thus, the dual of $P_0(\mathbf{q})$. One can write

$$\langle \bar{0} | \mathbf{H} = 0, \quad \text{with} \quad \langle \bar{0} | \mathbf{q} \rangle = 1. \quad (34.22)$$

The condition (34.21), which can then be expressed as

$$\int P_0(\mathbf{q}) d^d q = \langle \bar{0} | \bar{0} \rangle < \infty$$

is thus the condition that the vector $|\bar{0}\rangle$ has a finite norm, a necessary condition for the eigenvalue 0 to really belong to the spectrum. Since one can show that the eigenvalues of \mathbf{H} have a non-negative real part (for a hint of the proof, see Section A34.1), the eigenvalue corresponding to the vector $|\bar{0}\rangle$ has the lowest real part.

If no normalizable equilibrium distribution exists, all eigenvalues of \mathbf{H} have strictly positive real parts and the Langevin equation (34.1) has *runaway solutions*, in the sense that the probability of finding $\mathbf{q}(t)$ inside a ball of arbitrary but finite radius goes to 0 at large time:

- (i) algebraically, if the real part of the spectrum of \mathbf{H} has a continuous part extending up to the origin (this is the case with Brownian motion, for example),
- (ii) exponentially, with a rate which is the inverse of the real part of the eigenvalue with lowest real part, otherwise. This rate is called the *relaxation time* (for details, see Section A34.1).

Equilibrium distribution and Langevin equation. The sole knowledge of the equilibrium distribution, which we parametrize as

$$P_0(\mathbf{q}) = e^{-E(\mathbf{q})}, \quad (34.23)$$

does not determine the functions $\mathbf{f}(\mathbf{q})$ in equation (34.18) uniquely. Indeed, by demanding that $P_0(\mathbf{q})$ is a time-independent solution of the FP equation, one obtains only the condition ($\nabla \equiv \nabla_{\mathbf{q}}$)

$$\nabla \cdot [e^{-E(\mathbf{q})} (\Omega \nabla E(\mathbf{q}) - \mathbf{f}(\mathbf{q}))] = 0, \quad (34.24)$$

which, setting

$$\mathbf{V}(\mathbf{q}) = e^{-E(\mathbf{q})} (\Omega \nabla E(\mathbf{q}) - \mathbf{f}(\mathbf{q})), \quad (34.25)$$

can be rewritten as a current conservation equation $\nabla \cdot \mathbf{V}(\mathbf{q}) = 0$.

34.3.2 Correlation functions

We assume the existence of an equilibrium distribution and, for notational simplicity, we specialize to dimension $d = 1$.

So far, we have discussed only equal-time averages. However, the Langevin equation can also be used to calculate correlation functions of observables at different times, like

$$Z^{(n)}(t_0; t_1, t_2, \dots, t_n) = \langle q(t_1) q(t_2) \cdots q(t_n) \rangle_{\nu}, \quad (34.26)$$

where we first assume the boundary condition $q(t_0) = q_0$. We order times as $t_0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, and we then use the definition of the FP Hamiltonian to transform $Z^{(n)}$.

We first average over the noise corresponding to times $t > t_{n-1}$. We can then consider $q(t_{n-1})$ as the initial data and, thus,

$$Z^{(n)}(t_0, t_1, t_2, \dots, t_n) = \int dq_n q_n \langle P(q_n, t_n; q(t_{n-1}), t_{n-1}) q(t_1) q(t_2) \cdots q(t_{n-1}) \rangle_\nu .$$

We then average over the noise for times $t_{n-2} < t \leq t_{n-1}$, where now $q(t_{n-2})$ yields the boundary condition. We find

$$\begin{aligned} Z^{(n)}(t_1, t_2, \dots, t_n) &= \int dq_n dq_{n-1} q_n P(q_n, t_n; q_{n-1}, t_{n-1}) q_{n-1} \\ &\quad \times \langle P(q_{n-1}, t_{n-1}; q(t_{n-2}), t_{n-2}) q(t_1) q(t_2) \cdots q(t_{n-2}) \rangle_\nu . \end{aligned}$$

The procedure can be iterated until the averaging over the noise is completed:

$$\begin{aligned} Z^{(n)}(t_0; t_1, t_2, \dots, t_n) &= \int \left(\prod_{i=1}^n dq_i \right) q_n P(q_n, t_n; q_{n-1}, t_{n-1}) q_{n-1} \\ &\quad \times \cdots q_2 P(q_2, t_2; q_1, t_1) q_1 P(q_1, t_1; q_0, t_0). \end{aligned}$$

Finally, using the representation (34.14) and (34.11), and the notation and arguments of Section 2.5, we can express $Z^{(n)}(t_0; t_1, \dots, t_n)$ as an operator expectation value of the form (see also Section A2.3)

$$Z^{(n)}(t_0; t_1, t_2, \dots, t_n) = \langle \bar{0} | \hat{q} e^{-(t_n - t_{n-1})\mathbf{H}} \hat{q} \cdots \hat{q} e^{-(t_2 - t_1)\mathbf{H}} \hat{q} e^{-(t_1 - t_0)\mathbf{H}} | q_0 \rangle , \quad (34.27)$$

$$= \langle \bar{0} | \hat{Q}(t_n) \cdots \hat{Q}(t_2) \hat{Q}(t_1) e^{\mathbf{H} t_0} | q_0 \rangle , \quad (34.28)$$

(remember that $\hat{q}|q\rangle = q|\hat{q}\rangle$) where the vector $\langle \bar{0} |$ has been defined by equation (34.22) and implies $\langle \bar{0} | e^{t_n \mathbf{H}} = \langle \bar{0} |$. Moreover, to simplify the notation, we have introduced $\hat{Q}(t)$, the operator \hat{q} in the ‘Heisenberg’ representation:

$$\hat{Q}(t) = e^{\mathbf{H} t} \hat{q} e^{-\mathbf{H} t} .$$

Equilibrium correlation functions. When the boundary condition is set in the far past ($t_0 \rightarrow -\infty$), the correlation functions converge towards *equilibrium correlation functions*:

$$Z^{(n)}(t_1, t_2, \dots, t_n) = \langle \bar{0} | \hat{Q}(t_n) \cdots \hat{Q}(t_2) \hat{Q}(t_1) | \bar{0} \rangle , \quad (34.29)$$

which are time-translation invariant. We recognize the analogue of the representation (2.45) of correlation functions as vacuum or ground state expectation values of product of operators in QM, as discussed in Section 2.5. As shown in Section A2.3, these expressions can be symmetrized in time by introducing the time-ordering operation.

34.3.3 Time evolution of observables

Passing from the FP equation formulation to expressions in terms of time-dependent operators, we have changed from a Schrödinger picture to a Heisenberg picture. We recall that in the Heisenberg picture, one directly writes equations for the evolution of time-dependent operators.

For example, in terms of Heisenberg operators, the average of an observable $\mathcal{O}(q)$ at time t can be expressed as

$$\langle \mathcal{O}(q(t)) \rangle_\nu = \langle \bar{0} | \mathcal{O}[\hat{Q}(t)] | \bar{0} \rangle .$$

The operator $\mathcal{O}[\hat{Q}(t)]$ satisfies the evolution equation,

$$\frac{d}{dt}\mathcal{O}[\hat{Q}(t)] = [\mathbf{H}, \mathcal{O}[\hat{Q}(t)]]. \quad (34.30)$$

Using equation (34.30), $\langle \bar{0} | \mathbf{H} = 0$, and the representation (34.16)) of \mathbf{H} , one derives for the matrix elements

$$\mathcal{O}(q, t) \equiv \langle \bar{0} | \mathcal{O}[\hat{Q}(t)] | q \rangle,$$

the evolution equation

$$\dot{\mathcal{O}}(q, t) = \frac{1}{2} \left(\Omega \frac{\partial}{\partial q} - f(q) \right) \frac{\partial}{\partial q} \mathcal{O}(q, t). \quad (34.31)$$

Then, the average $\langle \mathcal{O}(q(t)) \rangle_\nu$ is simply obtained by integrating $\mathcal{O}(q, t)$ over q with the equilibrium distribution as measure,

$$\langle \mathcal{O}(q(t)) \rangle_\nu = \int dq \mathcal{O}(q, t) P_0(q). \quad (34.32)$$

34.4 A special class: Dissipative Langevin equations

When $\mathbf{f}(\mathbf{q})$ is a gradient, which we parametrize as (a non-trivial condition only for $d > 1$)

$$\mathbf{f}(\mathbf{q}) = \Omega \nabla E(\mathbf{q}), \quad (34.33)$$

the Langevin equation ($\partial_i \equiv \partial/\partial q_i$)

$$\dot{q}_i(t) = -\frac{1}{2} \Omega \partial_i E(\mathbf{q}(t)) + \nu_i(t), \quad (34.34)$$

is called *purely dissipative*. The driving force f_i is then called conservative. With the parametrization (34.33), Ω becomes an inverse time scale. The linear Langevin equation (34.43) provides the simplest example of such an equation, since $\omega q = (\partial/\partial q) (\frac{1}{2} \omega q^2)$.

In the limit of vanishing noise, the Langevin equation reduces to the gradient flow,

$$\dot{q}_i(t) = -\frac{1}{2} \Omega \partial_i E(\mathbf{q}(t)). \quad (34.35)$$

Taking the scalar product with the vector $\dot{\mathbf{q}}$ and integrating, one finds

$$\int_{t_0}^t \dot{\mathbf{q}}^2(t') dt' = -\frac{1}{2} \Omega [E(\mathbf{q}(t)) - E(\mathbf{q}(t_0))]. \quad (34.36)$$

Therefore, in this limit, $E(\mathbf{q}(t))$ is a monotonically decreasing function of time.

34.4.1 The FP Hamiltonian

The FP Hamiltonian (34.17) is then equivalent to a real symmetric positive Hamiltonian. Indeed, the transformation

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = e^{-E(\mathbf{q})/2} \langle \mathbf{q} | U(t, t_0) | \mathbf{q}_0 \rangle e^{E(\mathbf{q}_0)/2}, \quad (34.37)$$

introduced into equation (34.18), leads to

$$\partial_t \langle \mathbf{q} | U(t, t_0) | \mathbf{q}_0 \rangle = -\tilde{\mathbf{H}} \langle \mathbf{q} | U(t, t_0) | \mathbf{q}_0 \rangle, \quad (34.38)$$

where the transformed Hamiltonian is given by

$$\tilde{\mathbf{H}} = \frac{1}{2} \Omega \left[\hat{\mathbf{p}}^2 + \frac{1}{4} (\nabla E(\hat{\mathbf{q}}))^2 - \frac{1}{2} \nabla^2 E(\hat{\mathbf{q}}) \right] = \frac{1}{2} \Omega \sum_i A_i^\dagger A_i, \quad (34.39)$$

with the definition,

$$\mathbf{A}(q, \partial/\partial q) \equiv \nabla + \frac{1}{2} \nabla E(q) \equiv i \hat{\mathbf{p}} + \frac{1}{2} \nabla E(\hat{\mathbf{q}}), \quad (34.40)$$

and is an explicit positive operator.

The operator $U(t, t_0)$ then takes the form of a statistical operator: $U(t, t_0) = e^{-(t-t_0)\tilde{\mathbf{H}}}$. If the function $e^{-E(\mathbf{q})/2} \equiv \langle \mathbf{q} | \bar{0} \rangle$ is normalizable, that is,

$$\langle \bar{0} | \bar{0} \rangle = \int d\mathbf{q} e^{-E(\mathbf{q})} < \infty,$$

$|\bar{0}\rangle$ is the eigenvector and ground state corresponding to the eigenvalue 0:

$$\mathbf{A}|\bar{0}\rangle = 0 \quad \Rightarrow \quad \tilde{\mathbf{H}}|\bar{0}\rangle = 0,$$

and thus is the ground state of $\tilde{\mathbf{H}}$.

At large times, the operator $e^{-t\tilde{\mathbf{H}}}$ projects onto the ground state. The interpretation of this result in terms of the probability distribution $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ is then

$$\lim_{t \rightarrow \infty} P(\mathbf{q}, t; \mathbf{q}_0, t_0) = e^{-E(\mathbf{q})/2} e^{E(\mathbf{q}_0)/2} \langle \mathbf{q} | \bar{0} \rangle \langle \bar{0} | \mathbf{q}_0 \rangle = e^{-E(\mathbf{q})}.$$

For $t \rightarrow +\infty$, $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ converges towards the equilibrium distribution $e^{-E(\mathbf{q})}$. By contrast, if the wave function $e^{-E(\mathbf{q})/2}$ is not normalizable, $\tilde{\mathbf{H}}$ has only strictly positive eigenvalues: no equilibrium can be reached, and the Langevin equation (34.1) has only runaway solutions.

34.4.2 Detailed balance

Since the operator $U(t, t_0)$ is then also real symmetric, the purely dissipative Langevin equation (34.34) leads to *detailed balance*. Indeed,

$$\langle \mathbf{q} | U(t, t_0) | \mathbf{q}_0 \rangle = \langle \mathbf{q}_0 | U(t, t_0) | \mathbf{q} \rangle,$$

implies

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = e^{-E(\mathbf{q})+E(\mathbf{q}_0)} P(\mathbf{q}_0, t; \mathbf{q}, t_0), \quad (34.41)$$

which is the detailed balance condition (for discrete processes see Section A34.1.3).

When, in equation (34.34), the gradient term is explicitly time dependent:

$$\dot{\mathbf{q}}(t) = -\frac{1}{2}\Omega \nabla_q E(\mathbf{q}(t), t) + \boldsymbol{\nu}(t), \quad (34.42)$$

detailed balance is lost, but a more general relation, called Jarzynski's relation, survives (see Section 34.8).

34.5 The linear Langevin equation

In one dimension, the Langevin equation is always dissipative, and the simplest example is the linear Langevin equation

$$\dot{q}(t) = -\omega q(t) + \nu(t), \quad (34.43)$$

where ω is a constant, and $\nu(t)$ the Gaussian white noise defined by equation (34.3), where we set $\Omega = 1$ (we leave as an instructive exercise the two-dimensional example, ω being then a 2×2 real matrix). With the boundary condition $q(0) = q_0$, the solution of the Langevin equation is

$$q(t) = q_0 e^{-\omega t} + \int_0^t e^{-\omega(t-t')} \nu(t') dt'. \quad (34.44)$$

The random walk. The special case $\omega = 0$ describes the random walk, or Brownian motion, in the continuum. Taking the square of the relation,

$$q(t) - q(t') = \int_{t'}^t \nu(\tau) d\tau,$$

one infers, after averaging over the noise, the characteristic property

$$\langle (q(t) - q(t'))^2 \rangle_\nu = |t - t'|.$$

Moreover, the example exhibits an important property, also relevant for the general case. For $|t - t'| = \varepsilon$, this result implies that for $\varepsilon \rightarrow 0$, generic paths are not differentiable, but only continuous, or more precisely, on average satisfy

$$|q(t + \varepsilon) - q(t)| \underset{\varepsilon \rightarrow 0}{=} O(\sqrt{\varepsilon}). \quad (34.45)$$

This result is also valid for any Langevin equation of the form (34.1), because for $\varepsilon \rightarrow 0$ the driving term, which is of order ε , is negligible with respect to the noise, which is of order $\sqrt{\varepsilon}$. This property also defines the class of paths that contribute to the path integrals, discussed in Chapter 2. As for path integrals, the notation \dot{q} in the Langevin equation is, therefore, somewhat symbolic.

General situation. The average of equation (34.44) over the noise yields

$$\langle q(t) \rangle_\nu = q_0 e^{-\omega t}. \quad (34.46)$$

The second cumulant of the distribution of $q(t)$, calculated with the help of equation (34.44), is given by

$$\langle [q(t) - \langle q(t) \rangle_\nu]^2 \rangle_\nu = \frac{1}{2\omega} (1 - e^{-2\omega t}). \quad (34.47)$$

The expressions (34.46, 34.47) show that, for $\omega < 0$, the distribution of $q(t)$ has no large time limit. Moreover, for $q_0 \neq 0$, $\langle q(t) \rangle_\nu$ grows exponentially with time.

By contrast, for $\omega > 0$, both moments have a finite large-time limit.

Since $q(t)$ is linearly related to $\nu(t)$, the distribution $P(q, t; q_0, 0)$ is also Gaussian, and characterized by its two first moments $\langle q(t) \rangle$ and $\langle q^2(t) \rangle$. One infers,

$$P(q, t; q_0, 0) = \left[\frac{\pi}{\omega} (1 - e^{-2\omega t}) \right]^{-1/2} \exp \left[-\frac{\omega}{(1 - e^{-2\omega t})} (q - q_0 e^{-\omega t})^2 \right]. \quad (34.48)$$

The expression is analogous to expression (2.32). For $\omega > 0$ and for $t \rightarrow +\infty$, the distribution $P(q, t; q_0, 0)$ converges towards the *equilibrium distribution*

$$P(q, t; q_0, 0) \xrightarrow[t \rightarrow +\infty]{} \sqrt{\omega/\pi} e^{-\omega q^2}. \quad (34.49)$$

The two-point function and the sgn(0) problem. In Section 34.2.2, we have derived a general equation, the FP equation, satisfied by the probability distribution $P(q, t; q_0, t_0)$. To illustrate the difficulty one may encounter in the derivation of the equation (see also Section 3.3.1), we calculate, for $q_0 = 0$, the two-point function,

$$\langle q(t_1) q(t_2) \rangle_\nu = \frac{1}{2\omega} (e^{-\omega|t_2-t_1|} - e^{-\omega(t_1+t_2)}). \quad (34.50)$$

When the system is initially at equilibrium, time translation is restored, and the two-point function reduces to

$$\langle q(t_1)q(t_2) \rangle_\nu = \frac{1}{2\omega} e^{-\omega|t_2-t_1|}. \quad (34.51)$$

Differentiating with respect to t_1 , one finds

$$\langle \dot{q}(t_1)q(t_2) \rangle_\nu = -\frac{1}{2} \operatorname{sgn}(t_1 - t_2) e^{-\omega|t_2-t_1|}, \quad (34.52)$$

where $\operatorname{sgn}(t)$ is the sign function. The limit $t_1 \rightarrow t_2$ is not defined, because $\operatorname{sgn}(0)$ appears in the right-hand side (a similar problem is met in Section 3.3). By contrast, if one takes this limit in expression (34.50) first, and then differentiates, one finds

$$\frac{1}{2} \frac{d}{dt} \langle q^2(t) \rangle_\nu = \frac{1}{2} e^{-2\omega t}.$$

The difficulty is related to a property of the Langevin equation (34.1): $d\langle \mathbf{q}^2(t) \rangle / dt$ is defined, but $2\langle \mathbf{q}(t) \cdot \dot{\mathbf{q}}(t) \rangle$ is not. Since the noise two-point correlation function (equation (34.3)) is singular, the two operations, time differentiation and averaging, in general, do not commute.

However, the choice $\operatorname{sgn}(0) = 0$ in equation (34.52), that is taking the half sum of the derivative from above and below, at least ensures compatibility between the different results, and thus the commutation of averaging and time differentiation.

The FP Hamiltonian. After the transformation (34.37), the FP Hamiltonian associated with equation (34.43) takes the form

$$\mathbf{H} = \frac{1}{2}p^2 + \frac{1}{2}\omega^2q^2 - \frac{1}{2}\omega. \quad (34.53)$$

The eigenvalues ϵ_n of \mathbf{H} are

$$\epsilon_n = \left(n + \frac{1}{2}\right) |\omega| - \frac{1}{2}\omega, \quad n \geq 0. \quad (34.54)$$

If ω is positive, ϵ_0 vanishes. Correspondingly, the ground state eigenfunction is $e^{-\omega q^2/2}$, and one recovers the normalizable equilibrium distribution (34.49).

By contrast, if ω is negative, the lowest eigenvalue $\epsilon_0 = -\omega = |\omega|$ is positive, and the function $e^{-\omega q^2/2}$ is not normalizable. For large time, the FP distribution becomes

$$P(q, t; q_0, 0) \sim e^{-|\omega|t} e^{-|\omega|q_0^2} + O\left(e^{-2|\omega|t}\right). \quad (34.55)$$

The expression shows that the probability of finding q at a finite distance from the origin decreases exponentially at a rate $\tau = 1/|\omega|$.

Finally, the special case $\omega = 0$ corresponds to Brownian motion, the spectrum of the FP Hamiltonian is continuous and covers $[0, +\infty]$, and the probability of remaining at a finite distance from the origin decreases algebraically as $1/\sqrt{t}$.

34.6 Path integral representation

Applying the method described in Section 2.2 to expression (34.18), one derives a path integral representation for the probability distribution $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ (equation (34.8)) and averaged observables. Here, it is convenient to start directly from expression (34.19) and combine it with the Markov property (34.10). One obtains

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = \int_{\mathbf{q}(t_0) = \mathbf{q}_0}^{\mathbf{q}(t) = \mathbf{q}} [\mathrm{d}\mathbf{q}(\tau)] \exp [-\mathcal{S}(\mathbf{q})/\Omega], \quad (34.56)$$

with

$$\mathcal{S}(\mathbf{q}) = \lim_{\varepsilon \rightarrow 0} \sum_{k=1} \frac{1}{2\varepsilon} (\mathbf{q}_k - \mathbf{q}_{k-1} + \frac{1}{2}\varepsilon \mathbf{f}(\mathbf{q}_{k-1}))^2.$$

To study the small time-step limit, we expand in powers of ε up to order ε . We face problems we have discussed in Sections 3.2.2 and 3.3. Here we symmetrize the argument of \mathbf{f} . The important point is that the typical values of $\mathbf{q} - \mathbf{q}'$ are of order $\varepsilon^{1/2}$. Then,

$$\mathbf{q} - \mathbf{q}' + \frac{1}{2}\varepsilon \mathbf{f}(\mathbf{q}') = \mathbf{q} - \mathbf{q}' + \frac{1}{2}\varepsilon \mathbf{f}((\mathbf{q} + \mathbf{q}')/2) - \frac{1}{4}\varepsilon (\mathbf{q} - \mathbf{q}') \cdot \nabla_q \mathbf{f} + O(\varepsilon^2).$$

To eliminate the last term, we can change variables $\mathbf{q} \mapsto \tilde{\mathbf{q}}$, setting

$$\mathbf{q} - \mathbf{q}' - \frac{1}{4}\varepsilon (\mathbf{q} - \mathbf{q}') \cdot \nabla_q \mathbf{f} \mapsto \tilde{\mathbf{q}} - \mathbf{q}',$$

where we identify \mathbf{q} with \mathbf{q}_k (and \mathbf{q}' with \mathbf{q}_{k-1}) for successive values of k . The corresponding Jacobian is

$$\det \left[\frac{\partial}{\partial q_i} (q_j - q'_j - \frac{1}{4}\varepsilon (\mathbf{q} - \mathbf{q}') \cdot \nabla_q f_j) \right]^{-1} = \exp \left(\frac{1}{4}\varepsilon \nabla_q \cdot \mathbf{f}(\mathbf{q}) + O(\varepsilon^2) \right),$$

where the identity (A2.4), $\ln \det = \text{tr} \ln$, has been used. Finally, collecting all terms of order ε and taking the formal $\varepsilon \rightarrow 0$ limit, we obtain the dynamic action

$$\mathcal{S}(\mathbf{q}) = \int_{t_0}^t \frac{1}{2} \left\{ [\dot{\mathbf{q}}(\tau) + \frac{1}{2}\mathbf{f}(\mathbf{q}(\tau))]^2 - \frac{1}{2}\Omega \nabla \cdot \mathbf{f}(\mathbf{q}(\tau)) \right\} \mathrm{d}\tau. \quad (34.57)$$

Remarks

(i) Since the action contains a term linear in the time derivative, perturbative calculations involve the undefined quantity $\text{sgn}(0)$. Consistency with the choice of symmetrizing the argument of \mathbf{f} requires $\text{sgn}(0) = 0$.

(ii) Combining the representation (34.27), which assumes an initial equilibrium distribution, with the results of Section 2.5, one obtains the path integral representation of equilibrium correlation functions,

$$Z_{i_1 i_2 \dots i_n}^{(n)}(t_1, t_2, \dots, t_n) = \int [\mathrm{d}\mathbf{q}(\tau)] q_{i_1}(t_1) q_{i_2}(t_2) \cdots q_{i_n}(t_n) e^{-\mathcal{S}(\mathbf{q})/\Omega}, \quad (34.58)$$

with

$$\mathcal{S}(\mathbf{q}) = \int_{-\infty}^{\infty} \frac{1}{2} \left\{ [\dot{\mathbf{q}}(\tau) + \frac{1}{2}\mathbf{f}(\mathbf{q}(\tau))]^2 - \frac{1}{2}\Omega \nabla \cdot \mathbf{f}(\mathbf{q}(\tau)) \right\} \mathrm{d}\tau. \quad (34.59)$$

Perturbative expansion. The parameter Ω here plays the role \hbar plays in QM. It orders the perturbative expansion in the form of a loop expansion. At leading order, the ‘classical’ action reduces to

$$\mathcal{S}(\mathbf{q}) = \frac{1}{2} \int_{t_0}^{\infty} [\dot{\mathbf{q}}(\tau) + \frac{1}{2}\mathbf{f}(\mathbf{q}(\tau))]^2 d\tau. \quad (34.60)$$

The zeros of $\mathbf{f}(\mathbf{q})$ are the starting points of the perturbative expansion.

Dissipative Langevin equation. In the case of the dissipative Langevin equation with $\mathbf{f}(\mathbf{q}) = \Omega \nabla E(\mathbf{q})$ (equation (34.34)), the action (34.57) becomes

$$\begin{aligned} \mathcal{S}(\mathbf{q}) = & \frac{1}{2} \int_{t_0}^t d\tau \left\{ \dot{\mathbf{q}}^2(\tau) + \frac{1}{4}\Omega^2 [(\nabla E(\mathbf{q}(\tau))]^2 - \frac{1}{2}\Omega^2 \nabla^2 E(\mathbf{q}(\tau)) \right\} \\ & + \frac{1}{2}\Omega [E(\mathbf{q}(t)) - E(\mathbf{q}(t_0))]. \end{aligned} \quad (34.61)$$

The result is consistent with the equations (34.37, 34.39). Again, the convention $\text{sgn}(0) = 0$ is required, in order for differentiation and path integration to commute.

Perturbation theory, for $\Omega \rightarrow 0$ at ΩE fixed, has to be expanded around one of the extrema of $E(\mathbf{q})$. However, not all extrema are equivalent, once one takes into account the first correction due to the additional term $-\frac{1}{2}\Omega^2 \nabla^2 E(\mathbf{q})$. Indeed, as the example (expression (34.54)) of Section 34.5 shows, this term lifts the degeneracy between minima and maxima, and only minima are suitable starting points for a perturbative expansion.

34.7 BRST and supersymmetry

We now construct, by a different method, a path integral representation for the generating functional $\mathcal{Z}(\mathbf{b})$ of dynamic correlation functions of the path $\mathbf{q}(t)$, solution of the Langevin equation (34.34) of the special dissipative class, with the noise (34.3).

34.7.1 Dynamic action

The generating functional $\mathcal{Z}(\mathbf{b})$ of dynamic correlation functions of the path $q(t)$, solution of equation (35.1), is given by the noise expectation value,

$$\begin{aligned} \mathcal{Z}(\mathbf{b}) &= \left\langle \exp \left[\int dt \mathbf{b}(t) \cdot \mathbf{q}(t) \right] \right\rangle_{\nu}, \\ &= \int [d\nu] \exp \left[- \int dt \left(\frac{1}{2\Omega} \boldsymbol{\nu}^2(t) - \mathbf{b}(t) \cdot \mathbf{q}(t) \right) \right], \end{aligned} \quad (34.62)$$

where $\mathbf{q}(t)$ is solution to the Langevin equation (34.1), that we rewrite as

$$\mathbf{L}(q, \nu; t) \equiv \dot{\mathbf{q}}(t) + \frac{1}{2}\mathbf{f}(\mathbf{q}(t), t) - \boldsymbol{\nu}(t) = 0. \quad (34.63)$$

Generalizing to paths the method explained in Section 26.2, we insert the identity

$$\int [dq] \det \mathbf{M} \prod_t \delta[\mathbf{L}(q, \nu; t)] = 1, \quad (34.64)$$

where \mathbf{M} is the matrix and differential operator ($\partial_i \equiv \partial/\partial q_i$) with elements,

$$M_{ij}(t; t') = \frac{\delta L_i(q, \nu; t)}{\delta q_j(t')} = \left(\delta_{ij} \frac{\partial}{\partial t} + \frac{1}{2} \partial_j f_i(\mathbf{q}(t), t) \right) \delta(t - t'), \quad (34.65)$$

into expression (34.62). We then find,

$$\mathcal{Z}(\mathbf{b}) = \int [d\nu] [dq] \det \mathbf{M} \delta(\mathbf{L}(q, \nu; t)) \exp \left[- \int dt \left(\frac{1}{2\Omega} \boldsymbol{\nu}^2(t) - \mathbf{b}(t) \cdot \mathbf{q}(t) \right) \right]. \quad (34.66)$$

We express the functional δ function by a path Fourier representation:

$$\delta(\mathbf{L}(q, \nu; t)) = \int [d\lambda] \exp \left[\int dt \mathbf{L}(q, \nu; t) \cdot \boldsymbol{\lambda}(t) \right],$$

where $\boldsymbol{\lambda}(t)$ is imaginary.

A determinant can be expressed as a Grassmann integral (equation (1.69)). Introducing two Grassmann paths $\mathbf{c}(t)$ and $\bar{\mathbf{c}}(t)$, we can express $\det \mathbf{M}$ as

$$\det \mathbf{M} = \int [dc d\bar{c}] \exp \left[\int dt \sum_{i,j} c_i(t) \frac{\delta L_i(q, \nu; t)}{\delta q_j(t')} \bar{c}_j(t') \right].$$

The expression (34.62) then formally becomes

$$\mathcal{Z}(\mathbf{b}) = \int [dq d\nu d\lambda dc d\bar{c}] \exp \left[-\mathcal{S}(\mathbf{q}, \boldsymbol{\lambda}, \mathbf{c}, \bar{\mathbf{c}}, \boldsymbol{\nu}) + \int dt \mathbf{b}(t) \cdot \mathbf{q}(t) \right], \quad (34.67)$$

with

$$\begin{aligned} \mathcal{S}(\mathbf{q}, \boldsymbol{\lambda}, \mathbf{c}, \bar{\mathbf{c}}, \boldsymbol{\nu}) = & \frac{2}{\Omega} \int dt \left\{ \boldsymbol{\lambda}(t) \cdot [\dot{\mathbf{q}}(t) + \frac{1}{2}\mathbf{f}(\mathbf{q}(t), t) - \boldsymbol{\nu}(t)] \right. \\ & \left. - \sum_{i,j} c_i(t) \cdot [\delta_{ij} \partial_t + \frac{1}{2} \partial_j f_i(\mathbf{q}(t), t)] \bar{c}_j(t) + \frac{1}{4} \boldsymbol{\nu}^2(t) \right\}, \end{aligned} \quad (34.68)$$

where, for convenience, we have rescaled the Langevin equation by a factor $2/\Omega$.

After integration over the noise, one obtains

$$\begin{aligned} \mathcal{S}(\mathbf{q}, \boldsymbol{\lambda}, \mathbf{c}, \bar{\mathbf{c}}) = & \frac{2}{\Omega} \int dt \left\{ \boldsymbol{\lambda}(t) \cdot [\dot{\mathbf{q}}(t) + \frac{1}{2}\mathbf{f}(\mathbf{q}(t), t)] \right. \\ & \left. - \sum_{i,j} c_i(t) [\delta_{ij} \partial_t + \frac{1}{2} \partial_j f_i(\mathbf{q}(t), t)] \bar{c}_j(t) - \boldsymbol{\lambda}^2(t) \right\}. \end{aligned} \quad (34.69)$$

This construction provides an example of a Nicolai map [355].

It follows directly from the general analysis of Section 26.2, that the dynamic action has a BRST symmetry, corresponding to the transformations,

$$\delta_{\text{BRST}} \mathbf{q}(t) = \varepsilon \bar{\mathbf{c}}(t), \quad \delta_{\text{BRST}} \bar{\mathbf{c}}(t) = 0, \quad \delta_{\text{BRST}} \mathbf{c}(t) = \varepsilon \boldsymbol{\lambda}(t), \quad \delta_{\text{BRST}} \boldsymbol{\lambda}(t) = 0, \quad (34.70)$$

where ε is a Grassmann constant.

34.7.2 Grassmann coordinates: Supersymmetry

In Section 26.5.2, we have shown that when a stochastic equation has, in symbolic notation, the form $\delta \mathcal{A}/\delta \varphi = \nu$, the associated action $\mathcal{S}(\phi)$ has two independent BRST symmetries, which can be better represented by introducing two Grassmann coordinates. The Langevin equation in the special case (34.34):

$$\dot{\mathbf{q}}(t) = -\frac{1}{2} \Omega \nabla_q E(\mathbf{q}(t)) + \boldsymbol{\nu}(t), \quad (34.71)$$

almost shares the property. The dynamic action becomes

$$\begin{aligned} \mathcal{S}(\mathbf{q}, \boldsymbol{\lambda}, \mathbf{c}, \bar{\mathbf{c}}, \boldsymbol{\nu}) = & \frac{2}{\Omega} \int dt \left\{ \boldsymbol{\lambda}(t) \cdot [\dot{\mathbf{q}}(t) + \frac{1}{2} \Omega \nabla_q E(\mathbf{q}(t))] \right. \\ & \left. - \sum_{i,j} c_i(t) \cdot [\delta_{ij} \partial_t + \frac{1}{2} \Omega \partial_i \partial_j E(\mathbf{q}(t))] \bar{c}_j(t) - \boldsymbol{\lambda}^2(t) \right\}. \end{aligned} \quad (34.72)$$

The force term is a gradient, but not the time derivative $\dot{\mathbf{q}}(t)$. However, somewhat surprisingly, a slightly modified BRST transformation leaves the dynamic action (34.72) invariant. Combined, in a non-trivial way, with the first one, it provides the simplest example of *supersymmetry* [356].

We introduce two Grassmann coordinates $(\theta, \bar{\theta})$ (equation (26.68)), and the *superpath*,

$$\phi(t, \bar{\theta}, \theta) = \mathbf{q}(t) + \theta \bar{\mathbf{c}}(t) + \mathbf{c}(t)\bar{\theta} + \theta\bar{\theta}\boldsymbol{\lambda}(t). \quad (34.73)$$

One can verify that the action (34.69), corresponding to the Langevin equation (34.34) with the noise Gaussian distribution (34.3) expressed in terms of superpaths can be rewritten as (see also equation (26.71))

$$\mathcal{S}(\phi) = \int d\bar{\theta} d\theta dt \left[\frac{2}{\Omega} \bar{D}\phi(t, \bar{\theta}, \theta) \cdot D\phi(t, \bar{\theta}, \theta) + E(\phi(t, \bar{\theta}, \theta)) \right], \quad (34.74)$$

where, with respect to the supersymmetry, the operators,

$$\bar{D} = \frac{\partial}{\partial \bar{\theta}}, \quad D = \frac{\partial}{\partial \theta} - \bar{\theta} \frac{\partial}{\partial t}, \quad (34.75)$$

are *covariant derivatives*, and satisfy the anticommutation relations,

$$D^2 = \bar{D}^2 = 0, \quad D\bar{D} + \bar{D}D = -\frac{\partial}{\partial t}. \quad (34.76)$$

We also introduce the two other Grassmann-type differential operators

$$Q = \frac{\partial}{\partial \theta}, \quad \bar{Q} = \frac{\partial}{\partial \bar{\theta}} + \theta \frac{\partial}{\partial t}, \quad (34.77)$$

Both *anticommute* with D and \bar{D} and satisfy

$$Q^2 = \bar{Q}^2 = 0, \quad Q\bar{Q} + \bar{Q}Q = \frac{\partial}{\partial t}. \quad (34.78)$$

The pair Q, \bar{Q} are the generators of this simplest example of *supersymmetry* (see, for example, Section A26.4).

Let us verify the invariance of the action. We already know that Q (a translation of θ) is the generator of a BRST symmetry of the action (34.68). We now show that Q is the generator of an additional symmetry. We perform a variation of ϕ of the form ($\bar{\varepsilon}$ is a constant, additional element of the Grassmann algebra)

$$\delta_{\bar{\varepsilon}}\phi(t, \theta, \bar{\theta}) = \bar{\varepsilon}\bar{Q}\phi(t, \theta, \bar{\theta}), \quad (34.79)$$

which, in component form, reads (to be compared with expressions (26.67)),

$$\delta_{\bar{\varepsilon}}\mathbf{q}(t) = \mathbf{c}(t)\bar{\varepsilon}, \quad \delta_{\bar{\varepsilon}}\mathbf{c}(t) = 0, \quad \delta_{\bar{\varepsilon}}\bar{\mathbf{c}}(t) = (\boldsymbol{\lambda}(t) - \dot{\mathbf{q}}(t))\bar{\varepsilon}, \quad \delta_{\bar{\varepsilon}}\boldsymbol{\lambda}(t) = \dot{\mathbf{c}}(t)\bar{\varepsilon}. \quad (34.80)$$

The variation of the action density then is a total derivative. This is immediate for E , because it does not explicitly depend on t and $\bar{\theta}$. For the remaining term, the additional property that \bar{Q} anticommutes with D and \bar{D} has to be used:

$$\delta_{\bar{\varepsilon}} [\bar{D}\phi \cdot D\phi] = \bar{D} [\bar{\varepsilon}\bar{Q}\phi] \cdot D\phi + \bar{D}\phi \cdot D [\bar{\varepsilon}\bar{Q}\phi] = \bar{\varepsilon}\bar{Q} [\bar{D}\phi \cdot D\phi],$$

which is a total derivative.

The action (34.74) thus is supersymmetric. The supersymmetry is directly related to the property that the corresponding FP Hamiltonian (35.5) is then equivalent to a positive Hamiltonian of the form (34.39).

Remarks

(i) It is possible to emphasize the symmetric role played by $\bar{\theta}$ and θ by performing the substitution $t \mapsto t + \frac{1}{2}\theta\bar{\theta}$. Here, we find it more convenient to remain with the original variables.

(ii) The anticommutator (35.40) of \bar{Q} and Q generates time translations. Therefore, *supersymmetry implies time-translation invariance*.

(iii) Considering the fermions also as dynamic variables, one can interpret the Hamiltonian associated with the supersymmetric action in boson–fermion space. One concludes that the path integral describes both the Langevin equation and its time-reversed form.

(iv) Supersymmetry can be used to give a proof of detailed balance, alternative to the proof based on the FP equation given in Section 34.4.

(v) With equilibrium boundary conditions, supersymmetry makes it possible to prove that the equal-time correlation functions are the static correlation functions corresponding to the equilibrium distribution (see Section 35.5).

34.8 Gradient time-dependent force and Jarzynski's relation

We consider a Langevin equation of the form (34.42),

$$\dot{q}_i(t) = -\frac{1}{2}\Omega\partial_i E(\mathbf{q}(t), t) + \nu_i(t), \quad (34.81)$$

where $E(\mathbf{q}, t)$ is explicitly time dependent. Note that $E(\mathbf{q}, t)$ is determined by the Langevin equation only up to a function of time independent of \mathbf{q} .

In the case of a potential that is explicitly time dependent, as in Langevin equation (34.81), the FP equation (34.18) becomes

$$\frac{\partial P(\mathbf{q}, t; \mathbf{q}', t')}{\partial t} = \mathbf{H}(t)P(\mathbf{q}, t; \mathbf{q}', t'), \quad (34.82)$$

with

$$\mathbf{H}(t) = \Omega\nabla[\nabla + \nabla E(\mathbf{q}, t)], \quad (34.83)$$

or in operator form,

$$\frac{\partial \mathbf{P}(t, t')}{\partial t} = \mathbf{H}(t)\mathbf{P}(t, t'). \quad (34.84)$$

As a consequence, detailed balance (Section 34.4.2) is replaced by a more complicated relation.

Taking the Hermitian conjugate of equation (34.13), one obtains

$$\frac{\partial \mathbf{P}^\dagger(t, t')}{\partial t'} = \mathbf{H}^\dagger(t')\mathbf{P}^\dagger(t, t').$$

Using expression (34.83), one verifies the identity,

$$\mathbf{H}^\dagger(t') = e^{E(\hat{\mathbf{q}}, t')} \mathbf{H}(t') e^{-E(\hat{\mathbf{q}}, t')}.$$

Defining the operator

$$\mathcal{O}(t', t) = e^{-E(\hat{\mathbf{q}}, t')} \mathbf{P}^\dagger(t, t') e^{E(\hat{\mathbf{q}}, t)}, \quad (34.85)$$

one finds

$$\left(\frac{\partial}{\partial t'} + \frac{\partial E(\hat{\mathbf{q}}, t')}{\partial t'} \right) \mathcal{O}(t', t) = \mathbf{H}(t')\mathcal{O}(t', t),$$

with the boundary condition $\mathcal{O}(t, t) = \mathbf{1}$.

When $E(\mathbf{q}, t) \equiv E(\mathbf{q})$ is time independent, this relation again leads to detailed balance. More generally, taking matrix elements, one finds

$$\langle \mathbf{q} | \mathcal{O}(t', t) | \mathbf{q}' \rangle = e^{-E(\mathbf{q}, t')} \langle \mathbf{q}' | \mathbf{P}(t, t') | \mathbf{q} \rangle e^{E(\mathbf{q}', t)}, \quad (34.86)$$

where

$$\left(\frac{\partial}{\partial t'} + \frac{\partial E(\mathbf{q}, t')}{\partial t'} \right) \langle \mathbf{q} | \mathcal{O}(t', t) | \mathbf{q}' \rangle = -\mathbf{H}(t') \langle \mathbf{q} | \mathcal{O}(t', t) | \mathbf{q}' \rangle.$$

An interpretation of the matrix elements of the operator (34.85) will be provided. Jarzynski's relation then follows from equation (34.86).

Path integral representation. The cross-term in the expansion of the square in expression (34.57) can no longer be integrated and, instead yields

$$\int_{t'}^{t''} d\tau \dot{\mathbf{q}}(\tau) \cdot \nabla_q E(\mathbf{q}(\tau), \tau) = E(\mathbf{q}(t''), t'') - E(\mathbf{q}(t'), t') - \int_{t'}^{t''} d\tau \frac{\partial E(\mathbf{q}(\tau), \tau)}{\partial \tau}. \quad (34.87)$$

It follows that

$$\begin{aligned} & \frac{1}{2} \int_{t'}^{t''} d\tau [\dot{\mathbf{q}}(\tau) + \frac{1}{2}\Omega \nabla_q E(\mathbf{q}(\tau), \tau)]^2 + \Omega \int_{t'}^{t''} d\tau \frac{\partial E(\mathbf{q}(\tau), \tau)}{\partial \tau} \\ &= \frac{1}{2} \int_{t'}^{t''} d\tau [\dot{\mathbf{q}}(\tau) - \frac{1}{2}\Omega \nabla_q E(\mathbf{q}(\tau), \tau)]^2 + \Omega [E(\mathbf{q}(t''), t'') - E(\mathbf{q}(t'), t')]. \end{aligned}$$

With the same arguments as in the proof of detailed balance, we obtain

$$\left\langle \exp \left[- \int_{t'}^{t''} d\tau \frac{\partial E}{\partial \tau} \right] \right\rangle_\nu = e^{-[E(\mathbf{q}'', t'') - E(\mathbf{q}', t')]} P_*(\mathbf{q}', t''; \mathbf{q}'', t'),$$

where the average in the left-hand side is taken at \mathbf{q}' and \mathbf{q}'' fixed, and P_* corresponds to the process with $E(\mathbf{q}, -t)$.

We now assume that $E(\mathbf{q}, t)$ is time independent for $t \leq t'$, and that the system is at equilibrium at $t = t'$. Therefore, we can integrate the left-hand side over all \mathbf{q}' with the normalized weight $e^{-E(\mathbf{q}', t')} / \mathcal{Z}'$, where

$$\mathcal{Z}' = \int d^d q' e^{-E(\mathbf{q}', t')}.$$

Then,

$$\left\langle \exp \left[- \int_{t'}^{t''} d\tau \frac{\partial E}{\partial \tau} \right] \right\rangle_{\nu, \mathbf{q}'} = e^{-E(\mathbf{q}'', t'')} / \mathcal{Z}'.$$

Finally, after an integration over all \mathbf{q}'' , one obtains Jarzynski's relation [357],

$$\left\langle \exp \left[- \int_{t'}^{t''} d\tau \frac{\partial E}{\partial \tau} \right] \right\rangle_{\nu, \mathbf{q}', \mathbf{q}''} = \mathcal{Z}'' / \mathcal{Z}' = e^{\mathcal{W}'' - \mathcal{W}'}, \quad (34.88)$$

where \mathcal{W}' and \mathcal{W}'' are the initial and final free energies, respectively.

Remarks.

(i) Since, in terms of the Langevin equation, $E(\mathbf{q}, t)$ is defined only up to a function of time, the free-energy variation depends on a more precise definition of $E(\mathbf{q}, t)$.

(ii) Relation (34.88) can be modified by adding any total time derivative to $\partial E / \partial t$.

34.9 More general Langevin equations. Motion in Riemannian manifolds

We now discuss a more general class of Langevin equation. Since the main motivation is Brownian motion on Riemannian manifolds, it is convenient to adopt the notation and conventions of Chapter 28, to which we refer for details, with, in particular, *summation over repeated lower and upper indices for tensors*. We also set $\Omega = 1$.

We consider the general Langevin equation

$$\dot{q}^i(t) = -\frac{1}{2}f^i(\mathbf{q}(t)) + e_a^i(\mathbf{q}(t))\nu_a(t), \quad (34.89)$$

in which $\nu_a(t)$ is the Gaussian white noise defined by

$$\langle \nu_a(t) \rangle = 0, \quad \langle \nu_a(t)\nu_b(t') \rangle = \delta(t-t')\delta_{ab}. \quad (34.90)$$

We assume that $e_a^i(\mathbf{q})$ is a square matrix and $\det e_a^i(\mathbf{q}) \neq 0$, but the generalization of the situation where ν has more components than \mathbf{q} is simple.

To clarify the meaning of the Langevin equation, we integrate it between times t and $t + \varepsilon$ (see Section 34.1.1). For $\varepsilon \rightarrow 0$,

$$q^i(t + \varepsilon) = q'^i - \frac{1}{2}\varepsilon f^i(\mathbf{q}') + \int_t^{t+\varepsilon} d\tau e_a^i(\mathbf{q}(\tau))\nu_a(\tau) + O(\varepsilon^{3/2}). \quad (34.91)$$

(with the notation $\mathbf{q}(t) \equiv \mathbf{q}'$). The new feature is the *q dependence of the noise term*. Since the variation of \mathbf{q} is of order $\sqrt{\varepsilon}$, ν_a is also of order $\sqrt{\varepsilon}$, and we need all terms up to order ε , we have to expand

$$e_a^i(\mathbf{q}(\tau)) = e_a^i(\mathbf{q}') + (q^j(\tau) - q'^j)\partial_j e_a^i(\mathbf{q}') + O(\varepsilon).$$

At leading order, for $\varepsilon \rightarrow 0$, the variation of $\mathbf{q}(\tau)$ is dominated by the noise and, therefore,

$$q^j(\tau) - q'^j \sim e_b^j(\mathbf{q}') \int_t^\tau \nu_b(\tau') d\tau'.$$

At order ε , one finds

$$\begin{aligned} q^i(t + \varepsilon) &= q'^i - \frac{1}{2}\varepsilon f^i(\mathbf{q}') + e_a^i(\mathbf{q}') \int_t^{t+\varepsilon} d\tau \nu_a(\tau) \\ &\quad + \partial_j e_a^i(\mathbf{q}') e_b^j(\mathbf{q}') \int_t^{t+\varepsilon} d\tau \int_t^\tau d\tau' \nu_a(\tau) \nu_b(\tau') + O(\varepsilon^{3/2}). \end{aligned} \quad (34.92)$$

A term quadratic in the noise has appeared which, at this order, can be replaced by its average

$$\begin{aligned} \int_t^{t+\varepsilon} d\tau \int_t^\tau d\tau' \langle \nu_a(\tau) \nu_b(\tau') \rangle &= \delta_{ab} \int_t^{t+\varepsilon} d\tau \int_t^\tau d\tau' \delta(\tau - \tau'), \\ &= \delta_{ab} \varepsilon \theta(0), \end{aligned}$$

where $\theta(t) = \frac{1}{2}(1 + \text{sgn}(t))$ is the step function. We obtain a term proportional to $\theta(0)$, which is not defined.

The simplest choice, $\theta(0) = 0$ (Itô calculus), leads to the discretized Langevin equation

$$\dot{q}^i(t + \varepsilon) = q^i(t) - \frac{1}{2}\varepsilon f^i(\mathbf{q}') + \sqrt{\varepsilon} e_a^i(\mathbf{q}') \bar{\nu}_a(t), \quad (34.93)$$

where $\bar{\nu}(t)$ has been defined in equations (34.6, 34.7):

$$\langle \bar{\nu}_a(t) \rangle = 0, \quad \langle \bar{\nu}_a(t) \bar{\nu}_b(t') \rangle = \varepsilon \delta_{tt'} \delta_{ab}. \quad (34.94)$$

In equation (34.91), it corresponds to the replacement of the noise term by $\sqrt{\varepsilon}e_a^i(\mathbf{q}')\bar{\nu}_a(t)$.

By contrast, $\theta(0) = 1/2$ (Stratanovich convention) leads to

$$q^i(t + \varepsilon) = q^i(t) - \frac{1}{2}\varepsilon f^i(\mathbf{q}(t)) + \frac{1}{2}\varepsilon \partial_j e_a^i(\mathbf{q}(t))e_a^j(\mathbf{q}(t)) + \sqrt{\varepsilon}e_a^i(\mathbf{q}(t))\bar{\nu}_a(t). \quad (34.95)$$

One can verify that this choice corresponds to the replacement of the noise term by $\sqrt{\varepsilon}e_a^i[\frac{1}{2}(\mathbf{q}(t) + \mathbf{q}(t + \varepsilon))]\bar{\nu}(t)$. Indeed,

$$\begin{aligned} \sqrt{\varepsilon}e_a^i\left\{\frac{1}{2}[\mathbf{q}(t) + \mathbf{q}(t + \varepsilon)]\right\}\bar{\nu}_a(t) &= \sqrt{\varepsilon}e_a^i[\mathbf{q}(t)]\bar{\nu}_a(t) + \frac{1}{2}\varepsilon e_b^j[\mathbf{q}(t)]\partial_j e_a^i[\mathbf{q}(t)]\bar{\nu}_b(t)\bar{\nu}_a(t) \\ &\quad + O(\varepsilon^{3/2}), \\ &= \sqrt{\varepsilon}e_a^i[\mathbf{q}(t)]\bar{\nu}_a(t) + \frac{1}{2}\varepsilon e_a^j[\mathbf{q}(t)]\partial_j e_a^i[\mathbf{q}(t)] + O(\varepsilon^{3/2}), \end{aligned}$$

because, at leading order for $\varepsilon \rightarrow 0$, we can replace $\bar{\nu}_b(t)\bar{\nu}_a(t)$ by its average value δ_{ab} . Therefore, the two conventions are indistinguishable in the formal continuum limit.

Remark. In contrast with the Langevin equation (34.1), equation (34.89) is not precisely defined, the difficulty being analogous to the problem of ordering operators arising in the quantization of a classical Hamiltonian of the form (3.24).

In what follows, for simplicity, we use the convention of equation (34.93), although Stratanovich's convention is more symmetric, and has simpler transformation properties under a change of variables and, therefore, of coordinates on a manifold.

34.9.1 The FP equation

From the Langevin equation (34.93), we derive an equation for the probability distribution $P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t)$:

$$P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) = \langle \mathbf{q} - \mathbf{q}(t + \varepsilon) \rangle_{\bar{\nu}}.$$

Following the method of Section 34.1.1, we first calculate \tilde{P} , the Fourier transform of P :

$$\tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) = \int dq e^{-i\mathbf{p}\cdot\mathbf{q}} P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) = \left\langle e^{-i\mathbf{p}\cdot\mathbf{q}(t+\varepsilon)} \right\rangle_{\bar{\nu}}.$$

Using equation (34.93) and the noise distribution, we obtain

$$\tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) \propto \int d\bar{\nu} e^{-(\bar{\nu}_a)^2/2\varepsilon} \exp \left[-ip_i (q'^i - \frac{1}{2}\varepsilon f^i(\mathbf{q}') + e_a^i(\mathbf{q}')\nu_a) \right].$$

The Gaussian integration over the variables ν_a then leads to

$$\tilde{P}(\mathbf{p}, t + \varepsilon; \mathbf{q}', t) \propto \exp \left[-ip_i (q'^i - \frac{1}{2}\varepsilon f^i(\mathbf{q}')) - \frac{1}{2}\varepsilon g^{ij}(\mathbf{q}') p_i p_j \right], \quad (34.96)$$

where we have introduced the positive symmetric matrix

$$g^{ij}(\mathbf{q}) = e_a^i(\mathbf{q})e_a^j(\mathbf{q}). \quad (34.97)$$

We denote by g_{ij} its inverse, a standard notation, and its determinant by $g = \det(g_{ij})$,

$$g^{ij}g_{jk} = \delta_k^i, \quad g = \det(g_{ij}). \quad (34.98)$$

Then,

$$\det e_a^i(\mathbf{q}) = \sqrt{\det g^{ij}(\mathbf{q})}.$$

If we invert the Fourier transformation, we obtain

$$P(\mathbf{q}, t + \varepsilon; \mathbf{q}', t) = \frac{1}{(2\pi\varepsilon)^{d/2}} \frac{1}{\sqrt{g(\mathbf{q}')}} \exp\left(-\frac{1}{2\varepsilon} g_{ij}(\mathbf{q}') d^i d^j\right), \quad (34.99)$$

in which we have set

$$\mathbf{d}(\mathbf{q}, \mathbf{q}') = \mathbf{q} - \mathbf{q}' + \frac{1}{2}\varepsilon \mathbf{f}(\mathbf{q}'). \quad (34.100)$$

Again the distribution P can be written in terms of the matrix elements of an operator of the form $e^{-\varepsilon \mathbf{H}}$, where \mathbf{H} is a second order differential operator. Using the method of Section 34.2.2, we expand equation (34.96) to first order in ε . Collecting the terms of order ε , we infer the matrix elements of the corresponding Hamiltonian in the mixed \mathbf{p}, \mathbf{q}' representation:

$$\langle \mathbf{p} | \mathbf{H} | \mathbf{q}' \rangle = \frac{1}{2} e^{-i\mathbf{p} \cdot \mathbf{q}'} [g^{ij}(\mathbf{q}') p_i p_j - i p_i f^i(\mathbf{q}')]. \quad (34.101)$$

Inverting the Fourier transformation, we obtain the corresponding FP Hamiltonian, in quantum operator notation,

$$\mathbf{H} = \frac{1}{2} \hat{p}_i [\hat{p}_j g^{ij}(\hat{\mathbf{q}}) - i f^i(\hat{\mathbf{q}})]. \quad (34.102)$$

The FP equation for the time-dependent distribution $P(\mathbf{q}, t)$ follows:

$$\dot{P}(\mathbf{q}, t) = \frac{1}{2} \partial_i [\partial_j (g^{ij}(\mathbf{q}) P(\mathbf{q}, t)) + f^i(\mathbf{q}) P(\mathbf{q}, t)]. \quad (34.103)$$

The path integral representation. Starting from the Hamiltonian (34.102), one obtains the formal path integral representation in phase space (see Chapter 3),

$$P = \int [d^d p d^d q] e^{-S(\mathbf{p}, \mathbf{q})},$$

with

$$S(\mathbf{p}, \mathbf{q}) = \int dt [i\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - \frac{1}{2} p_i(t) p_j(t) g^{ij}(\mathbf{q}(t)) + \frac{1}{2} i\mathbf{p}(t) \cdot \mathbf{f}(\mathbf{q}(t))]. \quad (34.104)$$

The integration over $\mathbf{p}(t)$ is Gaussian and can be performed explicitly. However, we have noted, in Section 3.3.2, that the resulting path integral is more singular (see equation (34.107)) due to the functional measure, while the path integral in phase space has only a finite ambiguity corresponding to the order between quantum operators in products.

Brownian motion on a Riemannian manifold. Covariance under a change of parametrization of the manifold implies that the FP equation for the free-Brownian motion on a manifold is expected to have the form (for details see Chapter 28, in particular, Section 28.3.1),

$$\dot{D}(\mathbf{q}, t) = \frac{1}{2\sqrt{g(\mathbf{q})}} \partial_i [\sqrt{g(\mathbf{q})} g^{ij}(\mathbf{q}) \partial_j D(\mathbf{q}, t)], \quad (34.105)$$

in which $g_{ij}(\mathbf{q})$ is the metric tensor in the manifold, g the determinant of $g_{ij}(\mathbf{q})$ (equations (34.98)), and $D(\mathbf{q}, t)$ a scalar density, that is, a density normalized with the covariant measure $\sqrt{g} d^d q$.

Quite generally, the functions $P(\mathbf{q}, t)$ and $D(\mathbf{q}, t)$ are related by

$$\sqrt{g(\mathbf{q})}D(\mathbf{q}, t) = P(\mathbf{q}, t). \quad (34.106)$$

This also implies that e_a^i is the inverse of the vielbein (Section 28.5.1). If the manifold is compact, a constant scalar density D is obviously a static solution to equation (34.105) and, therefore, corresponds to the equilibrium distribution.

Comparing equations (34.103) and (34.105), we note that the free-Brownian motion, in general, does not correspond to $\mathbf{f} = 0$, since, as a short calculation shows, equations (34.103) and (34.105) are identical only for a particular form of $f^i(\mathbf{q})$, a form which depends on the choice of discretization of the Langevin equation. The equation shows that \mathbf{f} does not only depend on the geometry of the manifold, but also on the choice of the vielbein.

Finally, we note that equation (34.103) is the most general second order stochastic differential equation: the operator $\partial/\partial q^i$ is implied by total probability conservation, and the positivity of the coefficient of $\partial^2/\partial q^i \partial q^j$ by the condition that the solution remains a positive distribution at all times, as the corresponding path integral representation will show.

The density D has a path integral representation, which, following the discussion of Section 3.3.2, has the formally covariant form

$$D = \int \left[dq(\tau) g^{1/2}(\mathbf{q}(\tau)) \right] \exp [-\mathcal{S}(\mathbf{q})], \quad (34.107)$$

with

$$\mathcal{S}(\mathbf{q}) = \frac{1}{2} \int d\tau \dot{q}^i(\tau) g_{ij}(\mathbf{q}(\tau)) \dot{q}^j(\tau), \quad (34.108)$$

and where $g^{1/2}(\mathbf{q})$ is formally the covariant measure in the manifold. This form is obtained from the expression (34.104) by integrating over $\mathbf{p}(t)$.

We have already discussed in Section 3.3.2 and previously in the section the difficulties associated with this continuum expression.

Supersymmetric formulation. We expect the supersymmetric action (34.74) to have a generalization of the form

$$\mathcal{S}(\phi) = \int d\bar{\theta} d\theta dt \left[\frac{2}{\Omega} \bar{D}\phi^i(t, \bar{\theta}, \theta) g_{ij}(\phi(t)) D\phi^j(t, \bar{\theta}, \theta) + E(\phi(t, \bar{\theta}, \theta)) \right], \quad (34.109)$$

and the path integration measure to be multiplied by $\sqrt{g(\phi(t, \bar{\theta}, \theta))}$.

A34 Markov's stochastic processes: A few remarks

A34.1 Discrete spaces: Markov's processes, phase transitions

In this section, we recall a few simple properties of Markov's processes in discrete spaces, because they illustrate, in a simpler context, similar properties met while discussing Langevin or FP equations. For additional comments on numerical simulations and stochastic processes we refer to Refs. [358].

A34.1.1 Evolution or master equation

We first consider a finite number N of discrete states labelled in the following by indices like a , b and so on. A stochastic process is defined in terms of a matrix Π_{ab} ($\Pi_{ab} \geq 0, \forall a, b$), which gives the probability of a transition at any given time from a state b to a state a . The conservation of probabilities implies,

$$\sum_{a=1}^N \Pi_{ab} = 1. \quad (\text{A34.1})$$

The probability $P_n(a)$ to be at time n in a state a then satisfies the evolution (or master) equation,

$$P_{n+1}(a) = \sum_b \Pi_{ab} P_n(b). \quad (\text{A34.2})$$

Summing equation (A34.2) over the index a and, using equation (A34.1), one verifies that the total probability is conserved:

$$\sum_a P_{n+1}(a) = \sum_b P_n(b).$$

A34.1.2 The spectrum of the transition matrix

Equation (A34.1) also implies that the vector U_a with all components $U_a = 1$ is a left eigenvector of the matrix Π_{ab} with eigenvalue 1:

$$\sum_a U_a \Pi_{ab} = \sum_a \Pi_{ab} = 1 = U_b.$$

The corresponding right eigenvector,

$$\tilde{U}_a = \sum_b \Pi_{ab} \tilde{U}_b, \quad (\text{A34.3})$$

is a stationary solution of equation (A34.2).

More generally, let V_a be an eigenvector of Π_{ab} , and v the corresponding eigenvalue. Then,

$$\sum_b \Pi_{ab} V_b = v V_a. \quad (\text{A34.4})$$

Summing over a and using equation (A34.1), one obtains

$$\sum_b V_b = v \sum_a V_a,$$

an equation that implies either $v = 1$, or $\sum V_a = 0$.

Comparing the modulus of both sides of equation (A34.4), one infers

$$\sum_b \Pi_{ab} |V_b| \geq |v| |V_a|, \quad (\text{A34.5})$$

and thus, using equation (A34.1) again,

$$|v| \leq 1. \quad (\text{A34.6})$$

All eigenvalues of the matrix Π_{ab} have a modulus smaller than or equal to one. Equality is possible only if the inequality (A34.5) is an equality for all values of a :

$$\sum_b \Pi_{ab} |V_b| = |V_a|. \quad (\text{A34.7})$$

The equation implies that $|V_a|$ is a stationary probability distribution.

Let us assume that for some subset I of $\{1, \dots, N\}$, V_a vanishes. Then,

$$\forall a \in I : \sum_{b \notin I} \Pi_{ab} |V_b| = 0. \quad (\text{A34.8})$$

The equation implies that Π_{ab} then vanishes. As a consequence, if the initial state does not belong to I , it will never belong to it at a later time. In the sense that will be defined *the space of states is disconnected*.

Connectivity assumption. From now on, we assume that starting from any state a there is a non-zero probability to reach any other state b . This means that there exists a set of states c_1, \dots, c_r such that the product,

$$\forall a, b \exists \{c_1, \dots, c_r\} : \Pi_{bc_r} \Pi_{c_r c_{r-1}} \cdots \Pi_{c_1 a} > 0.$$

We then call such a space of states *connected*.

As shown previously, it follows that the stationary distribution has no vanishing component. Moreover, the eigenspace corresponding to the eigenvalue 1 has dimension 1 (if the dimension were larger than 1, one could form, by linear combination of eigenvectors, another stationary distribution with one vanishing component).

The stationary distribution $P(a)$ can thus be parametrized as

$$P(a) = e^{-E(a)} / \mathcal{Z}, \quad \text{with} \quad E(a) > 0 \quad \text{and} \quad \mathcal{Z} = \sum_a e^{-E(a)}. \quad (\text{A34.9})$$

If all other eigenvalues of the transition matrix have a modulus smaller than 1, then any distribution converges at infinite time towards the stationary distribution $P(a)$, which thus is the *equilibrium distribution*.

Eigenvalues with modulus 1. If there exists another eigenvalue v with modulus 1,

$$v = e^{i\theta}, \quad \theta \neq 0 \pmod{2\pi},$$

the corresponding eigenspace also has dimension 1, and the eigenvector V_a satisfies $|V_a| = P(a)$, where $P(a)$ is the stationary distribution (A34.9).

We then decompose the set of integers $\{1, N\}$ into a union of subsets I_k such that, within a subset, the components of V_a all have the same phase. The inequality (A34.5) implies that, at a fixed, all states b for which Π_{ab} is non-vanishing belong to the same subset. Moreover, the phase of V_a is then $e^{-i\theta}$ times the phase of V_b . Therefore, Π_{ab} has non-vanishing elements only between different subsets. The subsets I_k are characterized by a phase $e^{-i(k-1)\theta}$, and the non-vanishing elements of the matrix Π_{ab} are such that, if b belongs to I_k , then a belongs to I_{k+1} . Finally, since the number of subsets is finite, there exists an integer n such that $e^{in\theta} = 1$, that is, the eigenvalue is a root of unity.

The transition matrix thus induces a cyclic permutation between the n subsets.

Infinite number of states. In the limit in which the number of states increases to infinity, an important new phenomenon may occur: the components of the equilibrium distribution, normalized by $\sum_a P(a) = 1$, may all go to 0.

In this case, even with the preceding assumptions, there is no equilibrium distribution. This is an important issue in realistic examples.

We shall also encounter another situation in the study of critical phenomena in Chapters 14–36. In the case of ferromagnetic systems, the states correspond to configurations of spins on a lattice: as long as the spin system is in a finite volume, the space of states remains connected, but the space becomes disconnected in the infinite volume limit. In a finite volume, the equilibrium state is unique, but not in an infinite volume.

A34.1.3 Detailed balance

One is sometimes confronted with the following problem: one has chosen a priori an equilibrium distribution $P(a)$, and one wants to construct a stochastic process that converges towards this distribution. We assume here that the space of states is connected, and that all probabilities $P(a)$ are thus, strictly positive.

The construction can be simplified by imposing to the matrix Π_{ab} the condition,

$$\Pi_{ab}P(b) = \Pi_{ba}P(a), \quad \text{for all pairs } (a, b). \quad (\text{A34.10})$$

This condition is called *detailed balance* (see also equation (34.41)). It is a local condition involving only the states a and b .

The condition implies that $P(a)$ is a stationary distribution. Indeed,

$$\sum_b \Pi_{ab}P(b) = \sum_b \Pi_{ba}P(a) = P(a). \quad (\text{A34.11})$$

We then set

$$P_n(a) = \sqrt{P(a)} \tilde{P}_n(a). \quad (\text{A34.12})$$

The equation (A34.2) becomes

$$\tilde{P}_{n+1}(a) = \sum_b \tilde{\Pi}_{ab} \tilde{P}_n(b), \quad (\text{A34.13})$$

where the matrix

$$\tilde{\Pi}_{ab} = \Pi_{ab} \sqrt{P(b)} / \sqrt{P(a)}, \quad (\text{A34.14})$$

is symmetric as a consequence of equation (A34.10). Its spectrum is thus real.

The distribution $P(a)$ is an equilibrium distribution, except if Π_{ab} has -1 as eigenvalue. In this case, according to preceding discussion, the space of states is divided into two subsets I_+ and I_- , which the matrix Π_{ab} exchanges.

It is easy to see that the corresponding left eigenvector has components +1 for one subset and -1 for the others. This left eigenvector has to be orthogonal to the right eigenvector $P(a)$.

This implies the condition

$$\sum_{a \in I_+} P(a) = \sum_{a \in I_-} P(a).$$

This is a non-generic situation, which is likely to be realized only if there exists some discrete symmetry in the space of states.

A34.2 Stochastic process with prescribed equilibrium distribution

There are many ways to construct stochastic processes that converge towards a given equilibrium distribution. First, one has to construct the matrix Π_{ab} . We give here examples based on detailed balance in the case of a discrete set of states.

One possibility is to connect all initial states to the same number r of final states and then take

$$\begin{aligned} \Pi_{ab} &= \frac{1}{r}, & \text{if } P(a) \geq P(b) \text{ and } a \neq b, \\ \Pi_{ab} &= \frac{1}{r} \frac{P(a)}{P(b)}, & \text{if } P(a) < P(b). \end{aligned} \quad (A34.15)$$

These conditions imply detailed balance. In addition, as required,

$$\Pi_{bb} = 1 - \sum_{a \neq b} \Pi_{ab} > 0. \quad (A34.16)$$

Another idea is to take

$$\Pi_{ab} = p P(a) \theta_{ab}, \quad a \neq b, \quad (A34.17)$$

with

$$\theta_{ab} = \theta_{ba} \in \{0, 1\},$$

and p adjusted such that for any state b

$$p \sum_{a \neq b} P(a) \theta_{ab} < 1.$$

Depending on the structure of the space of states, there are many other methods.

Once the matrix Π_{ab} is given, in the discrete case, it is easy to construct the corresponding stochastic process, that is to describe a motion in the space of states such that, asymptotically, at large time the probability of being at state a is just $P(a)$. At fixed initial state b , to the matrix elements Π_{ab} , corresponds a partition of the interval $[0, 1]$. By drawing a random number with uniform probability on $[0, 1]$, one can select the final state a with probability Π_{ab} .

A34.3 Stochastic processes and phase transitions

To construct a stochastic process that converges towards the equilibrium distribution of the Ising model on a d -dimensional lattice $\Omega \in \mathbb{Z}^d$ of linear size L , one can use detailed balance (see Section A34.1) and choose for the transition probability p of a configuration $\{S'_r\}$ towards a configuration $\{S'_r\}$, with $\mathbf{r}, \mathbf{r}' \in \Omega$,

$$\begin{aligned} p(S_r, S'_r) &= e^{-\beta [E(S'_r) - E(S_r)]} && \text{for } E(S_r) < E(S'_r), \\ p(S_r, S'_r) &= 1, && \text{otherwise,} \end{aligned} \quad (A34.18)$$

in which $E(S_r)$ is the configuration energy (n.n. stands for nearest neighbour),

$$E(S) = -J \sum_{\substack{\text{n.n.} \\ \mathbf{r}, \mathbf{r}' \in \Omega \subset \mathbb{Z}^d}} S_r S'_{\mathbf{r}}, \quad (A34.19)$$

where the constant J is positive for ferromagnetic systems. For the argument that follows, the precise description of which configurations are directly connected by the matrix \mathbf{p} is irrelevant, provided the system is globally connected. The relevant property is that the probability to go from a configuration to a configuration of higher energy is, at low temperature, of the order of $e^{-\beta \Delta E}$, in which ΔE is the energy difference.

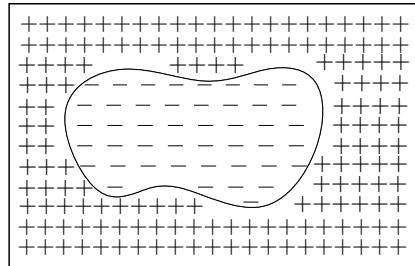


Fig. 34.1 A bubble of spins – in a background of spins +

Therefore, at low temperature, if the initial configuration corresponds, for example, to all spins equal to +1, the probability of creating a bubble of minus spins is proportional to $e^{-\beta J A}$, in which A is the area of the bubble surface (Fig. 34.1). If a large fluctuation creates a bubble that fills half of the total volume, then, clearly, there is a probability 1/2 that afterwards all spins become equal to -1. Therefore, the probability of reversing most of the spins is, at low temperature, related to the area of a surface which divides the volume Ω into two halves. Since Ω has linear size L , this probability is of the form $e^{-\sigma L^{d-1}}$.

For $d = 1$, the system is ergodic, and, as we have shown, no phase transitions can occur.

For $d > 1$, the same mechanism which makes the existence of several phases at low temperature possible, is responsible for the breaking of ergodicity.

Note that the argument generalizes only to local dynamics, that is, dynamics in which the probability of change of one spin on the lattice is only influenced by its neighbours (for more details, see Section 36.6.3).

35 Langevin field equations: Properties and renormalization

An important class of stochastic field equations in physics [359] corresponds to extensions to field theory of the Langevin equations introduced in Chapter 34. Langevin equations for fields have been proposed to describe the dynamics of critical phenomena (see Chapter 36) not too far from equilibrium, or as an alternative method of quantization, which could be useful in cases where straightforward methods lead to difficulties, like in gauge theories. Some of their general properties will be described in this chapter.

For a number of problems, in particular, related to perturbation theory, it is more convenient to work with an action and a field integral than with the equation directly, because standard methods of quantum field theory (QFT) then become available. For this purpose, generalizing the path integral formalism of Section 34.6 to field equations, in Section 35.2.1 we associate a field integral representation to the Langevin equation [354]. In this context, the corresponding action can be naturally called *dynamic action*.

In the framework of Sections 26.1, 26.2, the dynamic action can be interpreted as generated by the Langevin equation considered as a constraint equation. We have shown, quite generally, that the integral representation of constraint equations, including stochastic equations, leads to an action that has a Slavnov–Taylor (ST) symmetry and, in a different form, has an anticommuting type Becchi–Rouet–Stora–Tyutin (BRST) symmetry, a symmetry that involves anticommuting parameters. This symmetry has no geometric origin but is merely a consequence of associating a specific form of integral representations to the constraint equations, as explained in Section 26.2 [285]. This symmetry is used in a number of situations, for example, to prove the renormalizability of non-Abelian gauge theories in Chapter 26, or of models defined on homogeneous spaces (Section 29.3).

An analytic method to solve Langevin equations for fields involves starting from a linear approximation, and expanding the remainder in perturbation theory. Divergences appear in perturbative calculations, and it is then necessary to understand how Langevin equations renormalize. In the field integral representation, the dynamical actions exhibit ST or BRST symmetries. The BRST symmetry implies Ward–Takahashi (WT) identities, which can be used to prove that, under some general conditions, the structure of the Langevin equation is stable under renormalization [285].

35.1 Langevin and Fokker–Planck (FP) equations

To keep the notation simple, we discuss the properties of Langevin equations governing the stochastic evolution of a one-component scalar field $\varphi(t, x)$, where x is the position in the d -dimensional Euclidean space, and t the time. We consider a Langevin equation, first-order in time differential equation, of the generic form ($\dot{\varphi} \equiv \partial\varphi/\partial t$)

$$\dot{\varphi}(t, x) = -\frac{1}{2}L[\varphi(t, x)] + \nu(t, x), \quad (35.1)$$

where $L[\varphi(x)]$ is a local functional of $\varphi(x)$. Equation (35.1) is a straightforward generalization of the equation (34.1) discussed in Chapter 34. We assume that the stochastic noise field $\nu(t, x)$ has the Gaussian local distribution (Gaussian white noise), with constant width Ω ,

$$[d\rho(\nu)] = [d\nu] \exp \left[- \int d^d x dt \nu^2(t, x) / 2\Omega \right]. \quad (35.2)$$

Equivalently, the noise can be characterized by its one- and two-point functions,

$$\langle \nu(t, x) \rangle_\nu = 0, \quad \langle \nu(t, x) \nu(t', x') \rangle_\nu = \Omega \delta^{(d)}(x - x') \delta(t - t'). \quad (35.3)$$

We will discuss the properties of equation (35.1) with the noise (35.2). This is a simple example, but many results apply to more general equations and Gaussian distributions of noise (see, for example, Sections 35.6, 35.7).

The FP equation. Given the noise (35.2), and some initial conditions for the field $\varphi(t, x)$, the Langevin equation generates the time-dependent field distribution

$$P(t, \varphi) = \langle \delta(\varphi(t, x) - \varphi(x)) \rangle_\nu. \quad (35.4)$$

The derivation of Section 34.1.1 can immediately be generalized, and yields the FP equation, field-theory extension of equation (34.18),

$$\dot{P}(t, \varphi) = -\mathcal{H}_{\text{FP}} P(t, \varphi), \quad (35.5)$$

where \mathcal{H}_{FP} , the FP Hamiltonian, a functional differential operator, is given by

$$\mathcal{H}_{\text{FP}} \left(\varphi, \frac{\delta}{\delta \varphi} \right) = -\frac{1}{2} \int d^d x \frac{\delta}{\delta \varphi(x)} \left[\Omega \frac{\delta}{\delta \varphi(x)} + L(\varphi(x)) \right]. \quad (35.6)$$

The dissipative Langevin equation. In Section 35.4, a special case is discussed more thoroughly—the purely dissipative Langevin equation, which corresponds to the choice

$$L[\varphi(x)] = \Omega \frac{\delta \mathcal{A}}{\delta \varphi(x)}, \quad (35.7)$$

and generalizes equation (34.34) (Section 34.4). The functional $\mathcal{A}(\varphi)$ is a static (time-independent) Euclidean action for the scalar field φ , for example,

$$\mathcal{A}(\varphi) = \int d^d x \left[\frac{1}{2} (\nabla \varphi(x))^2 + V(\varphi(x)) \right],$$

and the equilibrium distribution corresponds to the static distribution $e^{-\mathcal{A}(\varphi)}$.

General equation. In Chapter 34, we have already discussed these equations for a finite number of degrees of freedom. All arguments based on purely algebraic identities can be generalized to field theory. Here, we discuss new problems which arise, because, as we will show, the Langevin equation, in general, requires renormalizations. In the case (35.7), when the action $\mathcal{A}(\varphi)$ is renormalizable, we expect that the renormalizations of the static theory, together with a time-scale renormalization, render the correlation functions generated by the Langevin equation finite. To discuss the problem, we have to set up a formalism more directly amenable to the ordinary methods of QFT. This can be done by constructing a field integral representation of the time-dependent φ -field correlation functions in terms of an associated local action, the *dynamic action*.

35.2 Time-dependent correlation functions and equilibrium

We construct a field integral representation for the generating functional $\mathcal{Z}(J)$ of dynamic correlation functions of the field $\varphi(t, x)$ solution of equation (35.1), and then discuss, in a special example, the relation between ST symmetry and equilibrium properties.

35.2.1 Dynamic action

The generating functional $\mathcal{Z}(J)$ of dynamic correlation functions of the field $\varphi(t, x)$ solution of equation (35.1) is given by the noise expectation value

$$\begin{aligned}\mathcal{Z}(J) &= \left\langle \exp \left[\int d^d x dt J(t, x) \varphi(t, x) \right] \right\rangle_{\nu}, \\ &= \int [d\nu] \exp \left[- \int d^d x dt \left(\frac{1}{2\Omega} \nu^2(t, x) - J(t, x) \varphi(t, x) \right) \right].\end{aligned}\quad (35.8)$$

To impose equation (35.1), which we symbolically denote as

$$E(\varphi, \nu; t, x) \equiv \dot{\varphi}(t, x) + \frac{1}{2} L[\varphi(t, x)] - \nu(t, x) = 0, \quad (35.9)$$

following the method explained in Section 34.7.1, we insert the identity

$$\int [d\varphi] \det M \prod_{t,x} \delta[E(\varphi, \nu; t, x)] = 1, \quad (35.10)$$

where M is the differential operator,

$$M(t, x; t', x') = \frac{\delta E(\varphi, \nu; t, x)}{\delta \varphi(t', x')} = \left(\frac{\partial}{\partial t} + \frac{1}{2} \frac{\delta L}{\delta \varphi(t, x)} \right) \delta(t - t') \delta^{(d)}(x - x'). \quad (35.11)$$

into expression (35.8). We then find

$$\mathcal{Z}(J) = \int [d\nu] [d\varphi] \det M \delta(E(\varphi, \nu)) \exp \left\{ \int d^d x dt \left[J(t, x) \varphi(t, x) - \frac{1}{2\Omega} \nu^2(t, x) \right] \right\}.$$

The δ -function can be used to integrate over the noise ν . One obtains

$$\mathcal{Z}(J) = \int [d\varphi] \det M \exp \left[- \int d^d x dt \left(\frac{1}{2\Omega} E^2(\varphi, 0; t, x) - J(t, x) \varphi(t, x) \right) \right]. \quad (35.12)$$

For a system with a discrete set of degrees of freedom (a $d = 0$ dimensional or a lattice regularized field theory), the determinant can be calculated using the identity

$$\det M \propto \exp \text{tr} \ln \left[1 + \left(\frac{\partial}{\partial t} \right)^{-1} \frac{1}{2} \frac{\delta L}{\delta \varphi} \right]. \quad (35.13)$$

As a consequence of the causality of the Langevin equation, the inverse of the operator $(\partial/\partial t)\delta(t - t')$ is the kernel $\theta(t - t')$ ($\theta(t)$ is the Heaviside step function). In an expansion in powers of L , all terms thus vanish when one takes the trace, except the first one that yields

$$\det M \propto \exp \left\{ \frac{\theta(0)}{2} \int dt d^d x \left. \frac{\delta L[\varphi(t, x)]}{\delta \varphi(x', t)} \right|_{x' = x} \right\}. \quad (35.14)$$

For the undefined quantity $\theta(0)$, we choose $\theta(0) = 1/2$, a choice symmetric in time, for reasons we have already explained in Section 34.6. The final expression then formally reads [354]

$$\mathcal{Z}(J) = \int [d\varphi] \exp \left[-\mathcal{S}(\varphi) + \int d^d x dt J(t, x) \varphi(t, x) \right], \quad (35.15)$$

$$\mathcal{S}(\varphi) = \frac{1}{2\Omega} \int d^d x dt \left\{ \dot{\varphi}(t, x) + \frac{1}{2} L([\varphi(t, x)]) \right\}^2 - \frac{1}{4} \int dt d^d x \left. \frac{\delta L[\varphi(t, x)]}{\delta \varphi(x', t)} \right|_{x' = x}. \quad (35.16)$$

The dynamic action (35.16) is the generalization to field theory of the action (34.57) and can thus be directly obtained by expressing the solution $P(t, \varphi)$ of the FP equation (35.5) as a field integral.

When the force in the Langevin equation derives from a static action (equation (35.7)), the term linear in $\dot{\varphi}$ in the action \mathcal{S} is a total time derivative and contributes only to boundary terms.

35.2.2 Slavov–Taylor symmetry and equilibrium correlation functions

The Slavov–Taylor (ST) symmetry discussed in Section 26.1 implies identities for correlation functions.

Here, we consider only the example (35.7). Then, the substitution into equation (35.5)

$$P = e^{-\mathcal{A}/2} \tilde{P}$$

transforms the Hamiltonian (35.6) into a Hermitian positive Hamiltonian. This property makes it possible to prove detailed balance (see Section 34.4), and the convergence of P at time $+\infty$ towards the static distribution $e^{-\mathcal{A}}$ (provided the latter is normalizable).

It is at the origin of the concept of stochastic quantization: the use of the Langevin equation to generate, at equilibrium, the static distribution $e^{-\mathcal{A}}$, as explained in Section 21.17.

Another proof relies on the ST symmetry (Section 26.1). The identity (35.10) implies that $\det M[d\varphi]$, where M is the operator,

$$M(t, x; x', t') = \left(\frac{\partial}{\partial t} + \frac{\Omega}{2} \frac{\delta^2 \mathcal{A}}{\delta \varphi(t, x) \delta \varphi(t', x')} \right) \delta(t - t') \delta^{(d)}(x - x'), \quad (35.17)$$

is the invariant measure for a group of non-linear φ transformations that translate $\nu(t, x)$ by a function $\mu(t, x)$ (Section 26.1.2). For $\mu(t, x)$ infinitesimal, the variation of φ is (see equation (26.5))

$$\delta \varphi(t, x) = \int dt' d^d x' M^{-1}(t, x; x', t') \mu(x', t'). \quad (35.18)$$

In an infinitesimal change of variables of the form of a transformation (35.18), the variations of the action and the source term in the field integral (35.12) are

$$\begin{aligned} \delta \left\{ \frac{1}{2\Omega} \int d^d x dt \left(\dot{\varphi} + \frac{1}{2}\Omega \frac{\delta \mathcal{A}}{\delta \varphi} \right)^2 \right\} &= \frac{1}{\Omega} \int d^d x dt \left[\dot{\varphi}(t, x) + \frac{1}{2}\Omega \frac{\delta \mathcal{A}}{\delta \varphi(t, x)} \right] \mu(t, x), \\ \delta \left[\int d^d x dt J(t, x) \varphi(t, x) \right] &= \int dt d^d x dt' d^d x' J(t, x) M^{-1}(x', t'; t, x) \mu(x', t'). \end{aligned}$$

The invariance of the field integral under changes of variables leads to the identity

$$\left[\frac{1}{\Omega} \frac{\partial}{\partial t} \frac{\delta}{\delta J(t, x)} + \frac{1}{2} \frac{\delta \mathcal{A}}{\delta \varphi(t, x)} \left(\frac{\delta}{\delta J} \right) \right] \mathcal{Z}(J) = \int dx' dt' J(x', t') \mathcal{Z}_M(x', t'; t, x), \quad (35.19)$$

with the definition,

$$\mathcal{Z}_M(t, x; x', t'; J) = M^{-1}(x', t'; t, x; \delta/\delta J) \mathcal{Z}(J), \quad (35.20)$$

or, equivalently,

$$\left\{ \frac{\partial}{\partial t} + \frac{\Omega}{2} \frac{\delta^2 \mathcal{A}}{[\delta \varphi(t, x)]^2} \left(\frac{\delta}{\delta J} \right) \right\} \mathcal{Z}_M(t, x; x', t'; J) = \delta^{(d)}(x - x') \delta(t - t') \mathcal{Z}(J). \quad (35.21)$$

Equation (35.19), which can also be derived directly from the Langevin equation, implies that the large time limit of equal-time correlation functions satisfies the field equations of the static theory.

To show it, we choose a special source of the form

$$J(t, x') = J(x)\delta(t' - \tau) \quad (35.22)$$

and set $t = \tau$ in equation (35.19). Then the equation involves only $\mathcal{Z}_M(x', \tau; x, \tau; J)$. In the expansion of \mathcal{Z}_M in powers of \mathcal{A} , due to the $\theta(t)$ step function, only the first term survives and, with $\theta(0) = \frac{1}{2}$, \mathcal{Z}_M becomes

$$\mathcal{Z}_M(x', \tau; x, \tau; J) = \frac{1}{2} \delta^{(d)}(x - x') \mathcal{Z}(J). \quad (35.23)$$

We have again used the causal properties of the Langevin equation. Also, it can be checked in perturbation theory that $\mathcal{Z}(J)$ has a limit, and that for a Langevin equation of type (35.7) the term

$$\left. \frac{\partial}{\partial t} \frac{\delta}{\delta J(t, x)} \mathcal{Z}(J) \right|_{t=\tau}$$

vanishes at large time. Then, as a consequence of equations (35.19) and (35.23), the limiting functional $\mathcal{Z}(J)$ satisfies

$$\frac{\delta \mathcal{A}}{\delta \varphi(x)} \left(\frac{\delta}{\delta J} \right) \mathcal{Z}(J) = J(x) \mathcal{Z}(J), \quad (35.24)$$

which implies

$$\mathcal{Z}(J) = \int [d\varphi(x)] \exp \left[-\mathcal{A}(\varphi) + \int d^d x J(x) \varphi(x) \right]. \quad (35.25)$$

From these considerations, we conclude that some form of equations (35.19) and (35.20) must be used in the proof of the renormalizability of the dynamic theory, that is, in the proof that, after renormalization the structure of the Langevin equation is preserved, and that, for time $t \rightarrow \infty$, the equal-time correlation functions have the correlation functions of the renormalized static theory corresponding to $\mathcal{A}(\varphi)$ as limits.

35.3 Renormalization and BRST symmetry

Divergences and the problem of renormalization. In dimension $d > 0$, the expression (35.16) is not defined when $L(\varphi)$ is a local functional, because the contribution of the determinant is formally proportional to ' $\delta^{(d)}(0)$ ':

$$\ln \det M \propto \int d^d x \left. \frac{\delta L[\varphi(t, x)]}{\delta \varphi(x', t)} \right|_{x'=x} \propto \cdot \delta^{(d)}(0)'.$$

We encounter the same problem in the discussion of homogeneous spaces in Chapters 19 and 29, where a divergent measure term also appears. The determinant thus has to be regularized. We have two options.

(i) By keeping this divergent term in some regularized form, it makes it possible to define the theory beyond perturbation theory and understand its geometric structure. This can be achieved with, for example, a space-lattice regularization.

(ii) With dimensional regularization, terms like $\delta^{(d)}(0)$ vanish and, therefore, the determinant, which presumably cancels other perturbative divergences, can be completely omitted [360]. However, this requires a proof that the cancellation really occurs.

With this assumption, the expression (35.15) can be used in practical perturbative calculations. However, this property also strongly suggests that the dynamic theory is not completely determined by its perturbative expansion.

35.3.1 Power counting analysis

As an example, we consider the dissipative Langevin equation (equation (35.7)) derived from the four-dimensional static action,

$$\mathcal{A}(\varphi) = \int d^4x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} m^2 \varphi^2(x) + (g/4!) \varphi^4(x) \right]. \quad (35.26)$$

From the dynamic action (35.16), we calculate the propagator Δ of the $\varphi(t, x)$ field. After Fourier transformation, as a function of ω and \mathbf{k} , the Fourier variables associated with time and space, respectively, the propagator reads

$$\tilde{\Delta}(\mathbf{k}, \omega) = \frac{\Omega}{\omega^2 + \frac{1}{4}\Omega^2(k^2 + m^2)^2}. \quad (35.27)$$

In the standard power-counting analysis, as presented in Chapter 8, it is generally assumed that the propagator is, in the Fourier representation and when all arguments become large, a homogeneous function. This is not the case here. The propagator has a homogeneous asymptotic behaviour only if the frequency ω is scaled as a momentum squared, or equivalently the time t as a distance squared (the Brownian motion behaviour). Then, the canonical (engineering) dimensions $[\varphi]$ of the field φ in the static and dynamic theories coincide:

$$[\varphi] = 1. \quad (35.28)$$

The interactions generated by $(\delta\mathcal{A}/\delta\varphi)^2$, with

$$\frac{\delta\mathcal{A}}{\delta\varphi(x)} = -\nabla^2\varphi(x) + m^2\varphi(x) + g\varphi^3(x)/6 \quad (35.29)$$

are of dimension 4 and 6. General renormalization theory implies only that the renormalized action is the most general local functional of canonical dimension 6, that is, containing all vertices of non-positive dimensions. In general, such a functional depends on more parameters than the bare action (35.16); in particular, it is not a square of a local functional, and can no longer be derived from a Langevin equation by the algebraic transformations (35.8–35.16). Therefore, to prove that the Langevin equation can be renormalized, we have to derive identities satisfied by correlation functions, which imply relations between counter-terms, and ensure that the structure of the action (35.16) is preserved by renormalization.

35.3.2 BRST symmetry

We consider the special form of the Langevin equation (35.1),

$$E(\varphi; t, x) \equiv \dot{\varphi}(t, x) + \frac{1}{2} [-\nabla^2\varphi(t, x) + V(\varphi(t, x))] - \nu(t, x) = 0, \quad (35.30)$$

in which $V(\varphi)$ is a polynomial in φ , and $\nu(t, x)$ the Gaussian noise with distribution (35.2). For notational simplicity, we have written the equation for a one-component field, but all arguments immediately generalize to a field φ with several components. We consider equation (35.30) (or equation (35.9)) as a constraint equation in the sense of equation (26.56). The dynamic action corresponding to equation (35.30) can be written as (see equation (34.69))

$$\mathcal{S}(\varphi, \bar{\varphi}, c, \bar{c}) = \int dt d^d x \left[-\frac{1}{2} \Omega \bar{\varphi}^2(t, x) + \bar{\varphi}(t, x) E(\varphi; t, x) - c(t, x) M \bar{c}(t, x) \right], \quad (35.31)$$

where $\bar{\varphi}$ is an imaginary scalar field, c, \bar{c} are spinless fermion fields, and M is the differential operator (35.11), here,

$$M(t, x; t', x') = \left[\frac{\partial}{\partial t} - \frac{1}{2} \nabla_x^2 + \frac{1}{2} \frac{\partial V[\varphi(t, x)]}{\partial \varphi} \right] \delta(t - t') \delta^{(d)}(x - x'). \quad (35.32)$$

In this form, the dynamic action (35.31) is BRST symmetric (see also equations (34.70)). We have seen that, in the case of the Langevin equation, the determinant $\det M$ can be formally calculated in such a way that the final dynamic action is in direct correspondence with the FP equation. However, to exhibit more clearly the general algebraic structure, it is useful to keep it in the form of a Grassmann integral.

35.3.3 Power counting and renormalization

In what follows, we denote the dimension of a quantity Q by $[Q]$, in the sense of power counting.

From the viewpoint of power counting, following the argument in Section 35.3.1, for any equation of the form (35.30), one has to assign to frequencies ω the canonical dimension of momentum squared, that is, 2 and thus to time the dimension -2 . The dynamic action density (35.31) then has dimension $d + 2$ since the integration measure here is $dt d^d x$. The dimensions $[\bar{\varphi}]$, $[\varphi]$, $[\bar{c}]$, and $[c]$ of the fields follow,

$$[\bar{\varphi}] = \frac{1}{2}(d + 2), \quad [\varphi] = \frac{1}{2}(d - 2), \quad [\bar{c}] + [c] = d. \quad (35.33)$$

The theory is thus renormalizable by power counting in dimension d if the dimension of $V(\varphi)$ is

$$[V(\varphi)] = \frac{1}{2}(d + 2).$$

The dimension of M is then $[M] = [V(\varphi)] - \frac{1}{2}(d - 2) = 2$. If the dimension d satisfies

$$2([\bar{c}] + [c]) = 2d > d + 2,$$

that is, if the dimension d of space is larger than 2, the renormalized action S_r is at most quadratic in the fermion fields.

Moreover, if the dimension d is larger than 2, the sum of dimensions $[\bar{c}] + [c] + [\bar{\varphi}]$ satisfies

$$[\bar{c}] + [c] + [\bar{\varphi}] = \frac{3}{2}d + 1 > d + 2,$$

and, therefore, M_r , the renormalized coefficient of $c\bar{c}$, cannot depend on $\bar{\varphi}$.

BRST symmetry being linear in the fields, it is simple to prove that the symmetry can be preserved by the renormalization process. In the parametrization (26.54) in terms of a Grassmann coordinate θ , this corresponds to a preservation of the translation symmetry $\theta \mapsto \theta + \varepsilon$. The structure of the dynamic action (35.31) is thus preserved by the renormalization. It follows that it can be derived from a renormalized Langevin equation.

We have thus shown that, if the canonical dimension of $V(\varphi)$ satisfies $[V] \leq \frac{1}{2}(d+2)$ and the dimension d is larger than 2, the Langevin equation (35.30) with the Gaussian noise (35.2) can be renormalized. Moreover, since the renormalized action remains quadratic in $\bar{\varphi}$, the renormalized noise is still Gaussian. Clearly, this result can be immediately generalized to several component fields.

The analysis can also be extended to more complicated cases, for example, to models with different noise two-point functions (see Chapter 36).

In the special case (35.7), more precise results can be obtained, which we describe in Section 35.4 (for a one-component field equation, the drift force in equation (35.30) can always be written as the variation of an action as in equation (35.7)).

35.4 Dissipative Langevin equation and supersymmetry

In previous sections, we have explained that, quite generally, the dynamic action associated to a Langevin equation has a BRST symmetry. This symmetry and its consequences in the form of WT identities have been used to prove that under some general conditions the structure of the Langevin equation is stable under renormalization.

A particular class of Langevin equations is of special interest: purely dissipative equations with Gaussian noise, in which the drift force derives from a static action and, thus is conservative (corresponding to stochastic processes with detailed balance).

We have shown in Section 34.7.2 that the dynamic action then displays the simplest example of a *supersymmetry* [356].

35.4.1 Supersymmetry

We introduce the superfields, generalizations of the superpaths (34.73)

$$\phi(t, x; \bar{\theta}, \theta) = \varphi(t, x) + \theta \bar{c}(t, x) + c(t, x) \bar{\theta} + \theta \bar{\theta} \bar{\varphi}(t, x). \quad (35.34)$$

One can verify that the action (35.31), corresponding to the Langevin equation

$$\dot{\varphi}(t, x) = -\frac{\Omega}{2} \frac{\delta \mathcal{A}(\varphi(t, x))}{\delta \varphi(x)} + \nu(t, x), \quad (35.35)$$

with the noise Gaussian distribution (35.2),

$$[d\rho(\nu)] = [d\nu] \exp \left[- \int d^d x dt \nu^2(t, x) / 2\Omega \right].$$

expressed in terms of the superfields (35.34), can be rewritten as

$$\mathcal{S}(\phi) = \int d\bar{\theta} d\theta dt \left[\frac{2}{\Omega} \int d^d x \bar{D}\phi(t, x; \bar{\theta}, \theta) D\phi(t, x; \bar{\theta}, \theta) + \mathcal{A}(\phi) \right], \quad (35.36)$$

(we have rescaled equation (35.35) by a factor $2/\Omega$) where $\mathcal{A}(\phi)$ means that, in the action density, we have replaced $\varphi(x)$ by $\phi(t, x; \bar{\theta}, \theta)$. In the action, we have introduced the *covariant derivatives*, with respect to supersymmetry (Section 34.7.2)

$$\bar{D} = \frac{\partial}{\partial \bar{\theta}}, \quad D = \frac{\partial}{\partial \theta} - \bar{\theta} \frac{\partial}{\partial t}. \quad (35.37)$$

They satisfy the anticommutation relations

$$D^2 = \bar{D}^2 = 0, \quad D\bar{D} + \bar{D}D = -\frac{\partial}{\partial t}. \quad (35.38)$$

The two other Grassmann-type differential operators

$$Q = \frac{\partial}{\partial \theta}, \quad \bar{Q} = \frac{\partial}{\partial \bar{\theta}} + \theta \frac{\partial}{\partial t}, \quad (35.39)$$

are supersymmetry generators, and *anticommute* with D and \bar{D} . They satisfy

$$Q^2 = \bar{Q}^2 = 0, \quad Q\bar{Q} + \bar{Q}Q = \frac{\partial}{\partial t}. \quad (35.40)$$

The pair Q, \bar{Q} provides the simplest example of generators of supersymmetry (see, for example, Sections A26.4 or 34.7.2).

This supersymmetry is directly related with the property that the corresponding FP Hamiltonian (35.5) is then equivalent to a positive Hamiltonian of the form (34.39).

35.4.2 WT identities

The first BRST symmetry simply implies that correlation functions are invariant under a translation of the coordinate θ . The transformation generated by \bar{Q} has a slightly more complicated form. It leads for the generating functional of connected correlation functions $\mathcal{W}(J)$ to the equation

$$\int dt d^d x d\bar{\theta} d\theta \bar{Q} J(t, x, \theta, \bar{\theta}) \frac{\delta \mathcal{W}}{\delta J(t, x, \theta, \bar{\theta})} = 0.$$

Connected correlation functions $W^{(n)}(t_i, x_i, \theta_i, \bar{\theta}_i)$, and vertex functions $\Gamma^{(n)}(t_i, x_i, \theta_i, \bar{\theta}_i)$, thus satisfy the WT identities,

$$\bar{Q} W^{(n)}(t_i, x_i, \theta_i, \bar{\theta}_i) = 0, \quad \bar{Q} \Gamma^{(n)}(t_i, x_i, \theta_i, \bar{\theta}_i) = 0, \quad (35.41)$$

where \bar{Q} is represented by

$$\bar{Q} = \sum_{k=1}^n \left(\frac{\partial}{\partial \bar{\theta}_k} + \theta_k \frac{\partial}{\partial t_k} \right). \quad (35.42)$$

After Fourier transformation over time, the operator \bar{Q} takes the form

$$\bar{Q} = \sum_{k=1}^n \left(\frac{\partial}{\partial \bar{\theta}_k} - i\omega_k \theta_k \right). \quad (35.43)$$

Example: Two-point functions. We now explore the implications of WT identities for a two-point function. As the relations (35.39, 35.40) show, supersymmetry implies invariance by translation in time and in θ . Therefore, any two-point function has the general form (we omit the space dependence, which plays no role)

$$\begin{aligned} W^{(2)}(t_1, \theta_1, \bar{\theta}_1; t_2, \theta_2, \bar{\theta}_2) &= A(t_1 - t_2) \\ &\quad + (\theta_1 - \theta_2) [(\bar{\theta}_1 + \bar{\theta}_2)B(t_1 - t_2) + (\bar{\theta}_1 - \bar{\theta}_2)C(t_1 - t_2)]. \end{aligned}$$

The WT identity (35.41) then implies

$$2B(t) = \frac{\partial A}{\partial t}.$$

The WT identity does not determine the function C . An additional constraint comes from the causality of the Langevin equation, which plays an important role. For the two-point function, causality implies that the coefficient of $\theta_1 \bar{\theta}_2$ vanishes for $t_1 < t_2$, and the coefficient of $\theta_2 \bar{\theta}_1$ for $t_2 < t_1$. The last function is thus determined, up to a possible distribution localized at $t_1 = t_2$. One finds

$$2C(t) = -\text{sgn}(t) \frac{\partial A}{\partial t}, \quad (35.44)$$

where $\text{sgn}(t)$ is the sign of t . This leads to the remarkable form

$$\begin{aligned} W^{(2)}(t_1, \theta_1, \bar{\theta}_1; t_2, \theta_2, \bar{\theta}_2) &= \left\{ 1 + \frac{1}{2}(\theta_1 - \theta_2) [\bar{\theta}_1 + \bar{\theta}_2 - (\bar{\theta}_1 - \bar{\theta}_2) \text{sgn}(t_1 - t_2)] \frac{\partial}{\partial t_1} \right\} A(t_1 - t_2). \quad (35.45) \end{aligned}$$

35.4.3 Renormalization of the dissipative Langevin equation

In the special case of the supersymmetric dynamic action (35.36), a comparison between the two explicit quadratic terms in ϕ of the action yields the relation between dimensions,

$$[t] - [\theta] - [\bar{\theta}] = 0 \Rightarrow [dt] + [d\theta] + [d\bar{\theta}] = 0. \quad (35.46)$$

(We recall that, since integration and differentiation over anticommuting variables are identical operations, the dimension of $d\theta$ is $-[\theta]$.)

Therefore, the term proportional to $\mathcal{A}(\phi)$ in the dynamic action and the static action have the same canonical dimension: the power counting in the static and the dynamic theory is the same, and the dynamic theory is always renormalizable in the same space dimension d as the static theory. Note that equation (35.46) also implies

$$2[\phi] = d + [t],$$

an equation that relates the dimensions of field and time.

In superfield notation, the most general form of the renormalized action \mathcal{S}_r , consistent with the results derived in Section 35.3.2 has the form

$$\mathcal{S}_r = \int d^d x dt d\bar{\theta} d\theta \left[\frac{2}{\Omega} \frac{\partial \phi}{\partial \bar{\theta}} \left(\mathcal{Z}' \frac{\partial \phi}{\partial \theta} - \mathcal{Z} \bar{\theta} \frac{\partial \phi}{\partial t} \right) + \bar{\theta} \frac{\partial \phi}{\partial \bar{\theta}} L(\phi) \right]. \quad (35.47)$$

The transformations generated by Q and \bar{Q} (equations (35.39)) are linearly represented on the fields ϕ and, therefore, the renormalized action remains supersymmetric. Performing the transformation generated by \bar{Q} , and expressing the invariance of the action, one obtains two equations. Identifying the coefficient of $(\partial \phi / \partial \bar{\theta})(\partial \phi / \partial t)$, one infers

$$\mathcal{Z}' = \mathcal{Z}. \quad (35.48)$$

The other equation comes from the variation of the last term of expression (35.47):

$$\int d\bar{\theta} \frac{\partial \phi}{\partial \bar{\theta}} L(\phi) = 0. \quad (35.49)$$

Thus $L(\phi)$ is a total derivative and can be written as

$$L(\phi) = \frac{\delta(\mathcal{A}_r)}{\delta \phi}. \quad (35.50)$$

Finally, an integration by parts over $\bar{\theta}$ of the last term in equation (35.47) yields the supersymmetric form of the renormalized action

$$\mathcal{S}_r(\phi) = \int d\bar{\theta} d\theta dt \left[\frac{2Z_t}{\Omega} \int d^d x \bar{D}\phi D\phi + \mathcal{A}_r(\phi) \right]. \quad (35.51)$$

After renormalization, the drift force in the Langevin equation is thus still of the form of the variation of an action.

Remark. Because the FP equation has static (time-independent) solutions which are not of the form (35.7), it is easy to construct bare Langevin equations which generate an equilibrium distribution characterized by a local static action, for which the dynamic action is not supersymmetric. Direct proofs that the renormalized equilibrium distribution still corresponds to a local static action have only been given in special cases.

35.5 Supersymmetry and equilibrium correlation functions

The supersymmetry of action (35.36) leads to a direct algebraic proof that the equal-time ϕ -field correlation functions converge at large times towards the static correlation functions corresponding to the action $\mathcal{A}(\phi)$ [361]. To simplify the notation, we consider the action (35.36), but the generalization to actions of the type examined in Sections 35.6 and 35.7 is simple. We assume that the initial conditions in the Langevin equation are given at time t' , and we calculate the equal-time correlation functions at later time t'' . We choose the source J associated with the field ϕ of the special form

$$J(t, x, \theta, \bar{\theta}) = J(x)\delta(t - t'')\delta(\theta)\delta(\bar{\theta}). \quad (35.52)$$

We consider the s -dependent dynamic action

$$\begin{aligned} \mathcal{S}(\phi, s) = & (1-s) \int_{t'}^{t''} d\bar{\theta} d\theta dt \left[\frac{2}{\Omega} \int d^d x \bar{D}\phi(t, x, \theta, \bar{\theta}) D\phi(t, x, \theta, \bar{\theta}) + \mathcal{A}(\phi) \right] \\ & + s\mathcal{A}(\phi(t'', \theta'' = \bar{\theta}'' = 0)). \end{aligned} \quad (35.53)$$

For $s = 0$, one recovers the dynamic action for time-dependent correlation functions relevant when all times belong to the interval $[t', t'']$, and for $s = 1$, because the source has the form (35.52), the action generates the static correlation functions. If we differentiate the connected correlation functions calculated with the action $\mathcal{S}(\phi, s)$ with respect to s , we generate the insertion of the operator

$$\begin{aligned} \frac{\partial}{\partial s} \mathcal{S}(\phi, s) = & - \int_{t'}^{t''} d\bar{\theta} d\theta dt \left[\frac{2}{\Omega} \int d^d x \bar{D}\phi(t, x, \theta, \bar{\theta}) D\phi(t, x, \theta, \bar{\theta}) + \mathcal{A}(\phi) \right] \\ & + \mathcal{A}(\phi(t'', \theta'' = \bar{\theta}'' = 0)). \end{aligned} \quad (35.54)$$

We first omit the breaking of supersymmetry due to the boundary condition at t' . Our arguments are thus only strictly correct in the case of a system already at equilibrium, that is, $t' = -\infty$. In this case, the insertion of the operator $\partial\mathcal{S}/\partial s$ in an equal-time correlation function generates a two-time correlation function. In Section 35.4.2, we have determined the most general two-time connected correlation functions satisfying the requirements of supersymmetry and causality. For $t'' > t$, and setting $\theta'' = \bar{\theta}'' = 0$, we infer from the generalization of equation (35.45) for two operators \mathcal{O}_1 and \mathcal{O}_2 ,

$$\langle \mathcal{O}_1(t)\mathcal{O}_2(t'') \rangle = \left(1 + \theta\bar{\theta} \frac{\partial}{\partial t} \right) A(t - t'') = A(t + \theta\bar{\theta} - t''). \quad (35.55)$$

Therefore, we can replace $\phi(t, \theta, \bar{\theta})$ by $\varphi(t + \theta\bar{\theta})$ in the operator insertion. The identity

$$\left(\frac{\partial}{\partial \theta} - \bar{\theta} \frac{\partial}{\partial t} \right) f(t + \theta\bar{\theta}) = 0,$$

then implies that the insertion of the first term in the right-hand side of equation (35.54) vanishes. We remain with the insertion of the operator

$$R(\phi) = - \int_{t'}^{t''} dt d\bar{\theta} d\theta \mathcal{A}[\varphi(t + \theta\bar{\theta})] + \mathcal{A}[\varphi(t'')]. \quad (35.56)$$

The integration over the variables θ and $\bar{\theta}$ generates the time derivative of the action. The last time integration then is immediate. The contribution coming from the upper bound of the integral cancels the second term in the right-hand side, and one remains with

$$R(\phi) = \mathcal{A}[\varphi(t')]. \quad (35.57)$$

Since $\varphi(t')$ is fixed, the insertion equals the correlation function itself multiplied by a factor independent of the correlation function. It thus corresponds to a change in the free energy or vacuum amplitude. Therefore, we have shown that the connected correlation functions are independent of s : the equal-time connected correlation functions are identical to the static correlation functions.

We now discuss the effect of breaking of supersymmetry due to the boundary condition at $t = t'$. To prove invariance of the action under transformation (34.80), one has to integrate by parts. Furthermore, since $\varphi(t')$ is fixed, one cannot perform the transformation for $t = t'$ and must, therefore, multiply the variations of the field in equation (34.80) by a function of time which is 1 everywhere, except close to t' . The result is that a supersymmetry transformation generates the insertion of an operator function of the field taken at $t = t'$. Due to cluster properties, connected correlation functions involving the insertion of such operators vanish in the large time-separation limit, $|t'' - t'| \rightarrow \infty$. This implies that equal-time correlation functions converge at large time towards static equilibrium correlation functions. In explicit calculations, we shall here always set the initial conditions in the Langevin equation at $t' = -\infty$. Then, at any time, the equal-time φ -field correlation functions are time independent and, therefore, equal to the static correlation functions.

35.6 Stochastic quantization of two-dimensional chiral models

From the analysis of Section 35.3.2, it follows that two-dimensional scalar field theories, which are special in the static case because the field is dimensionless, also have special dynamic properties. In particular, power counting allows quartic terms in the auxiliary fermion fields.

In Chapter 29, we have shown that two-dimensional models that depend on a finite number of parameters, and are strictly renormalizable are related to coset spaces G/H in which H is a subgroup of the group G . Of particular interest are the models defined on symmetric spaces discussed in Sections 29.4–29.6. Here, we describe Langevin equations for $G \times G/G$ chiral models, G being a simple unitary group, but the method can be generalized to other simple compact groups (the simplest example being the non-linear σ -model [362]).

A Langevin equation for chiral fields. In chiral models, the field $g(x)$ varies in a matrix representation $\mathcal{R}(G)$ of dimension N of a compact Lie group G . We have shown in Section 29.4 that a static action \mathcal{A} with only two derivatives has the form

$$\mathcal{A}(g) = -\frac{\beta}{2} \int d^d x \operatorname{tr} j_\mu^2(x), \quad (35.58)$$

where the G -current $j_\mu(x)$, given by

$$j_\mu(x) = g^{-1}(x) \partial_\mu g(x), \quad (35.59)$$

belongs to the Lie algebra $\mathcal{L}(G)$ of G , and β is a coupling constant, which plays the role of the inverse temperature in classical statistical field theory. The classical field equation $\nabla \cdot \mathbf{j}(x) = 0$ (equation (29.67)) expresses the current conservation.

A group covariant Langevin equation that leads to an equilibrium distribution corresponding to the static action (35.58), is (below the index 0 always refers to time)

$$j_0(t, x) = (\Omega\beta/2)\nabla \cdot \mathbf{j}(t, x) + g^\dagger(t, x)\nu(t, x) - \nu^\dagger(t, x)g(t, x), \quad (35.60)$$

where the current $j_0(t, x)$ is defined by

$$j_0(t, x) = g^{-1}(t, x)\partial_t g(t, x). \quad (35.61)$$

The noise $\nu(t, x)$ is a general complex $N \times N$ matrix with Gaussian probability distribution

$$[d\rho(\nu)] = [d\nu] \exp \left[-\frac{1}{2\Omega} \int d^d x dt \operatorname{tr}(\nu^\dagger(t, x)\nu(t, x)) \right]. \quad (35.62)$$

Note that $\nu(t, x)$, being a general complex matrix has more degrees of freedom than the current, which belongs to the Lie algebra and, which in our conventions is anti-Hermitian.

The dynamic action. To construct the dynamic action, we follow the same steps as in previous sections. However, since we have to integrate over $g(t, x)$ with a group-invariant measure, the expressions are slightly modified, as one expects from the analysis of Section 26.5. Similar expressions also appear in the quantization of non-Abelian gauge theories with gauge group G .

We introduce the covariant derivatives ∇_μ, ∇_0 associated with the currents j_μ, j_0 (for details, see Section 22.1):

$$\nabla_\mu = \partial_\mu + [j_\mu, \cdot] \quad \nabla_0 = \partial_t + [j_0, \cdot]. \quad (35.63)$$

In the infinitesimal space-time-dependent group transformation

$$g(t, x) \mapsto g(t, x)(1 + \ell(t, x)), \quad \ell(t, x) \in \mathcal{L}(G), \quad (35.64)$$

the currents transform like

$$j_\mu(t, x) \mapsto j_\mu(t, x) + \nabla_\mu \ell(t, x), \quad j_0(t, x) \mapsto j_0(t, x) + \nabla_0 \ell(t, x). \quad (35.65)$$

The analogue of the operator \mathbf{M} introduced in equation (35.32) is now defined, when acting on a field ℓ , as

$$\begin{aligned} [\mathbf{M}\ell](t, x) &= \nabla_0 \ell(t, x) - \frac{1}{2}\beta\Omega\partial_\mu\nabla_\mu \ell(t, x) + \ell(t, x)g^\dagger(t, x)\nu(t, x) \\ &\quad + \nu^\dagger(t, x)g(t, x)\ell(t, x). \end{aligned} \quad (35.66)$$

Introducing a field $\bar{\varphi}$ to impose the Langevin equation, and fermion fields \bar{c} and c to express $\det \mathbf{M}$, all belonging to the Lie algebra of G , we can write, in the normalization of equation (35.36), the dynamic action \mathcal{S} resulting from the integration over the noise as (using the identity $(\bar{c}c)^\dagger = -c\bar{c}$)

$$\begin{aligned} \mathcal{S}(g, \bar{\varphi}, c, \bar{c}) &= \int d^d x dt \frac{2}{\Omega} \operatorname{tr} \left[(\bar{\varphi}(t, x) + c(t, x)\bar{c}(t, x))(\bar{\varphi}(t, x) + \bar{c}(t, x)c(t, x)) \right. \\ &\quad \left. + (j_0(t, x) - \frac{1}{2}\Omega\beta\partial_\mu j_\mu(t, x))\bar{\varphi}(t, x) - c(t, x)(\nabla_0 - \frac{1}{2}\Omega\beta\partial_\mu\nabla_\mu)\bar{c}(t, x) \right]. \end{aligned} \quad (35.67)$$

Note the appearance of a quartic fermion term induced by the dependence of the noise term on the field $g(t, x)$ in equation (35.60). As one expects from the analysis of Section 26.3.1, the set of BRST transformations that leave the action invariant is now

$$\delta_{\text{BRST}} g(t, x) = \varepsilon \bar{c}(t, x)g(t, x), \quad \delta_{\text{BRST}} \bar{c}(t, x) = \varepsilon \bar{c}^2(t, x), \quad (35.68)$$

$$\delta_{\text{BRST}} c(t, x) = \varepsilon \bar{\varphi}(t, x), \quad \delta_{\text{BRST}} \bar{\varphi}(t, x) = 0, \quad (35.69)$$

(ε is a Grassmann constant). The BRST transformation of the current j_μ induced by equation (35.68) then becomes

$$\delta_{\text{BRST}} j_\mu(t, x) = \varepsilon \nabla_\mu \bar{c}(t, x). \quad (35.70)$$

The BRST transformations that appear in quantized non-Abelian gauge theories in Section 22.4.1 are identical to the transformations (35.68–35.70); the gauge field associated with the group G transforms as the current in equation (35.70).

We now show that the Langevin equation (35.60) is the natural generalization of equation (35.35). We introduce the superfield \mathcal{G} , formal element of the group G ,

$$\mathcal{G}(t, x, \theta, \bar{\theta}) = (1 + c(t, x)\bar{\theta} + \theta\bar{\theta}\bar{\varphi}(t, x))(1 + \theta\bar{c}(t, x))g(t, x). \quad (35.71)$$

The set of equations (35.68, 35.69) is then equivalent to

$$\delta_{\text{BRST}} \mathcal{G}(t, x, \theta, \bar{\theta}) = \varepsilon \frac{\partial}{\partial \theta} \mathcal{G}(t, x, \theta, \bar{\theta}). \quad (35.72)$$

From \mathcal{G} , we derive a set of associated currents:

$$J_\mu(t, x, \theta, \bar{\theta}) = [\mathcal{G}(t, x, \theta, \bar{\theta})]^{-1} \partial_\mu \mathcal{G}(t, x, \theta, \bar{\theta}), \quad (35.73)$$

and

$$J_{\bar{\theta}}(t, x, \theta, \bar{\theta}) = [\mathcal{G}(t, x, \theta, \bar{\theta})]^{-1} \bar{D}\mathcal{G}(t, x, \theta, \bar{\theta}), \quad (35.74)$$

$$J_\theta(t, x, \theta, \bar{\theta}) = [\mathcal{G}(t, x, \theta, \bar{\theta})]^{-1} D\mathcal{G}(t, x, \theta, \bar{\theta}), \quad (35.75)$$

where \bar{D}, D are the covariant derivatives (35.37).

The action (35.67) can then be written in a form analogous to expression (35.36) as

$$\mathcal{S} = - \int d^d x dt d\bar{\theta} d\theta \text{tr} \left(\frac{2}{\Omega} J_{\bar{\theta}}(t, x, \theta, \bar{\theta}) J_\theta(t, x, \theta, \bar{\theta}) + \frac{\beta}{2} J_\mu^2(t, x, \theta, \bar{\theta}) \right). \quad (35.76)$$

In this form, it is clear that \mathcal{S} is supersymmetric. Supersymmetry implies that equal-time $g(t, x)$ -field correlation functions converge at time $+\infty$ towards the correlation functions of the static action (35.58), justifying the choice of the Langevin equation (35.60). We show in the next section that equation (35.60) is a special example of a family of Langevin equations corresponding to two-dimensional models defined on Riemannian manifolds, which contain all models defined on homogeneous spaces.

The form of the renormalized action is then dictated by the structure of symmetric space $G \times G/G$ and supersymmetry (equation (34.79)): two renormalization constants are needed, the usual coupling constant renormalization of the static theory and again the time-scale renormalization. In addition, in general, the parametrization of the group elements is also renormalized.

35.7 Langevin equation and Riemannian manifolds

We now construct a Langevin equation on manifolds [363] by the following method: we consider smooth manifolds of dimension p , embedded in an Euclidean space of dimension N , and defined by a set of equations constraining the Euclidean coordinates χ_α , $\alpha = 1, \dots, N$:

$$E^s(\chi_\alpha) = 0, \quad s = p + 1, \dots, N. \quad (35.77)$$

We have already seen in Section 34.9 that it is impossible to define a Langevin equation that only depends on the geometry of a manifold.

To simplify calculations, we assume that we can solve locally these equations, and choose a set of coordinates χ^α such that the first p , which we denote by φ^i are independent, and the remaining $(N-p)$, which we denote by σ^s are functions of the independent components φ^i :

$$\chi_\alpha \equiv \{\varphi^i, \sigma^s\}, \quad \text{and} \quad E^s(\chi_\alpha) = 0 \Rightarrow \sigma^s = \sigma^s(\varphi). \quad (35.78)$$

The induced metric tensor g_{ij} on the manifold is then (see equation (A29.20))

$$g_{ij} = \delta_{ij} + \frac{\partial \sigma^s}{\partial \varphi_i} \frac{\partial \sigma^s}{\partial \varphi_j}. \quad (35.79)$$

To construct a Langevin equation on the manifold, we start from a Langevin equation of type (35.35) in the embedding Euclidean space,

$$\dot{\chi}_\alpha(t, x) = -\frac{\Omega}{2} \frac{\delta \mathcal{A}}{\delta \chi_\alpha(t, x)} + \nu_\alpha(t, x), \quad (35.80)$$

in which $x \in \mathbb{R}^d$, Ω is a positive constant, and $\nu_\alpha(t, x)$ is a Gaussian noise defined by

$$\langle \nu_\alpha(t, x) \rangle_\nu = 0, \quad \langle \nu_\alpha(t, x) \nu_\beta(t', x') \rangle_\nu = \Omega \delta(t - t') \delta^{(d)}(x - x') \delta_{\alpha\beta}. \quad (35.81)$$

As in the case of chiral fields, we are again naturally led to introduce a noise with more degrees of freedom than the field.

We project the equation onto the space tangent to the manifold at point $\chi_\alpha(t, x)$. For this purpose, we introduce some notation. In the manifold, the variations $\delta \chi_\alpha$ of the field χ_α are constrained by

$$\frac{\partial E^s}{\partial \chi_\alpha} \delta \chi_\alpha = 0. \quad (35.82)$$

We introduce an orthogonal basis $e_a^\alpha(\chi)$, with $1 \leq a \leq p$, for the variations of χ_α :

$$\frac{\partial E^s}{\partial \chi_\alpha} e_a^\alpha(\chi) = 0, \quad (35.83)$$

with the orthogonality conditions

$$e_a^\alpha(\chi) e_b^\alpha(\chi) = \delta_{ab}. \quad (35.84)$$

In terms of the independent coordinates φ^i , equation (35.83) can be written as

$$e_a^s(\varphi) - \partial_i \sigma^s(\varphi) e_a^i(\varphi) = 0, \quad (35.85)$$

and, therefore, equation (35.84) becomes, after elimination of the components e_a^s ,

$$e_a^i(\varphi) g_{ij}(\varphi) e_b^j(\varphi) = \delta_{ab}. \quad (35.86)$$

Equation (35.86) shows that the matrix e_a^i is the inverse of the vielbein (Section 28.5.1).

Projecting equation (35.80) onto the tangent plane, we now obtain the Langevin equation on the manifold:

$$\dot{\chi}_\alpha(t, x) = e_a^\alpha e_a^\beta \left(-\frac{1}{2} \frac{\delta \mathcal{A}(\chi)}{\delta \chi_\beta} + \nu_\beta \right). \quad (35.87)$$

In terms of the independent components φ^i , the equation reads

$$\begin{aligned} \dot{\varphi}^i(t, x) &= -\frac{\Omega}{2} e_a^i(\varphi(t, x)) \left(e_a^j(\varphi(t, x)) \frac{\delta \mathcal{A}}{\delta \varphi^j(t, x)} + \partial_j \sigma^s(\varphi(t, x)) e_a^j(\varphi(t, x)) \frac{\delta \mathcal{A}}{\delta \varphi^s(t, x)} \right) \\ &\quad + e_a^i(\varphi(t, x)) [e_a^j(\varphi(t, x)) \nu_j(t, x) + e_a^j(\varphi(t, x)) \partial_j \sigma^s(\varphi(t, x)) \nu_s(t, x)]. \end{aligned} \quad (35.88)$$

We observe that

$$\frac{\delta \mathcal{A}}{\delta \varphi^j}(\varphi, \sigma) + \partial_j \sigma^s(\varphi) \frac{\delta \mathcal{A}}{\delta \sigma^s} = \partial_j \mathcal{A}(\varphi, \sigma(\varphi)), \quad (35.89)$$

in which in the right-hand side $\partial_j \mathcal{A}$ now means total derivative of \mathcal{A} with respect to φ^j .

As a consequence of equation (35.86), the inverse metric tensor g^{ij} is given by

$$g^{ij}(\varphi) = e_a^i(\varphi) e_a^j(\varphi). \quad (35.90)$$

Equation (35.88) can thus be written as

$$\dot{\varphi}^i(t, x) = -\frac{1}{2} \Omega g^{ij}(\varphi(t, x)) \partial_j \mathcal{A}(\varphi(t, x)) + g^{ij}(\varphi(t, x)) t_j^\alpha(\varphi(t, x)) \nu_\alpha(t, x), \quad (35.91)$$

with the definition

$$t_i^\alpha \equiv \begin{cases} t_i^j = \delta_{ij}, \\ t_i^s = \partial_i \sigma^s. \end{cases} \quad (35.92)$$

Since the noise ν_α is multiplied by a function of φ , we have to specify the meaning of the product. We choose the covariant definition as in Section 34.9.

The quantities t_i^α satisfy

$$t_i^\alpha t_j^\alpha = g_{ij}, \quad \text{and} \quad \partial_i t_j^\alpha t_k^\alpha = g_{kl} \Gamma_{ij}^l, \quad (35.93)$$

in which Γ_{ij}^l is the Christoffel symbol (equation (28.43)). This equation can be rewritten in covariant form as

$$\nabla_i t_j^\alpha t_k^\alpha = 0, \quad (35.94)$$

in which ∇_i is the covariant derivative on the manifold.

Dynamic action. It is convenient to rewrite equation (35.91) as

$$g_{ij}(\varphi(t, x)) \dot{\varphi}^j(t, x) + \frac{1}{2} \Omega \partial_i \mathcal{A}(\varphi(t, x)) - t_i^\alpha(\varphi(t, x)) \nu_\alpha(t, x) = 0. \quad (35.95)$$

Introducing a Lagrange multiplier $\bar{\varphi}^i$ and fermion fields c^i and \bar{c}^i , we can express the corresponding dynamic action \mathcal{S} , before integration over the noise, as the sum of two contributions \mathcal{S}_0 and \mathcal{S}_1 :

$$\mathcal{S} = \mathcal{S}_0 + \mathcal{S}_1 \quad (35.96)$$

with

$$\mathcal{S}_0 = \int d^d x dt \left(\frac{1}{2\Omega} \nu_\alpha \nu_\alpha - \frac{2}{\Omega} t_i^\alpha \nu_\alpha \bar{\varphi}^i + \frac{2}{\Omega} c^i \partial_j t_i^\alpha \bar{c}^j \nu_\alpha \right), \quad (35.97)$$

$$\mathcal{S}_1 = \int d^d x dt \frac{2}{\Omega} \left[\bar{\varphi}^i (g_{ij} \dot{\varphi}^j + \frac{1}{2} \Omega \partial_i \mathcal{A}) + c^i g_{ij} \dot{\bar{c}}^j - c^i \partial_k g_{ij} \dot{\varphi}^j \bar{c}^k - \frac{1}{2} \Omega c^i \partial_i \partial_j \mathcal{A} \bar{c}^j \right]. \quad (35.98)$$

After an integration over the noise, and a short calculation, one finds

$$\mathcal{S} = \mathcal{S}_1 + \int d^d x dt \frac{2}{\Omega} (-\bar{\varphi}^i g_{ij} \bar{\varphi}^j + 2\bar{\varphi}^i g_{il} \Gamma_{jk}^l c^j \bar{c}^k + c^i \bar{c}^j c^k \bar{c}^l \partial_i \partial_j g_{kl}). \quad (35.99)$$

Introducing the superfield

$$\phi^i(t, x, \theta, \bar{\theta}) = \varphi^i(t, x) + \theta \bar{c}^i(t, x) + c^i(t, x) \bar{\theta} + \theta \bar{\theta} \bar{\varphi}^i(t, x), \quad (35.100)$$

and the covariant derivatives (35.37), one can rewrite the action as

$$\mathcal{S}(\phi) = \int dt d\bar{\theta} d\theta \left[\int d^d x \frac{2}{\Omega} \bar{D}\phi^i g_{ij}(\phi) D\phi^j + \mathcal{A}(\phi) \right]. \quad (35.101)$$

This supersymmetric form then implies that, at equilibrium, the equal-time field configurations are weighted by the measure $[d\rho(\varphi)]$:

$$[d\rho(\varphi)] = \prod_i [d\varphi^i] \sqrt{g(\varphi)} \exp [-\mathcal{A}(\varphi, \sigma(\varphi))]. \quad (35.102)$$

This was indeed the field distribution expected from equation (35.91). Note that, in contrast to the Langevin equation (35.91), the dynamic action (35.101) and, therefore, the correlation functions at equilibrium depend only on the geometry of the manifold. In the case of homogeneous spaces, the static action is simply

$$\mathcal{A}(\chi) = \frac{1}{2} \int (\nabla \chi_\alpha(x))^2 d^d x, \quad (35.103)$$

which, as a functional of φ , becomes on the manifold

$$\mathcal{A}(\varphi, \sigma(\varphi)) = \frac{1}{2} \int g_{ij}(\varphi(x)) \nabla \varphi^i(x) \nabla \varphi^j(x) d^d x. \quad (35.104)$$

One can verify that the expression (35.101) contains the expression (35.76) as a special example.

Finally, note that the field integral representation for the generating functional $\mathcal{Z}(J)$ can also be rewritten, introducing the superfield X in Euclidean coordinates

$$X_\alpha(t, x, \theta, \bar{\theta}) = \chi_\alpha(t, x) + \theta \bar{c}_\alpha(t, x) + c_\alpha(t, x) \bar{\theta} + \bar{\varphi}_\alpha(t, x) \theta \bar{\theta}, \quad (35.105)$$

in the form

$$\begin{aligned} \mathcal{Z}(J) &= \int [d\phi] \exp \left[-\mathcal{S}(\phi) + \int dx dt d\bar{\theta} d\theta J_i \phi^i \right] \\ &= \int [dX] \delta [E^s(X)] \exp \left(-\Sigma(X) + \int dx dt d\bar{\theta} d\theta J_i \phi^i \right), \end{aligned} \quad (35.106)$$

with, for $\Sigma(X)$, the expression

$$\Sigma(X) = \int d\bar{\theta} d\theta dt \left[\int d^d x \frac{2}{\Omega} \bar{D}X_\alpha D X_\alpha + \mathcal{A}(X) \right]. \quad (35.107)$$

These expressions emphasize the similarity between the properties of the static and the dynamic theory.

In the case of homogeneous spaces, the renormalization properties of the Langevin equation then follow from the general analysis of Chapter 29 and supersymmetry.

A35 The random field Ising model: Supersymmetry

It can be argued that the large-scale properties of an Ising model in a random magnetic field can be described by a stochastic field equation of the form

$$(-\nabla^2 + m^2)\varphi(x) + \frac{g}{3!}\varphi^3(x) = h(x), \quad (A35.1)$$

in which the field $\varphi(x)$ represents the Ising spin, and $h(x)$, the magnetic field, has the *quenched* Gaussian distribution (see Section A14.5) defined by (α is a positive constant),

$$\langle h(x)h(x') \rangle_h = \alpha d^d(x - x'), \quad \langle h(x) \rangle_h = 0. \quad (A35.2)$$

Equation (A35.1) is an example of a gradient equation of the form (26.66) and, therefore, leads to two BRST symmetries. Introducing the superfield

$$\phi(x, \bar{\theta}, \theta) = \varphi(x) + \theta\bar{c}(x) + \bar{c}(x)\theta + \theta\bar{\theta}\bar{\varphi}(x),$$

one can write the associated action $\mathcal{S}(\phi)$, after averaging over the magnetic field, as

$$\mathcal{S}(\phi) = \int d\bar{\theta} d\theta \left(\frac{\alpha}{2} \int d^d x \frac{\partial\phi}{\partial\bar{\theta}} \frac{\partial\phi}{\partial\theta} + \mathcal{A}(\phi) \right), \quad (A35.3)$$

with

$$\mathcal{A}(\varphi) = \int d^d x \left[\frac{1}{2} (\nabla\varphi(x))^2 + \frac{1}{2} m^2 \varphi^2(x) + \frac{g}{4!} \varphi^4(x) \right]. \quad (A35.4)$$

The action (A35.3) has a remarkable property: it has a large symmetry group: translations of $\bar{\theta}$ and θ as expected, but also ‘rotations’, which leave the line element

$$(ds)^2 = dx_\mu dx_\mu - (4/\alpha) d\bar{\theta} d\theta \quad (A35.5)$$

invariant. In addition to transformations internal to the $\{\bar{\theta}, \theta\}$ and x_μ spaces, one finds the two infinitesimal transformations, which we express in terms of the two sets of anti-commuting parameters, a_μ and \bar{a}_μ , elements of the Grassmann algebra as

$$\theta = \theta' + \alpha a_\mu x'_\mu, \quad x_\mu = x'_\mu - 2a_\mu \bar{\theta}', \quad (A35.6)$$

$$\bar{\theta} = \bar{\theta}' + \alpha \bar{a}_\mu x'_\mu, \quad x_\mu = x'_\mu - 2\theta' \bar{a}_\mu. \quad (A35.7)$$

The generators \bar{Q}_μ and Q_μ of the transformations (A35.6) and (A35.7), respectively, satisfy the anticommutation relations

$$\bar{Q}_\mu \bar{Q}_\nu + \bar{Q}_\nu \bar{Q}_\mu = -4\alpha \delta_{\mu\nu} \bar{\theta} \frac{\partial}{\partial \bar{\theta}}, \quad (A35.8)$$

$$Q_\mu Q_\nu + Q_\nu Q_\mu = 4\alpha \delta_{\mu\nu} \theta \frac{\partial}{\partial \theta}, \quad (A35.9)$$

$$\bar{Q}_\mu Q_\nu + Q_\nu \bar{Q}_\mu = 2\alpha \left[\delta_{\mu\nu} \left(\theta \frac{\partial}{\partial \theta} - \bar{\theta} \frac{\partial}{\partial \bar{\theta}} \right) + x_\mu \partial_\nu - x_\nu \partial_\mu \right]. \quad (A35.10)$$

One recognizes generators of internal transformations in the right-hand side, .

Using this symmetry, it is possible to prove a property of dimensional reduction for the action (A35.3) by a variant of the method of Section 35.5 [361, 364]. One shows that the correlation functions calculated in d dimensions with the action (A35.3) are the same as those calculated in $(d - 2)$ dimensions with the action $\mathcal{A}(\varphi(x))$. This result maps the random field Ising model in d dimensions onto the pure Ising model in $(d - 2)$ dimensions. Unfortunately, it has consequences which contradict physical intuition [365] and has been disproved [366]. A basic reason can be found in the starting point: equation (A35.1), in the region of interest, which is m^2 small, has certainly for some fields many solutions and, therefore, the whole method is not applicable without modifications. A similar problem called Gribov’s ambiguity arises in non-Abelian gauge theories.

36 Critical dynamics and renormalization group (RG)

In Chapters 14–18, we discuss the static equilibrium properties of critical statistical systems. Here, we study their time evolution in the critical domain (note that, from the point of view of particle physics, this time is completely non-physical and can be thought of as, for example, the computer time in Monte Carlo simulations).

Typical quantities of interest are relaxation rates towards equilibrium, time-dependent correlation functions, and transport coefficients.

The main motivation for such a study is that, in systems in which the dynamics is *local* (on short time scales, a modification of a dynamic variable has an influence only locally in space), when the correlation length becomes large, a large time scale emerges, which characterizes the rate of time evolution. This phenomenon called *critical slowing down* leads to universal behaviour and scaling laws for time-dependent quantities.

However, in contrast with the situation in static critical phenomena, there is no clean and systematic derivation of the dynamical equations governing the time evolution in the critical domain. One reason is that often the time evolution is influenced by conservation laws involving the order parameter, or other variables like energy, momentum, angular momentum, currents, and so on. Still, one can argue that, in the critical domain, the dynamics of all these quantities can be described by coupled Langevin equations of the type considered in Chapters 34 and 35.

We have already shown in Section 34.3.1 that the equilibrium distribution does not determine the driving force in the Langevin equation. Only the dissipative couplings, which are generated by the derivative of the equilibrium Hamiltonian, are related to the static properties. Indeed, the Langevin equation

$$\dot{\varphi}_i(t) = -\frac{1}{2}\beta \sum_j \Omega_{ij} \frac{\delta \mathcal{H}}{\delta \varphi_j} + F_i(\varphi(t)) + \nu_i(t), \quad (36.1)$$

(Ω_{ij} is a constant symmetric-invertible matrix) in which $\nu_i(t)$ is the usual Gaussian white noise:

$$\langle \nu_i(t) \rangle_\nu = 0, \quad \langle \nu_i(t) \nu_j(t') \rangle_\nu = \Omega_{ij} \delta(t - t'), \quad (36.2)$$

leads to the equilibrium distribution $e^{-\beta \mathcal{H}}$ if the ‘streaming’ term $F_i(\varphi)$ satisfies the conservation equation:

$$\frac{\partial}{\partial \varphi_i} \left[F_i(\varphi) e^{-\beta \mathcal{H}(\varphi)} \right] = 0, \quad (36.3)$$

as one can immediately verify with the corresponding Fokker–Planck (FP) equation (see, for example, equations (34.24)).

Equation (36.3) does not determine the streaming term. As a direct consequence, with each static universality class are associated an infinite number of dynamical universality classes, and critical dynamics has no longer the aesthetic simplicity of statics. In specific examples, one writes the simplest phenomenological evolution equation that has the required equilibrium properties, and that incorporates all the known physical conditions. Only a few typical examples, which appear in the classification of the review of Halperin and Hohenberg [367], are studied in this chapter.

They illustrate different physics requirement and, in a correlated way, different technical issues. The various techniques described in Chapter 35 are then useful to predict the RG properties, and to perform explicit perturbative calculations [368, 360]. Note that some of these techniques also apply to general local dynamical equations without reference to a possible equilibrium state.

Correlation and response functions. One can be interested in the critical relaxation towards equilibrium of time-dependent connected correlation functions (we omit the space dependence, and ν is the noise)

$$W_{i_1 \dots i_n}^{(n)}(t_1, \dots, t_n) = \langle \varphi_{i_1}(t_1) \cdots \varphi_{i_n}(t_n) \rangle_\nu, \quad (36.4)$$

and also, response functions, which characterize the response of the system to infinitesimal time-dependent perturbations. They can be generated by adding to the Hamiltonian $\mathcal{H}(\varphi(t))$ a source term linearly coupled to a perturbation $\mathcal{O}(\varphi(t))$,

$$\mathcal{H}(\varphi(t)) \mapsto \mathcal{H}(\varphi(t)) - \int dt h(t) \mathcal{O}(\varphi(t)). \quad (36.5)$$

The variation $R^{(n)}$ of the correlation function $W^{(n)}$ under an infinitesimal perturbation proportional to the function $\mathcal{O}(\varphi(t))$ then is given by

$$R_{i_1 \dots i_n}^{(n)}(t_0; t_1, \dots, t_n) = \frac{\delta}{\delta h(t_0)} \langle \varphi_{i_1}(t_1; h) \cdots \varphi_{i_n}(t_n; h) \rangle \Big|_{h=0}. \quad (36.6)$$

The causality of the Langevin equation implies

$$R^{(n)} = 0, \quad \text{for } t_i < t_0, \quad \forall 1 \leq i \leq n. \quad (36.7)$$

If one takes h constant, then the modification (36.5) of the Hamiltonian enables at equilibrium to generate $\mathcal{O}(\varphi)$ correlation functions. Therefore,

$$\lim_{T \rightarrow +\infty} \int dt R_{i_1, \dots, i_n}^{(n)}(t; T, T, \dots, T) = \langle \mathcal{O}(\varphi) \varphi_{i_1} \cdots \varphi_{i_n} \rangle, \quad (36.8)$$

in which the right-hand side now is an average taken with the equilibrium distribution $e^{-\beta \mathcal{H}(\varphi)}$. Higher order derivatives with respect to $h(t)$ generate generalized forms of the properties (36.7, 36.8).

36.1 Dissipative equation: RG equations in dimension $d = 4 - \varepsilon$

Notation. We will use the notation $\theta = (\bar{\theta}, \theta)$ for the pair of Grassmann coordinates $\theta, \bar{\theta}$.

We first discuss a purely dissipative dynamics (which thus satisfies detailed balance) and without conservation laws, the model A in the classification of Ref. [367]. We have described in Chapter 35 most of the techniques required for the derivation of RG equations for the dynamics [360, 285]. In Section A36.1.1, we calculate the dynamic RG function at two-loop order (for model A, see also Ref. [369]).

The N -vector model in dimension $d = 4 - \varepsilon$. The bare dissipative Langevin equation for the $O(N)$ -symmetric N -vector model has the form

$$\dot{\varphi}(t, x) = -\frac{\Omega_0}{2} \left[(-\nabla^2 + r_0) \varphi(t, x) + g_0 \frac{\Lambda^\varepsilon}{3!} \varphi(t, x) \varphi^2(t, x) \right] + \nu(t, x), \quad (36.9)$$

where ν is the Gaussian white noise characterized by

$$\langle \nu \rangle_\nu = 0, \quad \langle \nu_i(t, x) \nu_j(t', x') \rangle_\nu = \Omega_0 \delta(t - t') \delta^{(d)}(x - x'). \quad (36.10)$$

The equilibrium distribution is characterized by the Hamiltonian

$$\mathcal{H}(\boldsymbol{\varphi}) = \int d^d x \left[\frac{1}{2} (\nabla \boldsymbol{\varphi}(x))^2 + \frac{1}{2} r_0 \boldsymbol{\varphi}^2(x) + g_0 \frac{\Lambda^\varepsilon}{4!} (\boldsymbol{\varphi}^2(x))^2 \right]. \quad (36.11)$$

In terms of the superfield (see Section 35.4),

$$\phi(t, x, \boldsymbol{\theta}) = \boldsymbol{\varphi}(t, x) + \theta \bar{\mathbf{c}}(t, x) + \mathbf{c}(t, x) \bar{\theta} + \theta \bar{\theta} \boldsymbol{\lambda}(t, x), \quad (36.12)$$

the dynamic action $\mathcal{S}(\phi)$ takes the supersymmetric form

$$\mathcal{S}(\phi) = \int dt d\bar{\theta} d\theta \left[\int d^d x \frac{2}{\Omega_0} \bar{D}\phi(t, x, \boldsymbol{\theta}) \cdot D\phi(t, x, \boldsymbol{\theta}) + \mathcal{H}(\phi) \right], \quad (36.13)$$

(in $\mathcal{H}(\phi)$, $\boldsymbol{\varphi}(x)$ is replaced by $\phi(t, x, \boldsymbol{\theta})$) with

$$\bar{D} = \frac{\partial}{\partial \bar{\theta}}, \quad D = \frac{\partial}{\partial \theta} - \bar{\theta} \frac{\partial}{\partial t}. \quad (36.14)$$

36.1.1 Supersymmetry and the fluctuation-dissipation theorem

The addition to the static action of a perturbation proportional to the order parameter itself,

$$\mathcal{H}(\boldsymbol{\varphi}) \mapsto \mathcal{H}(\boldsymbol{\varphi}) - \int d^d x \mathbf{h}(x) \cdot \boldsymbol{\varphi}(x) \quad (36.15)$$

is equivalent, in the Langevin equation, to the translation of the noise (36.9):

$$\boldsymbol{\nu}(t, x) \mapsto \boldsymbol{\nu}(t, x) + \frac{1}{2} \Omega_0 \mathbf{h}(t, x). \quad (36.16)$$

Therefore, a differentiation with respect to $\mathbf{h}(t, x)$ generates $\boldsymbol{\lambda}$ -field correlation functions.

In Section 35.4.2, from supersymmetry, we have derived Ward–Takahashi (WT) identities satisfied by the connected time-dependent correlation functions (equation (35.41)):

$$\sum_{j=1}^n \left(\frac{\partial}{\partial \bar{\theta}_j} + \theta_j \frac{\partial}{\partial t_j} \right) \widetilde{W}^{(n)}(t_i, k_i, \boldsymbol{\theta}_i) = 0. \quad (36.17)$$

We have then obtained the general form of a two-point function consistent with supersymmetry and causality (equation (35.45)):

$$\widetilde{W}^{(2)}(t, k, \boldsymbol{\theta}; t', 0, \boldsymbol{\theta}') = \left\{ 1 + \frac{1}{2} (\theta - \theta') [\bar{\theta} + \bar{\theta}' - (\bar{\theta} - \bar{\theta}') \operatorname{sgn}(t - t')] \frac{\partial}{\partial t} \right\} A(t - t', k),$$

where $\operatorname{sgn}(t)$ is the sign function. In our example, $A(t, k)$ is in addition an even real function of t . In terms of the Fourier transform over time,

$$B(\omega, k) = \int_0^{+\infty} e^{i\omega t} A(t, k) dt, \quad (36.18)$$

the time-Fourier transform of $\widetilde{W}^{(2)}$ takes the form

$$\begin{aligned} \tilde{W}^{(2)}(\omega, k, \boldsymbol{\theta}, \boldsymbol{\theta}') &= B(\omega, k) + B^*(\omega, k) - i\omega (\theta - \theta') [\bar{\theta}' B(\omega, k) + \bar{\theta} B^*(\omega, k)] \\ &\quad + A(0, k) \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}'), \end{aligned} \quad (36.19)$$

in which B^* is the complex conjugate of B .

The sum $B + B^*$, which is the $\langle \boldsymbol{\varphi} \boldsymbol{\varphi} \rangle$ connected two-point function, is a real function. The response function $\langle \lambda \boldsymbol{\varphi} \rangle$ is the coefficient of $\theta' \bar{\theta}'$. Taking its imaginary part, we find the relation

$$\begin{aligned} \operatorname{Im} \left[\frac{\delta}{\delta h(\omega, k)} \langle \boldsymbol{\varphi}(-\omega, -k) \rangle \right] \Big|_{h=0} &\equiv \langle \lambda(\omega, k) \boldsymbol{\varphi}(-\omega, -k) \rangle \\ &= \frac{1}{2} \omega \langle \boldsymbol{\varphi}(\omega, k) \boldsymbol{\varphi}(-\omega, -k) \rangle, \end{aligned} \quad (36.20)$$

a form of what is known under the name of *fluctuation-dissipation theorem*.

36.1.2 RG equations at T_c for $d < 4$

In Section 35.4.3, we have shown that static and supersymmetric dynamic theories have the same upper-critical dimension. Therefore, fluctuations are only relevant for dimensions $d \leq 4$. From the discussion of Chapter 35, we know that the renormalized action $\mathcal{S}_r(\phi)$ then has the form

$$\mathcal{S}_r(\phi) = \int dt d\bar{\theta} d\theta \left[\int d^d x \frac{2}{\Omega} Z_\omega \bar{D}\phi(t, x, \theta) \cdot D\phi(t, x, \theta) + \mathcal{H}_r(\phi) \right], \quad (36.21)$$

where ϕ is now the renormalized superfield, and $\mathcal{H}_r(\phi)$ is the static renormalized Hamiltonian, where $\varphi(x)$ is replaced by $\phi(t, x, \theta)$.

In this section, we work in the scheme of dimensional regularization and renormalization by minimal subtraction (MS) at a physical scale μ .

To renormalize the action (36.13), we have introduced, in addition to the static renormalization constants, a renormalization of the parameter Ω :

$$\Omega_0 = \Omega Z/Z_\omega, \quad (36.22)$$

where Z is the field renormalization constant. The RG differential operator in equation (9.70) is replaced by

$$D_{RG} = \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} + \eta_\omega(g) \Omega \frac{\partial}{\partial \Omega} - \frac{n}{2} \eta(g), \quad (36.23)$$

where $\eta_\omega(g)$ is a new independent RG function, defined by

$$\eta_\omega(g) = \mu \frac{d}{d\mu} \Big|_{g_0, \Omega_0} \ln \Omega. \quad (36.24)$$

In the case of the MS scheme (Section 10.4), equation (36.24) becomes

$$\eta_\omega(g) = \beta(g) \frac{d}{dg} \ln(Z_\omega/Z). \quad (36.25)$$

The vertex functions of the *critical theory*, in the Fourier representation, satisfy the RG equations

$$D_{RG} \tilde{\Gamma}^{(n)}(\omega_i, p_i, \theta_i, \mu, \Omega, g) = 0. \quad (36.26)$$

At an infrared (IR) fixed point g^* (see e.g. equation (15.53)), they reduce to

$$\left(\mu \frac{\partial}{\partial \mu} + \eta_\omega \Omega \frac{\partial}{\partial \Omega} - \frac{n}{2} \eta \right) \tilde{\Gamma}^{(n)}(\omega_i, p_i, \theta_i, \mu, \Omega) = 0, \quad (36.27)$$

with

$$\eta_\omega = \eta_\omega(g^*).$$

We set

$$z = 2 + \eta_\omega. \quad (36.28)$$

Dimensional analysis implies

$$\tilde{\Gamma}^{(n)} \left(\rho \omega_i, \lambda p_i, \frac{\theta_i}{\sqrt{\rho}}, \lambda \mu, \frac{\rho \Omega}{\lambda^2} \right) = \lambda^{d-n(d-2)/2} \rho^{1-n} \tilde{\Gamma}^{(n)}(\omega_i, p_i, \theta_i, \mu, \Omega). \quad (36.29)$$

In the dimensional equation (36.29), we choose

$$\rho = \Omega \lambda^z \mu^{-\eta_\omega}. \quad (36.30)$$

Then, combining the solution of the RG equation with equation (36.29), we find the dynamic scaling form of vertex functions:

$$\tilde{\Gamma}^{(n)}(\omega_i, \lambda p_i, \theta_i, \mu = \Omega = 1) = \lambda^{d-n(d-2+\eta)/2-z(n-1)} F^{(n)}(\lambda^{-z} \omega_i, p_i, \theta_i \lambda^{z/2}). \quad (36.31)$$

A few algebraic transformations yield the corresponding dynamic scaling form of the connected correlation functions:

$$\widetilde{W}^{(n)}(\omega_i, \lambda p_i, \theta_i, \mu = \Omega = 1) = \lambda^{(d+z)(1-n)+n(d-2+\eta)/2} G^{(n)}(\omega_i \lambda^{-z}, p_i, \theta_i \lambda^{z/2}). \quad (36.32)$$

Setting $n = 2$ and $\theta = 0$, one obtains the scaling form of the φ -field two-point correlation function:

$$\widetilde{W}^{(2)}(\omega, p, \theta = 0) \sim p^{-2+\eta-z} G^{(2)}(\omega/p^z). \quad (36.33)$$

The equal-time correlation function is obtained by integrating over ω . One recovers the static scaling.

The dynamic critical two-point function thus depends on a frequency scale that vanishes at small momentum like p^z , or a time scale that diverges like p^{-z} . The RG function η_ω for this model is calculated in Section A36.1.1 (see equation (A36.13)):

$$\eta_\omega(\tilde{g}) = \frac{N+2}{72} [6 \ln(4/3) - 1] \tilde{g}^2 + O(\tilde{g}^3), \quad (36.34)$$

with, as usual

$$\tilde{g} = N_d g, \quad N_d = \frac{2}{\Gamma(d/2)(4\pi)^{d/2}}. \quad (36.35)$$

For the dimension $d = 4 - \varepsilon$, the IR fixed point corresponds, at leading order, to $g^* = 48\pi^2\varepsilon/(N+8)$ (see Section 15.7), and the dynamic critical exponent z then is given by

$$z = 2 + \frac{N+2}{2(N+8)^2} [6 \ln(4/3) - 1] \varepsilon^2 + O(\varepsilon^3). \quad (36.36)$$

At this order, it can also be written as

$$z = 2 + c\eta + O(\varepsilon^3), \quad (36.37)$$

where the exponent η (defined, *e.g.* in equation (14.67)) is related to the dimension d_φ of the field in the static theory as $\eta = 2d_\varphi - d + 2$. One finds (the exponent z for model A is given at order ε^3 in Ref. [370])

$$c = 6 \ln(4/3) - 1. \quad (36.38)$$

The scaling behaviour of the response function can be obtained by considering, for example, the coefficient of $\theta\bar{\theta}$ in $W^{(2)}$.

36.1.3 Correlation functions above T_c , in the critical domain

Notation. Because here t and τ denote times, we characterize the deviation $T - T_c$ from the critical temperature by the renormalized parameter $v \propto r_0 - r_{0,c}$.

The RG equations directly at g^* become (see equation (9.82))

$$\left(\mu \frac{\partial}{\partial \mu} + \eta_\omega \Omega \frac{\partial}{\partial \Omega} - \eta_2 v \frac{\partial}{\partial v} - \frac{n}{2} \eta \right) \tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, v, \mu, \Omega) = 0. \quad (36.39)$$

The generalization of the dimensional relation (36.29) is

$$\tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, v, \mu, \Omega) = \lambda^{d-n(d-2)/2} \rho^{1-n} \tilde{\Gamma}^{(n)} \left(\frac{\omega_i}{\rho}, \frac{p_i}{\lambda}, \boldsymbol{\theta}_i \sqrt{\rho}, \frac{v}{\lambda^2}, \frac{\mu}{\rho}, \frac{\Omega \lambda^2}{\rho} \right). \quad (36.40)$$

Combining this equation with the RG equation (36.39), and choosing

$$\lambda = v^\nu \mu^{\nu \eta/2} \sim \xi^{-1}, \quad \rho = \Omega \mu^{-\eta \omega} \lambda^z \sim \xi^{-z}, \quad (36.41)$$

where ξ is the correlation length, one obtains the scaling form

$$\begin{aligned} \tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, v, \mu = 1, \Omega = 1) &\sim \xi^{-d+n(d-2+\eta)/2+z(n-1)} \\ &\times F^{(n)}(\omega_i \xi^z, p_i \xi, \boldsymbol{\theta}_i \xi^{-z/2}). \end{aligned} \quad (36.42)$$

All times are expressed in unit of the *correlation time* τ , which diverges at the critical temperature as (a manifestation of *critical slowing down*)

$$\tau \propto \xi^z. \quad (36.43)$$

36.2 Dissipative dynamics: RG equations in dimension $d = 2 + \varepsilon$

Near two dimensions, the N -vector model is described by a Langevin equation of the form (see also Sections 35.6 and A36.1.2)

$$\begin{aligned} \dot{\varphi}(t, x) &= -\frac{1}{2} \Omega_0 \left[-\nabla^2 \varphi(t, x) + \varphi(t, x) (\varphi(t, x) \cdot \nabla^2 \varphi(t, x)) \right] \\ &+ \boldsymbol{\nu}(t, x) - \varphi(t, x) (\varphi(t, x) \cdot \boldsymbol{\nu}(t, x)), \end{aligned} \quad (36.44)$$

$\boldsymbol{\nu}(t, x)$ being the Gaussian noise (36.10). The quantity $\varphi^2(t, x)$ is time independent, and we choose

$$\varphi^2(t, x) = 1/g_0. \quad (36.45)$$

From equation (36.44), one derives the dynamic action

$$\mathcal{S}(\phi) = \int d\bar{\theta} d\theta dt \left[\int d^d x \frac{2}{\Omega_0} \bar{D}\phi(t, x, \boldsymbol{\theta}) \cdot D\phi(t, x, \boldsymbol{\theta}) + \mathcal{H}(\phi) \right],$$

where ϕ is the superfield defined by equation (36.12), and

$$\mathcal{H}(\phi) = \frac{1}{2} \int d^d x [\nabla \phi(\boldsymbol{\theta}, t, x)]^2. \quad (36.46)$$

After the rescaling $\phi \mapsto \phi/\sqrt{g_0}$, the superfield satisfies the constraint

$$\phi^2(\boldsymbol{\theta}, t, x) = 1. \quad (36.47)$$

The dynamic action then becomes

$$\mathcal{S}(\phi) = \frac{1}{g_0} \int d\bar{\theta} d\theta dt \left[\int d^d x \frac{2}{\Omega_0} \bar{D}\phi(t, x, \boldsymbol{\theta}) \cdot D\phi(t, x, \boldsymbol{\theta}) + \mathcal{H}(\phi) \right]. \quad (36.48)$$

In the superfield form (36.48), the dynamical actions for $(\phi^2)^2$ field theory and the nonlinear σ -model are related in the same way as the corresponding static Hamiltonians.

The renormalized action $\mathcal{S}_r(\phi)$ then reads

$$\mathcal{S}_r(\phi) = \frac{\mu^\varepsilon}{g} \int dt d\bar{\theta} d\theta \left[\int d^d x \frac{2}{\Omega} Z_\omega \bar{D}\phi(t, x, \boldsymbol{\theta}) \cdot D\phi(t, x, \boldsymbol{\theta}) + \mathcal{H}_r(\phi) \right],$$

in which μ is the renormalization scale and $d = 2 + \varepsilon$.

In the MS scheme, the new RG function $\eta_\omega(g)$ is given by

$$\eta_\omega(g) = \beta(g) \frac{d}{dg} \ln(Z_\omega Z_g / Z). \quad (36.49)$$

The RG equations for vertex functions in zero-magnetic field then read (see equation (19.92))

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} + \eta_\omega(g) \Omega \frac{\partial}{\partial \Omega} - \frac{n}{2} \zeta(g) \right] \tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, g, \mu, \Omega) = 0. \quad (36.50)$$

Dimensional analysis yields (equation (36.29))

$$\tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, g, \mu, \Omega) = \lambda^d \rho^{1-n} \tilde{\Gamma}^{(n)} \left(\frac{\omega_i}{\rho}, \frac{p_i}{\lambda}, \boldsymbol{\theta}_i \sqrt{\rho}, \frac{\Omega \lambda^2}{\rho}, \frac{\mu}{\lambda} \right). \quad (36.51)$$

In addition to the functions (see equations (19.105, 19.106))

$$M_0(g) = \exp \left[-\frac{1}{2} \int_0^g \frac{\zeta(g')}{\beta(g')} dg' \right], \quad (36.52)$$

$$\xi(g) = \mu^{-1} g^{1/\varepsilon} \exp \left[\int_0^g \left(\frac{1}{\beta(g')} - \frac{1}{\varepsilon g'} \right) dg' \right], \quad (36.53)$$

we introduce the new RG function

$$\tau(g) = \Omega^{-1} \xi^2(g) \exp \left[\int_0^g \frac{\eta_\omega(g')}{\beta(g')} dg' \right]. \quad (36.54)$$

Combining equations (36.50–36.54) and choosing

$$\lambda = \xi^{-1}(g), \quad \rho = \tau^{-1}(g), \quad (36.55)$$

one obtains the scaling form

$$\tilde{\Gamma}^{(n)}(\omega_i, p_i, \boldsymbol{\theta}_i, g, \mu = 1, \Omega = 1) = \tau^{n-1} \xi^{-d} M_0^{-n} F^{(n)}(\omega_i \tau, p_i \xi, \boldsymbol{\theta}_i \tau^{-1/2}).$$

Near the critical temperature g^* (see Section 19.12), this expression agrees with the scaling form (36.42). The exponent z then is given by

$$z = 2 + \eta_\omega(g^*). \quad (36.56)$$

Equation (A36.32) gives $\eta_\omega(g)$ at two-loop order. In dimension $d = 2 + \varepsilon$, at the critical and ultraviolet (UV) fixed point $g^* = 2\pi\varepsilon/(N-2) + O(\varepsilon^2)$ (equation (19.98) in Section 19.12), one finds the dynamic exponent

$$z = 2 + (1 - \ln(4/3)) \frac{\varepsilon^2}{N-2} + O(\varepsilon^3), \quad (36.57)$$

or again, in terms of the static exponent η (equation (14.67)),

$$z = 2 + (1 - \ln(4/3)) \varepsilon \eta + O(\varepsilon^3). \quad (36.58)$$

36.3 Conserved order parameter

A simple modification of equation (36.9) ensures that the order parameter is conserved, which is the statement

$$\frac{d}{dt} \int d^d x \varphi(t, x) = 0. \quad (36.59)$$

We consider the equation

$$\dot{\varphi}(t, x) = \frac{\Omega_0}{2} \nabla_x^2 \frac{\delta \mathcal{H}}{\delta \varphi(t, x)} + \nu(t, x), \quad (36.60)$$

where the centred Gaussian noise distribution is defined by

$$\langle \nu_i(t, x) \nu_j(t', x') \rangle_\nu = -\delta_{ij} \Omega_0 \delta(t - t') \nabla_x^2 \delta^{(d)}(x - x'). \quad (36.61)$$

The dynamic action $\mathcal{S}(\phi)$ is still supersymmetric:

$$\mathcal{S}(\phi) = \int dt d\bar{\theta} d\theta \left[- \int d^d x d^d y \frac{2}{\Omega_0} \bar{D}\phi(t, x, \theta) \Delta(x - y) D\phi(t, y, \theta) + \mathcal{H}(\phi) \right], \quad (36.62)$$

where $\Delta(x)$ is given by (see equation (10.5))

$$-\nabla_x^2 \Delta(x) = \delta^{(d)}(x) \Rightarrow \Delta(x) = \frac{\Gamma(d/2 - 1)}{4\pi^{d/2}} |x|^{2-d}.$$

This form implies that this Langevin equation generates the same equilibrium distribution as model A.

The appearance of a non-local term in the action leads to problems already met in the example of effective field theory (EFT) for uniaxial systems with dipolar forces (Section 17.5). Power counting is different. Since the propagator reads,

$$\tilde{\Delta}(\theta, \theta', \omega, k,) = \frac{\Omega k^2 [1 - \frac{1}{2} i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}') + \frac{1}{4} \Omega k^2 (k^2 + r) \delta^2(\theta - \theta')]}{\omega^2 + \frac{1}{4} \Omega^2 (k^2)^2 (k^2 + r)^2}, \quad (36.63)$$

ω has the dimension of k^4 . Above four dimensions, the characteristic frequency diverges like k^4 . This implies that the dynamical exponent z is 4, instead of 2 as in model A.

Above two dimensions, Feynman diagrams calculated with the propagator (36.63) are not singular at zero momentum for $\omega \neq 0$. Therefore, no counter-term singular in \mathbf{k} can be generated, and Ω_0 in the dynamic action is not renormalized. The renormalized action takes the form

$$\mathcal{S}_r(\phi) = \int dt d\bar{\theta} d\theta \left[- \int d^d x \frac{2}{\Omega} \bar{D}\phi(t, x, \theta) \cdot \nabla_x^{-2} D\phi(t, x, \theta) + \mathcal{H}_r(\phi) \right], \quad (36.64)$$

The field renormalization implies the relation

$$\Omega_0 = Z\Omega \quad (36.65)$$

and, therefore, the RG function $\eta_\omega(g)$ is

$$\eta_\omega(g) = -\beta(g) \frac{d}{dg} \ln Z = -\eta(g). \quad (36.66)$$

One then proceeds in a way quite similar what has been done in Section 36.1. However, the relation (36.66) implies that the exponent z is no longer an independent exponent but, by contrast, is related to the static exponent η by,

$$z = 4 - \eta. \quad (36.67)$$

36.4 Relaxational model with energy conservation

Still in the framework of the N -vector model, we now assume that the total energy is conserved. We know that in the critical domain the most singular part of the energy is φ^2 . Therefore, we couple a field $e(x)$ to $\varphi^2(x)$ and consider the equilibrium Hamiltonian

$$\begin{aligned} \mathcal{H}(\varphi, e) = & \int d^d x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} r \varphi^2(x) + \frac{u}{4!} (\varphi^2(x))^2 \right. \\ & \left. + \frac{1}{2} \Lambda^{\varepsilon/2} v e(x) \varphi^2(x) + \frac{1}{2} e^2(x) \right], \end{aligned} \quad (36.68)$$

where v is new coupling constant. The static properties are not affected by this modification, because after integration over the e -field, one recovers the usual $(\varphi^2)^2$ theory. We denote by $K(x)$ the source for the correlation functions of $e(x)$. If one integrates over $e(x)$, one then obtains the new Hamiltonian

$$\begin{aligned} \mathcal{H}(\varphi, K) = & \int d^d x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} r \varphi^2(x) + \frac{1}{4!} (u - 3\Lambda^\varepsilon v^2) (\varphi^2(x))^2 \right. \\ & \left. + \frac{1}{2} \Lambda^{\varepsilon/2} v K(x) \varphi^2(x) - \frac{1}{2} K^2(x) \right]. \end{aligned} \quad (36.69)$$

In the static theory, the e -field is directly related to the φ^2 field, which is the energy operator. The e -field amplitude and the v coupling constant renormalizations can be expressed in terms of φ^2 multiplicative and $\langle \varphi^2 \varphi^2 \rangle$ additive renormalization constants. More precisely, differentiating the field integral twice with respect to $K(x)$, one obtains the relation between correlation functions,

$$\langle e(x) e(y) \rangle = \delta^{(d)}(x - y) + \Lambda^\varepsilon v^2 \langle (\frac{1}{2} \varphi^2(x)) (\frac{1}{2} \varphi^2(y)) \rangle, \quad (36.70)$$

and finds that all other correlation functions of the field $e(x)$ are identical to the correlation functions of $-\Lambda^{\varepsilon/2} v \varphi^2(x)/2$. From the RG equations satisfied by the vertex functions with φ^2 insertions (equation (15.67)), one can derive RG equations with $e(x)$ insertions. At T_c , they take the form

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \beta_v(g, v) \frac{\partial}{\partial v} - \frac{l}{2} \eta_e(g, v) - \frac{n}{2} \eta(g) \right] \tilde{\Gamma}_e^{(l,n)}(q, p; g, v, \Lambda) = 0, \quad (36.71)$$

where $\Gamma_e^{(l,n)}$ is the correlation function with l $e(x)$ insertions, and

$$\eta_e(g, v) = \varepsilon + 2\beta_v/v + 2\eta_2(g). \quad (36.72)$$

Then, applying equation (36.71) to the relation (36.70), and again using equation (15.67), one finds

$$\varepsilon + 2\beta_v/v + 2\eta_2(g) = v^2 B(g), \quad (36.73)$$

and, therefore,

$$\eta_e = v^2 B(g), \quad (36.74a)$$

$$\beta_v = -\frac{1}{2} v [\varepsilon + 2\eta_2 - v^2 B(g)]. \quad (36.74b)$$

At the IR fixed point g^* of the static theory, the function β_v becomes

$$\beta_v = -\frac{1}{2}v [\alpha/\nu - v^2 B(g^*)]. \quad (36.75)$$

At leading order in the ε -expansion,

$$B(g^*) = \frac{N}{16\pi^2} + O(\varepsilon)$$

which is positive. Therefore, two cases have to be envisaged:

(i) $\alpha < 0$:

the origin $v = 0$ is the unique IR fixed point and the coupling to $e(x)$ is irrelevant; we are back to the model A of Section 36.1; and

(ii) $\alpha > 0$:

the IR fixed point now is given by

$$v = v^* \equiv \pm \left(\frac{\alpha}{\nu B(g^*)} \right)^{1/2}. \quad (36.76)$$

At leading order α vanishes for $N = 4$:

$$\alpha = \frac{4-N}{2(N+8)}\varepsilon + O(\varepsilon^2). \quad (36.77)$$

In three dimensions, numerical calculations (see Section 41.4, Table 41.3) show that α is already slightly negative for $N = 2$.

For $\alpha > 0$, the dynamics differs from the dynamics of Section 36.1, as we will show.

Still, in both cases, the values of η_e lead to a behaviour consistent with previous results concerning the $\langle \varphi^2 \varphi^2 \rangle$ correlation functions.

The Langevin equation. We now examine the dynamics of the model in the case $\alpha > 0$. The Langevin equations consistent with the energy-conservation equation

$$\frac{d}{dt} \int d^d x e(t, x) = 0$$

are

$$\dot{\varphi}(t, x) = -\frac{\Omega}{2} \frac{\delta \mathcal{H}}{\delta \varphi(t, x)} + \nu(t, x), \quad (36.78)$$

$$\dot{e}(t, x) = \frac{\Omega'}{2} \nabla_x^2 \frac{\delta \mathcal{H}}{\delta e(t, x)} + \nu'(t, x). \quad (36.79)$$

The new Gaussian noise $\nu'(t, x)$ is defined by

$$\langle \nu' \rangle_{\nu'} = 0, \quad \langle \nu'(t, x) \nu'(x', t') \rangle_{\nu'} = -\Omega' \delta(t - t') \nabla_x^2 \delta^{(d)}(x - x'). \quad (36.80)$$

The dynamic action, now expressed in terms of two superfields $\phi(t, x)$ and $E(t, x)$, is still supersymmetric and reads

$$\mathcal{S}(\phi, E) = \int d\bar{\theta} d\theta dt \left\{ \int d^d x \left[-\frac{2}{\Omega'} \bar{D}E \nabla_x^{-2} D E + \frac{2}{\Omega} \bar{D}\phi D\phi \right] + \mathcal{H}(\phi, E) \right\}. \quad (36.81)$$

As we have already explained in the example of the φ -field conservation, since the $\langle EE \rangle$ correlation function is not singular at zero momentum, the parameter Ω' remains unrenormalized. However, the renormalization of Ω is now modified by the presence of loops with E fields.

Setting again

$$\Omega = \Omega_r / Z_\omega \quad (36.82)$$

at one-loop order one finds

$$Z_\omega = 1 - \frac{\Omega}{\Omega + \Omega'} N_d v^2 \ln \Lambda, \quad (36.83)$$

while Ω' is given by relation,

$$\Omega' = Z_e \Omega'_r, \quad (36.84)$$

in which Z_e is the e -field renormalization. The corresponding RG function η_e is given by equation (36.74a).

The RG equations then read

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \beta_v \frac{\partial}{\partial v} + \eta_\omega \Omega \frac{\partial}{\partial \Omega} + \eta'_\omega \Omega' \frac{\partial}{\partial \Omega'} - \frac{l}{2} \eta_e - \frac{n}{2} \eta \right] \times \tilde{\Gamma}_e^{(l,n)}(q, p; g, v, \Lambda, \Omega, \Omega') = 0. \quad (36.85)$$

Equation (36.84) leads to the identity

$$\eta'_\omega = \eta_e, \quad (36.86)$$

while at leading order

$$\eta_\omega = \frac{1}{8\pi^2} \frac{\Omega}{\Omega + \Omega'} v^2. \quad (36.87)$$

To separate variables, we set

$$s = \Omega'/\Omega, \quad (36.88)$$

and take, for example, s and Ω' as independent variables. We infer

$$\eta_\omega \Omega \frac{\partial}{\partial \Omega} + \eta_e \Omega' \frac{\partial}{\partial \Omega'} = \eta_e \Omega' \frac{\partial}{\partial \Omega'} + \beta_s \frac{\partial}{\partial s}, \quad (36.89)$$

with

$$\beta_s = \eta_e - \eta_\omega. \quad (36.90)$$

At the IR fixed point in the (g, v) plane, the function β_s takes the form

$$\beta_s = \frac{\alpha}{\nu} \left(1 - \frac{2}{N} \frac{1}{1+s} \right) s + O(\varepsilon^2). \quad (36.91)$$

The fixed points in the variables s are

$$s^* = 0, \quad s^* = 2/N - 1, \quad s = \infty. \quad (36.92)$$

Since α has been assumed to be positive, the fixed point $s = 0$ is stable for $N > 2 + O(\varepsilon)$. However, the $s = 0$ limit is a singular limit, and it is not clear whether this result is consistent with the ε -expansion (and α is negative for $N \geq 2$).

The fixed point $s = \infty$ is never stable for $\alpha > 0$. It also corresponds to a decoupling of the E sector.

Finally, for $N < 2 + O(\varepsilon)$, $s^* = 2/N - 1$ is the stable fixed point. The ratio Ω'/Ω is finite and, therefore, the dynamics of E and ϕ are coupled. The function η_e corresponds to η_ω in model A and, therefore, since η_e at the fixed point has the value α/ν , the exponent z is

$$s^* = 2/N - 1 + O(\varepsilon) \Rightarrow z = 2 + \alpha/\nu. \quad (36.93)$$

36.5 A non-relaxational model

General remarks. We have argued that if in the Langevin equation (36.1) (we have set $\beta = 1$)

$$\dot{\varphi}(t) = -\frac{1}{2} \sum_j \Omega_{ij} \frac{\delta \mathcal{H}}{\delta \varphi_j(t)} + F_i(t) + \nu_i(t), \quad (36.94)$$

with the centred Gaussian noise (36.2), $\langle \nu_i(t) \nu_j(t') \rangle_\nu = \Omega_{ij} \delta(t - t')$, the function F_i satisfies equation (36.3),

$$\sum_i \frac{\partial}{\partial \varphi_i} \left[F_i(\varphi) e^{-\mathcal{H}(\varphi)} \right] = 0, \quad (36.95)$$

then, the equilibrium distribution is $e^{-\mathcal{H}(\varphi)}$. A particular class of solutions to equation (36.95) is provided by

$$F_i(\varphi) = \sum_j \left(\frac{\partial}{\partial \varphi_j} R_{ij}(\varphi) - R_{ij}(\varphi) \frac{\partial \mathcal{H}}{\partial \varphi_j} \right), \quad (36.96)$$

where R_{ij} is an antisymmetric matrix, $R_{ij} = -R_{ji}$.

In particular, the mode-coupling of Kawasaki and Kadanoff–Swift has this form. In concrete examples, the matrix R_{ij} is linear in the field φ_i , and associated with transformations corresponding to symmetries of the Hamiltonian.

Example. We now give a physical example of a model with non-dissipative couplings. Since the dynamic action in such cases is no longer supersymmetric, we expect the number of independent renormalizations to increase substantially and, therefore, the analysis to become more complex. Moreover, by losing the supersymmetry, we lose a powerful and elegant technique to solve the renormalization problem.

We consider the so-called model E: the order parameter is a complex field $\varphi(x)$ and, as in the preceding section, there is a conserved density $e(x)$. However, the Langevin equation reads

$$\dot{\varphi}(t, x) = -\frac{\Omega}{2} \frac{\delta \mathcal{H}}{\delta \varphi^*(t, x)} - i s \varphi(t, x) \frac{\delta \mathcal{H}}{\delta e(t, x)} + \nu(t, x), \quad (36.97)$$

$$\dot{e}(t, x) = \frac{\Omega'}{2} \nabla^2 \frac{\delta \mathcal{H}}{\delta e(t, x)} + i s \left[\varphi^*(t, x) \frac{\delta \mathcal{H}}{\delta \varphi^*(t, x)} - \varphi(t, x) \frac{\delta \mathcal{H}}{\delta \varphi(t, x)} \right] + \nu'(t, x), \quad (36.98)$$

where s is an additional parameter. The Hamiltonian is

$$\mathcal{H}(\varphi, e) = \int d^d x \left[|\nabla \varphi(x)|^2 + r |\varphi(x)|^2 + \frac{1}{3!} u |\varphi(x)|^4 + \frac{1}{2} e^2(x) \right]. \quad (36.99)$$

The two-point functions of the centred Gaussian noise are the same as in Section 36.4 (for the $N = 2$ case):

$$\begin{aligned} \langle \nu(t, x) \nu^*(t', x') \rangle_\nu &= \Omega \delta(t - t') \delta^{(d)}(x - x'), \\ \langle \nu \nu \rangle_\nu &= \langle \nu^* \nu^* \rangle_\nu = 0, \\ \langle \nu'(t, x) \nu'(t', x') \rangle_\nu &= -\Omega' \delta(t - t') \nabla_x^2 \delta^{(d)}(x - x'). \end{aligned} \quad (36.100)$$

It is simple to verify that this model provides one example of Langevin equation (36.94) with a streaming term of the form (36.96), which ensures that $e^{-\mathcal{H}(\varphi)}$ remains the equilibrium distribution.

The model has a $U(1)$ symmetry corresponding to the multiplication of φ by a phase. From the invariance of the Hamiltonian under an infinitesimal $U(1)$ transformation follows,

$$\int d^d x \left(\varphi^*(t, x) \frac{\delta \mathcal{H}}{\delta \varphi^*(t, x)} - \varphi(t, x) \frac{\delta \mathcal{H}}{\delta \varphi(t, x)} \right) = 0$$

and, therefore, $e(t, x)$ is a conserved quantity.

To simplify the discussion, we have omitted the generically expected $e|\varphi^2|$ coupling and this generates an additional reflection symmetry $e(x) \mapsto -e(x)$. According to the analysis of Section 36.4, this simplification is really justified only if the exponent of the specific heat α is negative. The analysis can be generalized to the case where the $e|\varphi^2|$ coupling is included. The interested reader is referred to the literature.

Renormalization. Power counting tells us that the canonical dimensions of the fields are

$$[\varphi] = 1, \quad [e] = 2,$$

and that the theory is renormalizable in four dimensions. From the general results of Section 35.3.2 based on the BRST symmetry of the dynamic action, we know that the Langevin equations renormalize as predicted by power counting. The form of the renormalized equations is further restricted by the $U(1)$ symmetry, the parity symmetry and the conservation of $e(t, x)$. However, these conditions are not sufficiently restrictive. They do not forbid a term proportional to $\nabla^2(\varphi^*\varphi)$ in equation (36.98) and do not imply the equality of the coupling constants s in equations (36.97) and (36.98). Here, we need a rather indirect argument: since the regularized dynamic theory has the regularized static theory as equilibrium distribution, the same must be true for the renormalized theories. We then have three renormalization constants given by the static properties, Z_φ , $Z_e = 1$, and Z_u . In addition, we have to renormalize Ω , Ω' and s . However, a WT identity follows from the remark that $e(t, x)$ is coupled to phase transformations on the field φ . Indeed, if we perform the transformation

$$\varphi(t, x) = e^{i\alpha(t)} \varphi'(t, x), \quad e(t, x) = e'(t, x) - \dot{\alpha}(t)/s,$$

equation (36.97) is unchanged, while an additional field independent term is added to the right-hand side of equation (36.98): $\ddot{\alpha}/s$. Therefore, the renormalization of s is connected to the renormalization of the Lagrange multiplier λ_e associated with e in the dynamic action. It is then simple to verify that s is not renormalized.

The RG β -functions. The model depends on three independent dimensionless coupling constants, which we can choose to be u ,

$$v = s^2/\Omega\Omega', \quad \text{and} \quad w = \Omega/\Omega'. \tag{36.101}$$

The function β_u is given by the static properties and determines the IR fixed point value $u^* = 3\varepsilon/40\pi^2 + O(\varepsilon^2)$. From the previous discussion, and dimensional considerations, it follows that the two other β -functions can be written in dimension $d = 4 - \varepsilon$ as

$$\beta_v = -v(\varepsilon + \eta_\omega + \eta_{\omega'}), \tag{36.102}$$

$$\beta_w = w(\eta_\omega - \eta_{\omega'}). \tag{36.103}$$

The coupling constant v has one obvious fixed point value $v = 0$, which decouples $\varphi(t, x)$ and $e(t, x)$.

Then, $\eta_{\omega'} = 0$ and η_{ω} assumes the value of model A (equation (36.34)) and is thus positive. The stability matrix constructed with the derivatives of the β -functions has an eigenvalue ω given by

$$\omega = (\partial \beta_v / \partial v) = -(\varepsilon + \eta_{\omega}) < 0,$$

showing that model A is unstable with respect to the introduction of the coupling s .

If v does not vanish, equation (36.102) implies

$$\varepsilon + \eta_{\omega} + \eta_{\omega'} = 0. \quad (36.104)$$

Equation (36.103) has three type of solutions, $w = \infty$, $w = 0$, or $\eta_{\omega} = \eta_{\omega'}$. To find the stable fixed points, we then need the RG functions at leading order:

$$\eta_{\omega} = -\frac{v}{8\pi^2(1+w)}, \quad \eta_{\omega'} = -\frac{v}{16\pi^2}. \quad (36.105)$$

It follows that:

(i) The fixed point $w = \infty$ is unstable, because the stability matrix has one negative eigenvalue $-\varepsilon$.

(ii) The fixed point $w = 0$, at leading order, also appears to be unstable, because one eigenvalue is negative $\omega = -\varepsilon/3$. However, the next term in the ε -expansion has been calculated and is positive. Therefore, one cannot exclude that this fixed point becomes stable for $\varepsilon = 1$. This fixed point exhibits an interesting violation of dynamic scaling, since the renormalized ratio of the time scales Ω^{-1} and Ω'^{-1} vanishes.

(iii) The last fixed point $w = \varepsilon/2 + O(\varepsilon^2)$ corresponds to normal dynamic scaling, since the two time scales have a finite relation. All eigenvalues of the stability matrix are positive at leading order in ε . In this case, $\eta_{\omega} = \eta_{\omega'}$ and equation (36.104) hold. Therefore, to all orders

$$\eta_{\omega} = \eta_{\omega'} = -\varepsilon/2. \quad (36.106)$$

The exponent z is then exactly known:

$$z = d/2. \quad (36.107)$$

For a more detailed discussion, and a selection of other models, we refer the interested reader to the literature [367, 371].

36.6 Finite size effects and dynamics

In Chapter 35, and so far in Chapter 36, we have discussed dynamics from the RG point of view. In Chapter 32, we have analysed static finite size effects. The combined RG considerations also apply to the dynamics in a finite volume. As in the case of static properties, finite-size effects are characterized by the dependence in the scaling variable $L/\xi(T)$ (Section 32.1.1). If the IR fixed point is not Gaussian, in the critical domain the correlation time $\tau(T, L)$ in a finite volume of linear size L has the scaling form [372], which generalizes equation (36.43),

$$\tau(T, L) = L^z f \left[(T - T_c)L^{1/\nu} \right]. \quad (36.108)$$

For example, in a periodic hypercube, physical quantities can be calculated by methods analogous to those used in the case of the cylindrical geometry in previous sections, the time being now the physical time, instead of one of the spatial directions. Here, we illustrate these remarks by calculations of correlation times in the simple case of a *purely dissipative models* without conservation laws of Sections 36.1 and 36.2.

36.6.1 The $(\varphi^2)^2$ field theory

We consider the Langevin equation (36.9) with the Gaussian white noise (36.10). To the Langevin equation is associated a renormalized dynamic action $\mathcal{S}(\phi)$, which in terms of the superfield

$$\phi(t, x, \theta) = \varphi(t, x) + \theta \bar{c}(t, x) + c(t, x) \bar{\theta} + \theta \bar{\theta} \lambda(t, x) \quad (36.109)$$

takes the supersymmetric form (36.21)

$$\mathcal{S}(\phi) = \int dt d\bar{\theta} d\theta \left[\int d^d x \frac{2Z_\omega}{\Omega} \bar{D}\phi(t, x, \theta) \cdot D\phi(t, x, \theta) + \mathcal{H}_r(\phi) \right], \quad (36.110)$$

where $\mathcal{H}_r(\varphi)$ is the static Euclidean action (which be inferred from the action (9.77))

$$\mathcal{H}(\varphi) = \int d^d x \left[\frac{1}{2} Z (\nabla \varphi(x))^2 + \frac{1}{2} (r_c + Z_2 v) \varphi^2(x) + \frac{u Z_u}{4!} (\varphi^2(x))^2 \right]. \quad (36.111)$$

In a periodic hypercube, at the critical temperature, we have to separate the zero mode. At leading order, the effective action for the zero mode is just obtained by specializing the action (36.110) to a space-independent field. We then recognize the action associated with a stochastic differential equation of a type studied in Sections 34.4–34.7. The corresponding FP Hamiltonian H_{FP} , in Hermitian form is (equation (34.39))

$$H_{\text{FP}} = \frac{\Omega}{2L^d} \left[\mathbf{p}^2 + \frac{1}{4} (\nabla E(\mathbf{q}))^2 - \frac{1}{2} \nabla^2 E(\mathbf{q}) \right], \quad (36.112)$$

\mathbf{q} representing the field-zero mode. At leading order, since $r_c = 0$, $Z = Z_2 = Z_u = 1$,

$$E(\mathbf{q}) = L^d \left(\frac{1}{2} v \mathbf{q}^2 + \frac{u}{4!} (\mathbf{q}^2)^2 \right). \quad (36.113)$$

After the rescaling

$$\mathbf{q} \mapsto \mathbf{q} (L^d u)^{-1/4}, \quad (36.114)$$

we find that the eigenvalues E_i of H_{FP} , as functions of L and the deviation of the critical temperature v , take the form

$$E_i = \Omega u^{1/2} L^{-d/2} e_i(u^{-1/2} L^{d/2} v). \quad (36.115)$$

For $d > 4$, the relaxation time τ , which is the inverse of the difference of the two first eigenvalues, has the scaling form

$$\tau(T, L) = \Omega^{-1} u^{-1/2} L^{d/2} f(u^{-1/2} L^{d/2} v). \quad (36.116)$$

An analysis, similar to the one performed in the static case, shows that loop corrections do not modify the scaling form for $d > 4$. Like for the finite-size correlation length (see equation (32.59)), the naive extrapolation of the scaling form valid for $d < 4$ is incorrect.

Like for the correlation length, we examine the behaviour of τ when $v L^{d/2}$ is large. For $v > 0$, we find the expected limit $\tau^{-1} \rightarrow \Omega v$. For $v < 0$ and $N > 1$, we obtain

$$\tau \sim -\frac{12}{(N-1)g} \Omega^{-1} L^d v.$$

For $N = 1$, the behaviour is quite different, as we discuss in Section 36.6.3.

The relaxation time in $d = 4 - \varepsilon$ dimensions. We now calculate the relaxation time in $4 - \varepsilon$ dimensions, at the IR fixed point. The calculations are performed in the MS scheme using the supersymmetric formulation of Section A36.1.1. Following the lines of Section 32.4, we perform a rescaling in the action, equivalent to equation (36.114),

$$\phi \mapsto \phi (L^d u)^{-1/4}, \quad t \mapsto t (L^d / u)^{1/2}, \quad \theta \mapsto \theta (L^d / u)^{1/4}, \quad (36.117)$$

and calculate the powers of ε . Again, we verify that only the coefficients of ϕ^2 , $v\phi^2$, and $(\phi^2)^2$ are relevant at one-loop order. Moreover, the scaling of the time and Grassmann variables shows that only the contributions proportional to $(1, v) \int dt d\bar{\theta} d\theta \phi^2$ and $\int dt d\bar{\theta} d\theta (\phi^2)^2$ are required. One verifies that the calculation then becomes identical to the static calculation in a finite volume. The main result is given by equation (32.50). The relaxation time, at the IR fixed point, in the one-loop approximation, has the form

$$\tau(v, L) = \Omega'^{-1} L^z f(L^{1/\nu} v' + b), \quad (36.118)$$

in which $f(z)$ is the function implicitly defined by equation (36.116), Ω'^{-1} a renormalized time scale, and v' a renormalized deviation from the critical temperature.

36.6.2 The non-linear σ -model: The bare RG

With the non-linear σ -model considered in Section A36.1.2, we can calculate, for example, the relaxation time at fixed temperature below T_c , close to T_c in an $\varepsilon = (d - 2)$ expansion or in two dimensions. The bare dynamic action, with some change of notation, can be inferred from the expression (36.48),

$$\mathcal{S}(\phi) = \frac{\Lambda^\varepsilon}{g} \int d\bar{\theta} d\theta dt \left[\int d^d x \frac{2}{\Omega} \bar{D}\phi(t, x, \theta) \cdot D\phi(t, x, \theta) + \frac{1}{2} (\nabla\phi(t, x, \theta))^2 \right], \quad (36.119)$$

where Λ is the cut-off, g is dimensionless and plays the role of the temperature, and

$$\phi^2(t, x, \theta) = 1. \quad (36.120)$$

The RG equations of the model have been discussed in Section 36.2. The relaxation time, defined in the infinite volume by equation (36.54) satisfies the bare RG equation

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \eta_\omega(g) \Omega \frac{\partial}{\partial \Omega} \right) \tau(\Lambda, L, g, \Omega) = 0. \quad (36.121)$$

The finite-size relaxation time τ satisfies the dimensional relation

$$\tau(\Lambda, L, g, \Omega) = \Omega^{-1} \Lambda^{-2} \tau(1, \Lambda L, g, 1). \quad (36.122)$$

The RG equation (36.121) can thus be rewritten as

$$\left(L \frac{\partial}{\partial L} + \beta(g) \frac{\partial}{\partial g} - 2 - \eta_\omega(g) \right) \tau(\Lambda, L, g, \Omega) = 0. \quad (36.123)$$

We now set $\Lambda = 1$ and simplify the notation $\tau(\Lambda, L, g, \Omega) \mapsto \tau(L, g)$, because the dependence in Ω is trivial.

Solving this equation by the method of characteristics, we find

$$\tau(L, g) = L^2 \zeta(L, g) \tau(1, g_L), \quad (36.124)$$

with the notation $g_L \equiv g(1/L)$ (see Section 32.1.2), and

$$\ln L = \int_{g_L}^g \frac{dg'}{\beta(g')}, \quad (36.125)$$

$$\zeta(L, g) = \exp \left[\int_{g_L}^g \frac{\eta_\omega(g')}{\beta(g')} dg' \right]. \quad (36.126)$$

Equation (36.124) can also be rewritten as

$$\tau(L, g) = L^2 \zeta(L, g) \mathcal{T}(L/\xi(g)), \quad (36.127)$$

in which $\xi(g)$ is the infinite volume correlation length. As we have already shown in Section 32.1, at fixed coupling $g < g_c$, the effective temperature g_L goes to 0. Therefore, τ can be derived from perturbation theory.

Finally, note that the function $\eta_\omega(g)$ begins at order g^2 (see Section A36.1.2). Since we calculate only at one-loop order, the function $\zeta(L, g)$ can be replaced in what follows by a constant renormalization factor

$$\zeta(g) = \exp \left[\int_0^g \frac{\eta_\omega(g')}{\beta(g')} dg' \right]. \quad (36.128)$$

To calculate the relaxation time τ , we use the method explained in Section 32.5.2. For $N > 2$, we can parametrize the field $\phi(t, x, \theta)$ as

$$\phi(t, x, \theta) = \begin{cases} \cos \alpha(t, \theta) \sigma_1(t, x, \theta) - \sin \alpha(t, \theta) \sigma_2(t, x, \theta), \\ \sin \alpha(t, \theta) \sigma_1(t, x, \theta) + \cos \alpha(t, \theta) \sigma_2(t, x, \theta), \\ \boldsymbol{\pi}(t, x, \theta), \end{cases} \quad (36.129)$$

where σ_2 and $\boldsymbol{\pi}$ have no zero mode and determine the one-loop effective action for the field $\alpha(\theta, t)$.

Eliminating σ_1 through the relation

$$\sigma_1 = (1 - \sigma_2^2 - \boldsymbol{\pi}^2)^{1/2},$$

we again find that the part of the action relevant at one-loop order reduces to

$$\begin{aligned} \mathcal{S}(\alpha, \boldsymbol{\pi}) &= \frac{2\Lambda^\varepsilon}{\Omega g} \int d\bar{\theta} d\theta dt \bar{D}\alpha(t, \theta) D\alpha(t, \theta) \left[L^d - \int d^d x \boldsymbol{\pi}^2(t, x, \theta) \right] \\ &+ \frac{\Lambda^\varepsilon}{g} \int d\bar{\theta} d\theta dt \int d^d x \left[\frac{2}{\Omega} \bar{D}\boldsymbol{\pi}(t, x, \theta) \cdot D\boldsymbol{\pi}(t, x, \theta) + \frac{1}{2} (\nabla \boldsymbol{\pi}(t, x, \theta))^2 \right]. \end{aligned} \quad (36.130)$$

Neglecting the non-zero modes, and using the RG, we immediately obtain the form of the relaxation time, at leading order for large L :

$$\tau(L, g) \sim \Omega'^{-1}(g) \frac{L^2}{g_L}, \quad \text{with } \Omega'(g) = \Omega \zeta^{-1}(g) \frac{N-1}{2}, \quad (36.131)$$

where $\zeta(g)$ is defined by equation (36.128).

The integral over π , expanded for α small, simply yields a renormalization of the leading order α action:

$$\frac{L^d}{2\Omega g} \mapsto \frac{L^2}{2\Omega} \left(\frac{L^{d-2}}{g} - \frac{N-2}{4\pi^2} \sum_{\mathbf{k} \neq \mathbf{0}} \frac{1}{\mathbf{k}^2} \right). \quad (36.132)$$

The relaxation time follows:

$$\frac{\Omega\tau(g, L)}{L^2} = \frac{2}{N-1} \left(\frac{L^{d-2}}{g} - \frac{N-2}{4\pi^2} \sum_{\mathbf{k} \neq \mathbf{0}} \frac{1}{\mathbf{k}^2} \right). \quad (36.133)$$

The sum has to be understood with a cut-off. As we have explained in Section 32.3.3, if we subtract to the sum its infinite-size limit, we obtain a finite result (equation (32.47)). We then introduce the size-dependent temperature g_L (equation (32.22) with $t \mapsto g$ and $\Lambda = 1$) and find

$$\frac{\tau(g, L)}{L^2} = \Omega'^{-1}(g) \left[\frac{1}{g_L} + \frac{\beta_2(d)}{d-2} - \frac{N-2}{4\pi} \int_0^\infty ds (\vartheta_0^d(s) - 1 - s^{-d/2}) \right], \quad (36.134)$$

where $\Omega'(g)$ defined in equation (36.131) and the function ϑ_0 is defined by equation (32.44). We have thus obtained the first correction to the leading term for $g < g_c$, where g_c is the critical temperature and the zero of the β -function. Again, this expression has a finite limit when $d \rightarrow 2$. Indeed,

$$\int_0^\infty ds (\vartheta_0^d(s) - 1 - s^{-d/2}) = -\frac{2}{d-2} + \int_0^\infty ds (\vartheta_0^2(s) - 1 - \theta(1-s)/s), \quad (36.135)$$

where $\theta(s)$ is the step function. Hence, we also obtain the form of the leading correction for $d = 2$ and $\xi(t)/L$ large. Finally, we can calculate the value at g_c in an ε -expansion, but the result is proportional to a time scale.

36.6.3 Dynamics in the ordered phase

We consider a purely dissipative dynamics associated with a static action of the form (A32.13), for example,

$$\mathcal{A}(\varphi) = \int d^d x \left[\frac{1}{2} (\nabla \varphi(x))^2 + g (\varphi^2(x) - M_0^2)^2 \right]. \quad (36.136)$$

We assume that our system evolves inside a hypercube of linear size L with periodic boundary conditions.

In Section 39.3, it is shown that the difference between the two first eigenvalues of the corresponding FP Hamiltonian is of the order of

$$e^{-\Delta\mathcal{H}}, \quad \text{with} \quad \Delta\mathcal{H} = \mathcal{H}_{\max} - \mathcal{H}_{\min},$$

in which \mathcal{A}_{\min} is the value of $\mathcal{H}(\varphi)$ at the degenerate minima, and \mathcal{H}_{\max} the value at the saddle point which separates them. We have already briefly analysed this problem in Section A34.3.

We start from a configuration in which φ is closed to M_0 and create a bubble of the phase $\varphi = -M_0$. Since the cost in energy is proportional to the area of the bubble, the saddle point corresponds to the situation in which the hypercube is evenly divided between the two phases. Due to the periodic boundary conditions, the minimal area surface, which evenly divides the hypercube, consists in two parallel sections perpendicular to the axes. Such a configuration corresponds to an instanton–anti-instanton pair of the static action. Denoting by σ the instanton action in one dimension, we thus find that the relaxation time τ_L , which is the inverse of the second eigenvalue, behaves like

$$\tau_L \propto e^{2\sigma L^{d-1}}. \quad (36.137)$$

Comparing with equation (A32.15), we find the simple relation between the relaxation time in a cubic geometry and the finite-size correlation length ξ_L in a cylindrical geometry,

$$\tau_L \propto \xi_L^2. \quad (36.138)$$

A36 RG functions at two loops

For illustration purpose, we calculate here, at two-loop order, the dynamic RG functions corresponding to two examples of a purely dissipative Langevin equation (35.1, 35.7), the $O(N)$ -symmetric $(\phi^2)^2$ field theory of Section 36.1, and the non-linear σ -model of Section 36.2. In the purely dissipative case, the form of the renormalized dynamical action shows that only one new RG function appears associated with the time-scale renormalization.

We then examine the problem of finite size effects in dynamics in the example of the hypercubic geometry.

A36.1 Supersymmetric perturbative calculations at two loops

In perturbative calculations of dynamic quantities, the formalism of Section 35.2.1 can be used, dimensional regularization eliminating the determinant. In the particular case of the dissipative Langevin equation, it is also possible to use the method of super-diagrams, treating the Grassmann coordinates in the same way as the usual commuting coordinates. Then, supersymmetry is explicit at all steps. Moreover, dynamic and static perturbation theories become remarkably similar, the topology and weight factors of Feynman diagrams being the same. It is also convenient to take the boundary condition in the Langevin equation at time $-\infty$, in such a way that the system is at equilibrium at any finite time and, therefore, time translation invariance is secured.

WT identities. Vertex functions, after Fourier transformation, satisfy the WT identities (35.41, 35.43) corresponding to supersymmetry transformations (34.79),

$$\sum_{i=1}^n \left(\frac{\partial}{\partial \bar{\theta}_i} - i\omega_i \theta_i \right) \tilde{\Gamma}^{(n)}(\boldsymbol{\theta}, \omega, p) = 0. \quad (\text{A36.1})$$

These identities provide checks in perturbative calculations.

A36.1.1 The $(\phi^2)^2$ field theory: Dynamic exponent

To the action (36.13) corresponds, in the Fourier representation, the propagator

$$\tilde{\Delta}(\boldsymbol{\theta}, \boldsymbol{\theta}', \omega, \mathbf{k}) = \frac{\Omega [1 - \frac{1}{2}i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}') + \frac{1}{4}\Omega(k^2 + m^2)\delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}')] }{\omega^2 + \frac{1}{4}\Omega^2(k^2 + m^2)^2}. \quad (\text{A36.2})$$

We have omitted the factor δ_{ij} corresponding to group indices. For practical calculations, it is actually often more convenient to use a mixed representation for the propagator using Fourier variables for space, but not for time:

$$\tilde{\Delta}(\boldsymbol{\theta}, \boldsymbol{\theta}', t, \mathbf{k}) = \left\{ 1 + \frac{1}{4}\Omega(k^2 + m^2)(\theta - \theta') [\bar{\theta} - \bar{\theta}' - (\bar{\theta} + \bar{\theta}') \operatorname{sgn}(t)] \right\} A(t, \mathbf{k}), \quad (\text{A36.3})$$

with

$$A(t, \mathbf{k}) = \frac{1}{k^2 + m^2} \exp \left[-\frac{1}{2}\Omega(k^2 + m^2)|t| \right],$$

in which $\operatorname{sgn}(t)$ is the sign function. Note that, in agreement with the analysis of Section 35.4.2 (equation (35.45)), $\tilde{\Delta}$ can also be written as

$$\tilde{\Delta}(\boldsymbol{\theta}, \boldsymbol{\theta}', t, \mathbf{k}) = \left\{ 1 + \frac{1}{2}(\theta - \theta') [\bar{\theta} + \bar{\theta}' - (\bar{\theta} - \bar{\theta}') \operatorname{sgn}(t)] \frac{\partial}{\partial t} \right\} A(t, \mathbf{k}).$$

This form follows directly from the supersymmetry and the causality of the Langevin equation (see also equations (36.17–36.20)).

From the supersymmetric form (36.13) of the dynamical action, we infer that the dynamic and the static theory have similar perturbative expansions and differ mainly by the form of the propagator. For $N = 1$, the interaction vertex $V^{(4)}$ has the form

$$V^{(4)} = m^\varepsilon \frac{g}{4!} \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3) \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_4). \quad (A36.4)$$

We have already calculated all static renormalization constants to two-loop order in Sections 10.5, 10.6. To determine the new renormalization constant at leading order (*i.e.* two-loop order), only the two-point function is required. Using the expressions of Section 10.6, we find

$$\begin{aligned} \tilde{\Gamma}^{(2)} &= -\frac{2}{\Omega} [2 - i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}')] + (k^2 + m^2) \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}') \\ &\quad + \frac{N+2}{6} m^\varepsilon g D_1 - \frac{(N+2)^2}{36} m^{2\varepsilon} g^2 D_2 - \frac{N+2}{18} m^{2\varepsilon} g^2 D_3, \end{aligned} \quad (A36.5)$$

in which the three diagrams D_1 , D_2 and D_3 are given by

$$D_1 = \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}') \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 + m^2}, \quad (A36.6)$$

$$\begin{aligned} D_2 &= D_1 \int \frac{d^d p}{(2\pi)^d} dt d^2 \bar{\theta}'' e^{-\Omega(p^2 + m^2)|t|} \frac{1}{p^2 + m^2} \\ &\quad \times \left\{ \frac{1}{p^2 + m^2} + \frac{\Omega}{2} \delta(\theta'' - \theta) [\bar{\theta}'' - \bar{\theta} - \text{sgn}(t)(\bar{\theta}'' + \bar{\theta})] \right\}, \end{aligned} \quad (A36.7)$$

$$\begin{aligned} D_3 &= \frac{1}{(2\pi)^{2d}} \int dt e^{i\omega t} \int d^d p_1 d^d p_2 e^{-\Omega s(p_i)|t|/2} \prod_{i=1}^3 \frac{1}{p_i^2 + m^2} \\ &\quad \times \left\{ 1 + \frac{\Omega}{4} s \{p_i\} \delta(\theta - \theta') [\bar{\theta} - \bar{\theta}' - \text{sgn}(t)(\bar{\theta} + \bar{\theta}')] \right\}, \end{aligned} \quad (A36.8)$$

with the definitions

$$\mathbf{p}_3 = -(\mathbf{k} + \mathbf{p}_1 + \mathbf{p}_2), \quad (A36.9)$$

$$s(p_i) = \sum_{i=1}^3 (p_i^2 + m^2). \quad (A36.10)$$

The diagrams D_1 and D_2 (after integration over t) are identical to the diagrams of static theory. Only D_3 contains a new dynamic divergence contributing to the renormalization of Ω . Actually, D_3 can be written as the sum of two terms, $D_3^{(1)}$ proportional to $\delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}')$, and which contains the static two-loop divergence, and another one $D_3^{(2)}$, which after an integration by parts over t , can be written as

$$\begin{aligned} D_3^{(2)} &= [1 - \frac{1}{2} i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}')] \int dt e^{i\omega t} \\ &\quad \times \int \frac{d^d p_1}{(2\pi)^d} \frac{d^d p_2}{(2\pi)^d} e^{-\Omega s(p_i)|t|/2} \prod_{i=1}^3 \frac{1}{p_i^2 + m^2}. \end{aligned} \quad (A36.11)$$

The divergent part of the integral is a constant which can be calculated for $\omega = 0$ and $\mathbf{k} = 0$. Finally, integrating over t , p_2 and p_1 successively, one obtains

$$Z_\omega = 1 - N_d^2 g^2 \ln(4/3) \frac{N+2}{24\varepsilon} + O(g^3), \quad (A36.12)$$

and, therefore, using the results of Section 10.6,

$$\eta_\omega(g) = N_d^2 \frac{N+2}{72} g^2 (6 \ln(4/3) - 1) + O(g^3). \quad (A36.13)$$

At the IR fixed point $g^* = 48\pi^2\varepsilon/(N+8) + O(\varepsilon^2)$ (see Section 15.7), one obtains in dimension $4-\varepsilon$ (equation (36.36)),

$$z = 2 + \eta_\omega = 2 + \frac{N+2}{2(N+8)^2} (6 \ln(4/3) - 1) \varepsilon^2 + O(\varepsilon^3).$$

A36.1.2 The non-linear σ -model

We now consider the $O(N)$ -symmetric non-linear σ -model, for $N > 2$ (as discussed in Chapter 19), still with a supersymmetric dynamics, to illustrate the discussion of Sections 35.6, 35.7, and 36.2. The dynamic bare action in a magnetic field, in terms of the superfield (see equation (36.48)), can be written as

$$\mathcal{S}(\phi_0) = \frac{1}{g_0} \int d\bar{\theta} d\theta dt \left[\int d^d x \frac{2}{\Omega_0} \bar{D}\phi_0(t, x, \boldsymbol{\theta}) \cdot D\phi_0(t, x, \boldsymbol{\theta}) + \mathcal{H}(\phi_0) \right], \quad (A36.14)$$

with

$$\mathcal{H}(\phi_0) = \int d^d x \left[\frac{1}{2} (\nabla \phi_0(t, x, \boldsymbol{\theta}))^2 - \mathbf{h}_0 \cdot \phi_0(t, x, \boldsymbol{\theta}) \right]. \quad (A36.15)$$

Note that, for practical reasons, we have adopted normalizations that differ from those of Sections 35.6 and 35.7.

The renormalized theory is defined by the substitutions,

$$\phi_0(t, x, \boldsymbol{\theta}) = \sqrt{Z} \{ \sigma(t, x, \boldsymbol{\theta}), \boldsymbol{\pi}(t, x, \boldsymbol{\theta}) \}, \quad (A36.16)$$

with

$$\sigma(t, x, \boldsymbol{\theta}) = \sqrt{Z^{-1} - (\boldsymbol{\pi}(t, x, \boldsymbol{\theta}))^2}, \quad (A36.17)$$

and

$$g_0 = \mu^{-\varepsilon} g Z_g, \quad \Omega_0 g_0 = \Omega g \mu^{-\varepsilon} Z / Z_\omega, \quad h_0 \sqrt{Z} / g_0 = h \mu^\varepsilon / g. \quad (A36.18)$$

We have denoted by μ the renormalization scale and set $\varepsilon = d - 2$.

In the framework of dimensional regularization and MS, the dynamic RG function $\eta_\omega(g)$ then is given by

$$\eta_\omega(g) = \mu \frac{d}{d\mu} \Big|_{g_0, \Omega_0} \ln \Omega = \beta(g) \frac{d}{dg} \ln(Z_\omega Z_g / Z). \quad (A36.19)$$

Perturbation theory. The propagator is the same as in equation (A36.2) up to a global normalization:

$$\tilde{\Delta}(\boldsymbol{\theta}, \boldsymbol{\theta}', \omega, \mathbf{k}) = g \mu^{-\varepsilon} \frac{\Omega [1 - \frac{1}{2} i \omega (\theta - \theta') (\bar{\theta} + \bar{\theta}') + \frac{1}{4} \Omega (k^2 + h) \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}')] }{\omega^2 + \frac{1}{4} \Omega^2 (k^2 + h)^2}. \quad (A36.20)$$

To calculate the two-point function at two-loop order, we need the π^4 and the π^6 vertices:

$$V^{(4)} = \frac{\mu^\varepsilon}{8g} \delta_{i_1 i_2} \delta_{i_3 i_4} \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) \delta^2(\boldsymbol{\theta}_3 - \boldsymbol{\theta}_4) \\ \times \left\{ \frac{2}{\Omega} \left[-2 + i(\omega_1 + \omega_2)(\theta_1 - \theta_3)(\bar{\theta}_1 + \bar{\theta}_3) \right] + ((p_1 + p_2)^2 + h) \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3) \right\}, \quad (A36.21)$$

$$V^{(6)} = \frac{\mu^\varepsilon}{16g} \delta_{i_1 i_2} \delta_{i_3 i_4} \delta_{i_5 i_6} \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) \delta^2(\boldsymbol{\theta}_3 - \boldsymbol{\theta}_4) \delta^2(\boldsymbol{\theta}_3 - \boldsymbol{\theta}_5) \delta^2(\boldsymbol{\theta}_3 - \boldsymbol{\theta}_6) \\ \times \left\{ \frac{2}{\Omega} \left[-2 + i(\omega_1 + \omega_2)(\theta_1 - \theta_3)(\bar{\theta}_1 + \bar{\theta}_3) \right] + \delta^2(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3) ((p_1 + p_2)^2 + h) \right\}. \quad (A36.22)$$

Two-point function at one-loop order. The values of the two diagrams of Fig. 19.1 are $\frac{1}{2}(N-1)D_1$ and D_2 , with

$$D_1 = \mu^\varepsilon h \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}') I(h), \quad (A36.23)$$

$$D_2 = \mu^\varepsilon \left\{ -\frac{2}{\Omega} \left[2 - i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}') \right] + k^2 \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}') \right\} I(h), \quad (A36.24)$$

and we recall (equations (19.124, 18.16)),

$$I(h) = \Omega_d(\sqrt{h}) = \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 + h} = N_d \left(-\frac{1}{\varepsilon} - \frac{1}{2} \ln h \right) + O(\varepsilon). \quad (A36.25)$$

The two-point function at one-loop order follows:

$$\frac{g}{\mu^\varepsilon} \tilde{\Gamma}^{(2)} = -\frac{2Z_\omega}{\Omega} \left[2 - i\omega(\theta - \theta')(\bar{\theta} + \bar{\theta}') \right] + \left(\frac{Z}{Z_g} k^2 + h \sqrt{Z} \right) \delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}') \\ + \frac{1}{2}(N-1)gD_1 + gD_2 + O(g^2). \quad (A36.26)$$

We recover the one-loop static results,

$$Z = 1 + (N-1)g/(2\pi\varepsilon) + O(g^2), \quad Z_g = 1 + (N-2)g/(2\pi\varepsilon) + O(g^2),$$

and in addition, we find

$$Z_\omega = Z/Z_g + O(g^2). \quad (A36.27)$$

This equation together with equation (A36.19) implies the absence of dynamics renormalization at this order:

$$\eta_\omega(g) = O(g^2). \quad (A36.28)$$

Two-loop calculation. We now have to calculate the one-loop diagrams with the renormalization constants expanded at one-loop order, and the various two-loop diagrams generated by two four-point vertices and one six-point vertex, as displayed in Fig. 19.2. Since the static renormalization constants have already been calculated, to calculate $\eta_\omega(g)$ at two-loop order, we need only $\tilde{\Gamma}^{(2)}$ for vanishing arguments: $\mathbf{k} = 0, \omega = 0, \boldsymbol{\theta} = 0$. Then, the corresponding contributions are

$$\begin{bmatrix} 2(-4I^2), & (N-1)(-4I^2), & 0, & 0, & -(-4I^2 + I/\pi), \\ -\frac{1}{2}(N-1)(-I/\pi), & 0, & 0, & -(-8I^2 + 16J), & -\frac{1}{2}(N-1)(4I^2 - 32J). \end{bmatrix}$$

The integral J is given by

$$\begin{aligned} J &= \frac{1}{(2\pi)^{2d}} \int \frac{d^d p d^d q}{(p^2 + h)(p^2 + q^2 + (p+q)^2 + 3h)} \\ &= \frac{N_d^2}{4\varepsilon^2} \left(1 + \frac{1}{2}\varepsilon \ln(3/4) + O(\varepsilon^2) \right). \end{aligned} \quad (A36.29)$$

Therefore, we obtain,

$$\begin{aligned} \tilde{\Gamma}^{(2)} &= -\frac{4Z_\omega}{\Omega} \left[1 + gZ_g (hZ_g/Z)^{\varepsilon/2} \Omega_d + \frac{1}{2}(3N-5)\Omega_d^2 g^2 J g^2 \right. \\ &\quad \left. - 4(N-2)Jg^2 - \frac{1}{4}(N-3)N_d\Omega_d g^2 + O(g^3) \right]. \end{aligned} \quad (A36.30)$$

Finally, expanding all terms, we obtain the expression of Z_ω at two-loop order:

$$Z_\omega = 1 + \frac{g}{2\pi\varepsilon} + \left(\frac{N-1}{2\varepsilon^2} \right) \frac{g^2}{(2\pi)^2} - \frac{N-2}{2\varepsilon} \ln \frac{4}{3} \frac{g^2}{(2\pi)^2} + O(g^3). \quad (A36.31)$$

The expansion of $\eta_\omega(g)$ at order g^2 follows:

$$\eta_\omega(g) = (N-2)[1 - \ln(4/3)] \frac{g^2}{(2\pi)^2} + O(g^3). \quad (A36.32)$$

In dimension $d = 2 + \varepsilon$, at the critical and UV fixed point $g^* = 2\pi\varepsilon/(N-2) + O(\varepsilon^2)$ (Section 19.12), one obtains the result (36.57)

$$z = 2 + (1 - \ln(4/3)) \frac{\varepsilon^2}{N-2} + O(\varepsilon^3).$$

37 Instantons in quantum mechanics (QM)

In general, we calculate Euclidean functional integrals by the steepest-descent method always looking, in the absence of external sources, for saddle points in the form of constant solutions to the classical field equations. However, classical field equations may have non-constant solutions. In Euclidean stable field theories, non-constant solutions always have an action larger than the action of minimal constant solutions, because the gradient term gives an additional positive contribution.

In what follows, we are mainly interested in the structure of the ground state, and thus in the zero-temperature limit of the partition function. For a given constant solution, we will focus on the non-constant solutions whose relative action remains finite in this limit. These solutions are called *instanton* solutions, and are the saddle points relevant for a calculation, by the steepest-descent method, of *barrier penetration effects* [373, 374]. In this chapter, we consider the simple example of non-relativistic QM, where instanton calculus is an alternative to the semi-classical Wentzel–Kramers–Brillouin (WKB) method but, in the coming chapters, we show how the instanton method can be generalized to field theory.

We explain the role of instantons in some metastable systems in QM. In particular, we show that instantons determine, in the semi-classical limit, the decay rate of metastable states initially confined in a relative minimum of a potential and decaying through barrier penetration.

Using the technical tools developed in Ref. [375], we first discuss the quartic anharmonic oscillator with negative coupling and calculate the contributions of instantons at leading order. We then generalize the method to a large class of analytic potentials, and obtain explicit expressions, at leading order, for one-dimensional systems.

In the appendix, we give an exact expression for the Jacobian, due to collective coordinates, in the case of path integrals. We describe how semi-classical expressions can be derived from calculations based on the WKB method.

Finally, let us point out that, although we only deal here with Euclidean theories, many aspects of the techniques we describe also apply to the calculation of effects coming from finite energy solutions of the real-time field equations, called *soliton* solutions in the literature.

37.1 The quartic anharmonic oscillator for negative coupling

We consider the quantum Hamiltonian of the quartic anharmonic oscillator (2.74),

$$H = -\frac{1}{2} (\mathrm{d}/\mathrm{d}q)^2 + \frac{1}{2} q^2 + \frac{1}{4} g q^4, \quad (37.1)$$

where, initially, g is a positive parameter. The ground state energy $E_0(g)$ can be obtained from the large β limit of the partition function $\mathrm{tr} e^{-\beta H}$ (β is the inverse temperature),

$$E_0(g) = \lim_{\beta \rightarrow +\infty} -\frac{1}{\beta} \ln \mathrm{tr} e^{-\beta H}.$$

Moreover, a systematic expansion of the partition function for β large also yields the whole spectrum (for rigorous results concerning the spectrum, see Refs. [376]).

Since the partition function has the path integral representation (equation (2.33)),

$$\text{tr } e^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [dq(t)] \exp [-\mathcal{S}(q(t))], \quad (37.2)$$

where $\mathcal{S}(q)$ is the Euclidean action ($\dot{q} \equiv dq/dt$):

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\frac{1}{2} \dot{q}^2(t) + \frac{1}{2} q^2(t) + \frac{1}{4} g q^4(t) \right] dt, \quad (37.3)$$

we can use it to calculate the eigenvalues of H .

A generalization of arguments applicable to finite dimensional integrals, alternative to methods based on the Schrödinger equation [377], indicates that the path integral (37.2) defines a function of g analytic in the half plane $\text{Re}(g) > 0$. In this domain, for g small, the path integral is dominated by the saddle point $q(t) \equiv 0$. Therefore, it can be calculated by expanding the integrand in powers of g and integrating term by term (Section 2.7). This generates the perturbative expansion of the partition function, and thus of the ground state $E_0(g)$ by taking the large β limit.

Remarks

(i) We always expand in g before taking the large β limit. Since $E_N(g)$, the N th eigenvalue of H , satisfies

$$E_N(g) = N + \frac{1}{2} + O(g),$$

the perturbative expansion can be written as

$$\text{tr } e^{-\beta H} = \sum_{N=0} e^{-\beta E_N(g)} = \sum_{N=0} e^{-(N+1/2)\beta} \sum_{k=0} \frac{1}{k!} (-\beta)^k \left(E_N - \frac{1}{2} - N \right)^k. \quad (37.4)$$

We note that $E_N(g)$ can be inferred from the coefficient of $e^{-(N+1/2)\beta}$, that the coefficient of g^k is a polynomial of degree k in β .

(ii) As we have already mentioned when discussing the ϕ^4 field theory, by the rescaling

$$q(t) \mapsto q(t)g^{-1/2},$$

we factorize the whole dependence in g in front of the action:

$$\mathcal{S}(q, g) = \frac{1}{g} \mathcal{S}(q\sqrt{g}). \quad (37.5)$$

The coupling constant g plays the same formal role as \hbar in the semi-classical or loop expansion. Thus, for $g \rightarrow 0$, the integral can be evaluated by the steepest-descent method.

Continuation to negative coupling. For $g < 0$, the Hamiltonian is unbounded from below for all values of g . Therefore, the energy levels, considered as analytic functions of g , must have a singularity at $g = 0$: the perturbative expansion in powers of g is always a divergent series [377].

A quantum state, initially localized at time $t = 0$ (t is here the *real physical time* of the Schrödinger equation) in the well of the potential near $q = 0$, then decays due to barrier penetration. To determine the decay rate, in the semi-classical approximation, we can use the following method: we calculate the ground state energy E_0 , and the corresponding time-dependent wave function $\psi_0(t)$ for g positive. The time-dependence of the solution $\psi_0(t, q)$ of the Schrödinger equation is

$$\psi_0(t, q) \propto e^{-iE_0 t}.$$

We then proceed by analytic continuation in the complex g plane from $g > 0$ to $g < 0$, in the direction such that $\text{Im } E$ remains negative. After analytic continuation, E_0 becomes complex, and thus $\|\psi_0(t)\|$ decreases exponentially with time at a rate

$$\|\psi_0(t)\| \sim e^{-|\text{Im } E_0|t}.$$

$|\text{Im } E_0|$ is the inverse lifetime of the wave function $\psi_0(t, q)$. Actually, the decay of $\psi(t, q)$ also involves the imaginary parts of the continuations of all excited states. However, we expect on intuitive grounds that, when the real part of the energy increases, the corresponding lifetime decreases (this can easily be verified by examples). Thus, at large times, only the component corresponding to the ground state survives. Therefore, hereafter we calculate $\text{Im } E_0$ for g small and negative.

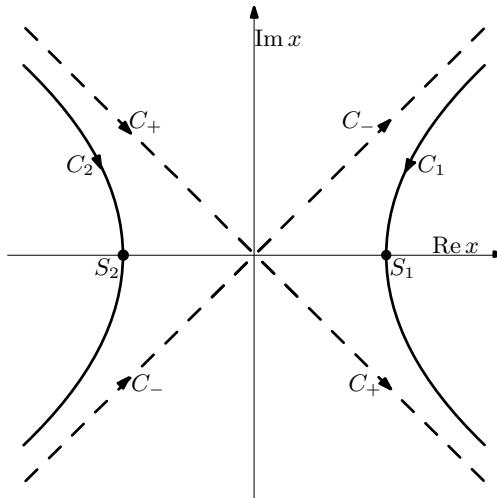


Fig. 37.1 The contours of integration C_+ , C_- , C_1 , and C_2

37.2 A toy model: A simple integral

To give an idea of how $E_0(g)$ can be defined and evaluated for g negative, we consider a simple integral with an analogous structure: the ‘zero-dimensional ϕ^4 field theory’. The coefficients of the expansion of the integral

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x^2/2+gx^4/4)} dx, \quad (37.6)$$

in powers of g , count the number of Feynman diagrams contributing to the partition function (or vacuum amplitude) in the general ϕ^4 field theory.

The integral defines the function for $\text{Re } g \geq 0$, but the function $I(g)$ is analytic in a cut plane. Its analytic continuation to $\text{Re } g < 0$ can be obtained by rotating the contour of integration C in the complex plane as one changes the argument of g :

$$C : \arg x = -\frac{1}{4} \arg g \pmod{\pi},$$

so that $\text{Re}(gx^4)$ remains positive.

In this way, one obtains two different expressions for $I(g)$ depending on the direction of rotation in the g -plane:

$$\begin{aligned} \text{for } g = -|g| + i0 : I(g) &= \int_{C_+} e^{-(x^2/2+gx^4/4)} dx, \text{ with } C_+ : \arg x = -\frac{\pi}{4} \pmod{\pi}, \\ \text{for } g = -|g| - i0 : I(g) &= \int_{C_-} e^{-(x^2/2+gx^4/4)} dx, \text{ with } C_- : \arg x = \frac{\pi}{4} \pmod{\pi}. \end{aligned}$$

For g positive and small, the integral is dominated by the saddle point at $x = 0$, and $I(g) = 1 + O(g)$. For $g \rightarrow 0_-$, the two integrals are still dominated by the saddle point at the origin, since the contribution of the other saddle points:

$$x + gx^3 = 0 \Rightarrow x^2 = -1/g \quad (37.7)$$

are of the order of

$$e^{-(x^2/2+gx^4/4)} \sim e^{1/4g} \ll 1. \quad (37.8)$$

However, the discontinuity of $I(g)$ on the cut is given by the difference between the two integrals:

$$I(g+i0) - I(g-i0) = 2i \operatorname{Im} I(g) = \frac{1}{\sqrt{2\pi}} \int_{C_+ - C_-} e^{-(x^2/2+gx^4/4)} dx. \quad (37.9)$$

It corresponds to the contour $C_+ - C_-$ which, as Fig. 37.1 shows, can be deformed into the sum of the contours C_1 and C_2 , which are dominated by the lower non-trivial saddle points S_1 and S_2 : $x = \pm 1/\sqrt{-g}$. This implies that the contribution of the saddle point at $x = 0$ cancels. The contributions of the saddle points S_1 and S_2 then yields

$$\operatorname{Im} I(g) \sim 2^{-1/2} e^{1/4g}. \quad (37.10)$$

Thus, for g negative and small, the real part of the integral is given by perturbation theory, while the exponentially small imaginary part is given by the contribution of non-trivial saddle points.

37.3 QM: Instantons

The method used in the example of the integral (37.6) can be adapted to the path integral (37.2). We rotate the contour in the functional $q(t)$ space, as we change the argument g from g positive to g negative:

$$q(t) \mapsto q(t) e^{-i\theta},$$

in which θ is time independent. Returning to the definition of the path integral as a limit of integrals in a discretized time (see Chapter 2), one can verify that this procedure makes sense.

However, there is one difference with respect to the simple integral: the contour has to stay within the domain in which $\operatorname{Re} [\dot{q}^2(t)] > 0$, since, as we have discussed in Chapter 2, the kinetic part $\int \dot{q}^2(t) dt$ selects continuous paths and ensures, therefore, the existence of the continuum limit of the discretized path integral.

For g negative, we thus integrate along a path

$$\arg q(t) = -\theta, \quad \text{with } \frac{1}{8}\pi < \theta < \frac{1}{4}\pi, \quad (37.11)$$

which satisfies the two conditions

$$\operatorname{Re} [gq^4(t)] > 0, \quad \operatorname{Re} [\dot{q}^2(t)] > 0. \quad (37.12)$$

For $g \rightarrow 0_-$, the path integrals corresponding to the two analytic continuations are still dominated by the saddle point at the origin $q(t) = 0$ but, in the difference, this contribution cancels. We have to look for non-trivial saddle points, which are solutions of the Euclidean classical equation of motion for $g < 0$,

$$-\ddot{q}(t) + q(t) + gq^3(t) = 0, \quad (37.13)$$

$$\text{with } q(-\beta/2) = q(\beta/2). \quad (37.14)$$

The contribution of the constant saddle point $q^2(t) = -1/g$ is of the order of $e^{\beta/4g}$ and, therefore, negligible in the large β limit. We have to look for solutions that have an action that remains finite for $\beta \rightarrow +\infty$. These are called *instantons*.

The solutions of equations (37.13, 37.14) correspond to a periodic motion in *real-time* in the potential $-V(q)$

$$V(q) = \frac{1}{2}q^2 + \frac{1}{4}gq^4. \quad (37.15)$$

Solutions exist which correspond to oscillations around each of the minima $q = \pm\sqrt{-1/g}$ of $-V$. Integrating once equation (37.13), one obtains

$$\frac{1}{2}\dot{q}^2(t) - \frac{1}{2}q^2(t) - \frac{1}{4}gq^4(t) = \epsilon,$$

where the constant ϵ is negative.

Denoting by q_0 and q_1 the points, with $q > 0$, where the velocity \dot{q} vanishes, one finds for the period of such a solution,

$$\beta = 2 \int_{q_0}^{q_1} \frac{dq}{\sqrt{q^2 + \frac{1}{2}gq^4 + 2\epsilon}}.$$

β can only become large if the constant ϵ , and thus q_0 go to 0. With increasing β , the classical trajectory comes closer to the origin. In the infinite β limit, the classical solution becomes

$$q_c(t) = \pm \left(-\frac{2}{g} \right)^{1/2} \frac{1}{\cosh(t - t_0)}. \quad (37.16)$$

The corresponding classical action is

$$\mathcal{S}(q_c) = -\frac{4}{3g} + O(e^{-\beta}/g). \quad (37.17)$$

Since the Euclidean action is invariant under time translations, the classical solution depends on a free parameter t_0 , which, for β , finite varies between 0 and β : $0 \leq t_0 < \beta$. Therefore, in contrast to the simple integral, we do not find two degenerate saddle points, but *two one-parameter families*.

We could have also considered trajectories oscillating n times around $q^2 = -1/g$ in the time interval β . It is easy to verify that the corresponding action in the infinite β limit becomes

$$\mathcal{S}(q_c) = -n \frac{4}{3g}, \quad (37.18)$$

and yields, therefore, a contribution proportional to $e^{n4/3g}$. For g small, the path integral is dominated by the term $n = 1$. Similarly, trajectories with $\epsilon > 0$ degenerate into the sum of two solutions with an action $-n8/3g$.

Remark. Although we emphasize the role of finite action configurations, the action corresponding to the paths that contribute to the path integral is always infinite, because the paths are not differentiable and, thus, the kinetic term diverges. However, the leading configurations are close to the saddle points.

37.4 Instanton contributions at leading order

The Gaussian approximation. To evaluate the contribution of the saddle points at leading order, the usual strategy consists in expanding the action around a saddle point, setting

$$q(t) = q_c(t) + r(t),$$

and calculating for $g \rightarrow 0_-$, $\beta \rightarrow \infty$ the Gaussian integral (one-loop order):

$$\begin{aligned} \text{Im tr } e^{-\beta H} &= \frac{1}{i} e^{4/3g} \int [dr(t)] \exp \left[-\frac{1}{2} \int dt (\dot{r}^2(t) + r^2(t) + 3gq_c^2(t)r^2(t)) \right], \\ &\equiv \frac{1}{i} e^{4/3g} \int [dr(t)] \exp \left(-\frac{1}{2} \int dt_1 dt_2 r(t_1) M(t_1, t_2) r(t_2) \right), \end{aligned}$$

where M is the differential operator,

$$M(t_1, t_2) = \left. \frac{\delta^2 \mathcal{S}}{\delta q(t_1) \delta q(t_2)} \right|_{q=q_c} = \left[-(d_{t_1})^2 + 1 + 3gq_c^2(t_1) \right] \delta(t_1 - t_2). \quad (37.19)$$

The path integral is normalized by dividing it by the partition function of the harmonic oscillator.

The zero mode. Differentiating equation (37.13) with respect to t , one finds

$$-(d_t)^2 \dot{q}_c(t) + \dot{q}_c(t) + 3gq_c^2(t)\dot{q}_c(t) \equiv [M\dot{q}](t) = 0. \quad (37.20)$$

Since the function $\dot{q}_c(t)$ is square integrable, this equation implies that $\dot{q}_c(t)$ is an eigenvector of M with eigenvalue 0. Hence, the naive Gaussian approximation yields a result proportional to $(\det M)^{-1/2}$, which is infinite!

The problem should have been expected: as we have noted previously, translation invariance in time implies the existence of two one-parameter families of continuously connected degenerate saddle points. An infinitesimal variation of $q(t)$ that corresponds to a variation of the parameter t_0 , that is, proportional to \dot{q}_c leaves the action unchanged. The problem that we face here is by no means special to path integrals, as the following example shows.

Zero modes in finite-dimensional integrals. We consider the integral,

$$I_2(g) = \int_{\mathbb{R}^\nu} d^\nu \mathbf{x} e^{\mathbf{x}^2 - g(\mathbf{x}^2)^2}, \quad \text{with } g > 0, \quad (37.21)$$

in which \mathbf{x} is a ν -component vector ($\nu > 1$), and the integrand is $O(\nu)$ invariant. For g small, this integral can be calculated by the steepest-descent method. The saddle points are given by

$$\mathbf{x}_c (1 - 2g\mathbf{x}_c^2) = 0, \quad (37.22)$$

and, since $\mathbf{x}_c = 0$ corresponds to a minimum, the relevant solutions are $|\mathbf{x}_c| = (2g)^{-1/2}$. We find a $(\nu - 1)$ parameter family of degenerate saddle points, since the saddle point equation determines only the length of the vector \mathbf{x}_c . If we single out one saddle point and evaluate its contribution in the Gaussian approximation, we are led to calculate the determinant of the matrix

$$M_{\alpha\beta} = 8gx_\alpha x_\beta, \quad (37.23)$$

which is the projector on \mathbf{x} and has, therefore, $(\nu - 1)$ vanishing eigenvalues.

Here, it is clear how to solve the problem: it is necessary to factorize the integration measure into a measure corresponding to angular variables, and a measure for the integration over the radial variable. The integration over angular variables has to be done exactly; only the integral over the radial variable can be evaluated by the steepest-descent method.

Similarly, in the case of the path integral, it is necessary to factorize the integration measure over the integration constants that parametrize the saddle points, in the preceding example, the time-translation constant. Then, the integration over these parameters has to be done exactly. The integration over the other path modes can be done by the steepest-descent method. This is the method of so-called *collective coordinates* [378].

Remark. We have already studied theories invariant under a continuous symmetry group in which the classical minimum is not invariant under the group, for example, the $O(N)$ -symmetric $(\phi^2)^2$ field theory in the ordered phase:

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} r \phi^2(x) + \frac{1}{4!} g(\phi^2(x))^2 \right], \quad \text{for } r < 0. \quad (37.24)$$

In such a situation, we have generally chosen one classical minimum, and made a systematic expansion around it. However, this procedure is justified only if the *symmetry is spontaneously broken*. Actually, we have noted in Chapter 19 that the absence of symmetry breaking manifests itself, in perturbation theory, by the appearance of infrared singularities. In the case of instanton solutions, the propagator M^{-1} in an instanton background has an isolated pole at the origin, which also leads to divergences in perturbation theory. We conclude that time-translation symmetry is not spontaneously broken, and that it is necessary to sum over all degenerate saddle points (see also Chapter 14).

37.4.1 Collective coordinates and Gaussian integration

The problem of factorizing the measure corresponding to the time-translation variable (a *collective coordinate*) is slightly more subtle than in the case of simple integrals, because a path corresponds to an infinite number of variables. A method, inspired by the Faddeev–Popov quantization method of gauge theories [219] can be used. We call t_0 the time-collective coordinate and start from the identity

$$1 = \int \frac{dt_0}{\sqrt{2\pi\xi}} \left[\int dt \dot{q}_c(t) \dot{q}(t+t_0) \right] \exp \left\{ -\frac{1}{2\xi} \left[\int dt \dot{q}_c(t) (q(t+t_0) - q_c(t)) \right]^2 \right\}, \quad (37.25)$$

which one can verify by changing variables,

$$t_0 \mapsto \lambda, \quad \text{with } \lambda = \int dt \dot{q}_c(t) (q(t+t_0) - q_c(t)).$$

The constant ξ has been introduced mainly for cosmetic reasons, but is of order g .

We insert the identity (37.25) into the path integral. The new action

$$\mathcal{S}(q) + \frac{1}{2\xi} \left[\int dt \dot{q}_c(t) (q(t+t_0) - q_c(t)) \right]^2$$

is no longer time-translation invariant. It leads to the saddle point equation

$$\frac{\delta \mathcal{S}}{\delta q(t)} + \frac{1}{\xi} \dot{q}_c(t-t_0) \int dt' \dot{q}_c(t'-t_0) (q(t') - q_c(t'-t_0)) = 0. \quad (37.26)$$

The equation is clearly satisfied for $q(t) = q_c(t - t_0)$. The determinant generated by the Gaussian integration around the saddle point is the determinant of the modified operator,

$$M'(t_1, t_2) = M(t_1, t_2) + \frac{1}{\xi} \dot{q}_c(t_1 - t_0) \dot{q}_c(t_2 - t_0),$$

or in the bra–ket notation of QM,

$$M' = M + \mu |1\rangle \langle 1|, \quad (37.27)$$

where we have denoted by $|1\rangle$ the eigenvector proportional to \dot{q}_c with unit norm, and μ is given by

$$\mu = \|\dot{q}_c\|^2 / \xi.$$

All the eigenvalues of the operators M' and M (equation (37.19)) are the same, except one: the eigenvalue that corresponds to the eigenvector \dot{q}_c is $\|\dot{q}_c\|^2 / \xi$, instead of 0.

To normalize the path integral, we compare it to its value at $g = 0$, which is the partition function $\mathcal{Z}(\beta)$ of the harmonic oscillator, and which, in the large β limit, reduces to $e^{-\beta/2}$. At $g = 0$, the operator M reduces to the operator

$$M_0(t_1, t_2) = \left[-(d_{t_1})^2 + 1 \right] \delta(t_1 - t_2). \quad (37.28)$$

As we explain in Section 37.4.2, a well-defined quantity that can be calculated is the determinant of the ratio of operators $\det(M + \varepsilon)(M_0 + \varepsilon)^{-1}$, where ε is an arbitrary constant. For $\varepsilon \rightarrow 0$, it vanishes like ε , and we thus set

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \det(M + \varepsilon)(M_0 + \varepsilon)^{-1} \equiv \det' MM_0^{-1}. \quad (37.29)$$

By contrast, the quantity that one needs to evaluate has the form

$$\det(M + \mu |1\rangle \langle 1|) M_0^{-1} = \lim_{\varepsilon \rightarrow 0} \det(M + \varepsilon + \mu |1\rangle \langle 1|)(M_0 + \varepsilon)^{-1}.$$

Then,

$$\begin{aligned} \det(M + \varepsilon + \mu |1\rangle \langle 1|)(M_0 + \varepsilon)^{-1} &= \det(M + \varepsilon)(M_0 + \varepsilon)^{-1} \\ &\quad \times \det(1 + \mu |1\rangle \langle 1| (M + \varepsilon)^{-1}) \\ &= \det(M + \varepsilon)(M_0 + \varepsilon)^{-1} (1 + \mu/\varepsilon). \end{aligned}$$

One thus finds

$$\det' MM_0^{-1} \|\dot{q}_c\|^2 / \xi.$$

Note that all functions depend only on $t - t_0$ and, therefore, t_0 can be eliminated from the determinant. In the first factor in (37.25), at leading order, we replace $q(t + t_0)$ by $q_c(t)$. The integral does not depend on t_0 anymore, and we find the factor

$$\frac{1}{\sqrt{2\pi\xi}} \beta \|\dot{q}_c\|^2.$$

Therefore, the result of the integration over the fluctuations around the saddle point is

$$\frac{\beta}{\sqrt{2\pi}} \mathcal{Z}_0(\beta) \|\dot{q}_c\| (\det' MM_0^{-1})^{-1/2}.$$

Taking into account the two families of saddle points and dividing by $2i$ to obtain the imaginary part, we find for $\beta \rightarrow \infty$:

$$\text{Im tr } e^{-\beta H} \sim \frac{2}{2i} \frac{\beta}{\sqrt{2\pi}} e^{-\beta/2} \|\dot{q}_c\| \left[\det' M (\det M_0)^{-1} \right]^{-1/2} e^{4/3g}. \quad (37.30)$$

37.4.2 The result at leading order

In Section 37.6.1, we calculate the determinant for a general, analytic potential in one-dimensional systems and, in Section A37.2, we compare this calculation with the corresponding WKB calculation. We show indirectly that, for all systems for which we can solve explicitly the classical equations of motion with arbitrary boundary conditions, we can also explicitly calculate the determinant of the operator governing the small fluctuations around the classical trajectory. In the special case considered here, M is a Hamiltonian with a Bargmann potential, whose spectrum is known exactly, and one finds [379],

$$\det(M + \varepsilon)(M_0 + \varepsilon)^{-1} = \frac{\sqrt{1+\varepsilon}-1}{\sqrt{1+\varepsilon}+1} \frac{\sqrt{1+\varepsilon}-2}{\sqrt{1+\varepsilon}+2}. \quad (37.31)$$

General arguments show that the ground-state wave function has no node, the wave function of the first excited state one node, and so on. The wave function $\dot{q}_c(t)$ vanishes once at the turning point and corresponds to the first excited state. Therefore, M has one negative eigenvalue corresponding to the ground state, as expression (37.31) confirms.

Then,

$$\det' M (\det M_0)^{-1} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \det(M + \varepsilon)(M_0 + \varepsilon)^{-1} = -\frac{1}{12}. \quad (37.32)$$

The square root of the determinant thus is imaginary, and the final result is real, as expected. However, the sign of the square root of expression (37.31) can only be resolved by following the analytic continuation from g positive to g negative.

The norm $\|\dot{q}_c\|$ is easily calculable. It has an important property: it is proportional to $1/\sqrt{g}$. As we shall note later, this is the first example of a general situation: when the instanton solution breaks a continuous symmetry of the classical action, the solution depends on parameters generated by the action of the symmetry group on the solution. Each parameter has to be taken as an integration variable, and the corresponding Jacobian generates as a factor the loop expansion parameter to the power $-1/2$. Here, we find

$$\|\dot{q}_c\| = \frac{2}{\sqrt{3}} \frac{1}{\sqrt{-g}}. \quad (37.33)$$

The expression then becomes

$$\text{Im} \text{tr } e^{-\beta H} = -\beta e^{-\beta/2} \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{-g}} e^{4/3g} [1 + O(g, e^{-\beta})], \text{ for } g \rightarrow 0_-, \beta \rightarrow \infty. \quad (37.34)$$

For β large, the left-hand side has the form

$$\text{Im} \text{tr } e^{-\beta H} \sim \text{Im} e^{-\beta E_0(g)} \equiv \text{Im} e^{-\beta(\text{Re } E_0(g) + i \text{Im } E_0(g))}, \text{ for } g \rightarrow 0_-, \beta \rightarrow \infty. \quad (37.35)$$

For g -small, the imaginary part of E_0 is exponentially small. Since the small g limit has always to be taken before the large β limit, we can write

$$\text{Im} \text{tr } e^{-\beta H} \sim -\beta \text{Im}(E_0(g)) e^{-\beta \text{Re } E_0(g)} \sim -\beta e^{-\beta/2} \text{Im } E_0. \quad (37.36)$$

Equation (37.34) then leads to

$$\text{Im } E_0(g) = \frac{4}{\sqrt{2\pi}} \frac{e^{4/3g}}{\sqrt{-g}} [1 + O(g)], \quad g \rightarrow 0_-. \quad (37.37)$$

Remark. We have derived the behaviour of the imaginary part of the ground state energy for g small and negative and, therefore, the decay rate of a metastable state localized in the unbounded potential corresponding to the anharmonic oscillator with negative coupling. In Section 40.1.1, we derive from this result an evaluation of the large-order behaviour of the perturbation series for the quartic anharmonic oscillator.

37.5 General analytic potentials: Instanton contributions

WE now generalize the methods described in previous sections to a general class of one-dimensional analytic potentials. We calculate, at leading order in the semi-classical limit, the decay rate of a quantum state located at initial time around a relative minimum of a potential and decaying through barrier penetration.

To guide the intuition, we imagine that we start from a situation in which a given minimum of a potential is an absolute minimum, and after an analytic continuation, becomes a relative minimum of the potential. As we have argued in Section 37.1, the corresponding ground state energy becomes complex in the analytic continuation, and its imaginary part yields the inverse lifetime of a state initially concentrated around the relative minimum of the potential. In the semi-classical limit, the imaginary part is again related to finite action, that is, instanton solutions of the Euclidean classical equations of motion.

The instanton solution. We consider Hamiltonians of the form [380],

$$H = -\frac{1}{2} (\mathrm{d}/\mathrm{d}q)^2 + g^{-1}V(q\sqrt{g}), \quad (37.38)$$

where $V(q)$ is an analytic function of q , which, for q small, behaves like

$$V(q) = \frac{1}{2}q^2 + O(q^3). \quad (37.39)$$

We assume that $q = 0$ corresponds to a *relative minimum* of the potential.

Again, in the Hamiltonian (37.38) the potential has been parametrized in such a way that g plays the formal role of \hbar and is a loop expansion parameter.

The path integral representation of the partition function is

$$\mathrm{tr} e^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q(t)] \exp[-\mathcal{S}(q)], \text{ with} \quad (37.40)$$

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\frac{1}{2}\dot{q}^2(t) + g^{-1}V(q(t)\sqrt{g}) \right] \mathrm{d}t. \quad (37.41)$$

In the situation that we are considering, we know that instanton solutions exist. Because $q = 0$ is only a relative minimum of the potential, the function $V(q)$, which we have assumed to be analytic and thus continuous has at least another zero. For β infinite, an instanton solution $q_c(t)$ starts from the origin at time $-\infty$, is reflected on a zero of the potential and returns to the origin at time $+\infty$.

A variation of $\mathcal{S}(q)$ yields the Euclidean equation of motion

$$\ddot{q}_c(t) = \frac{1}{\sqrt{g}} V'(q_c(t)\sqrt{g}). \quad (37.42)$$

Integrating once, we obtain for a finite action solution (for $\beta = \infty$):

$$\frac{1}{2}\dot{q}_c^2(t) - g^{-1}V(q_c(t)\sqrt{g}) = 0. \quad (37.43)$$

Denoting by x_0 the relevant zero of $V(x)$, we can write the corresponding action as

$$\mathcal{S}(q_c) = \int_{-\infty}^{+\infty} \dot{q}_c^2(t) \mathrm{d}t = \frac{a}{g}, \quad \text{with } a = 2 \int_0^{x_0} \sqrt{2V(x)} \mathrm{d}x. \quad (37.44)$$

We note that the classical action is positive and proportional to $1/g$.

The Gaussian integration. To calculate the instanton contribution at leading order, we have to integrate over the paths close to the saddle point $q_c(t)$, in the Gaussian approximation. However, as we have explained in Section 37.4, we must first separate a collective coordinate corresponding to time translation and restrict the Gaussian integration to the other modes. The zero mode yields a factor β and, at leading order, the Jacobian

$$J = \left[\int_{-\infty}^{+\infty} \dot{q}_c^2(t) dt \right]^{1/2} = \left(\frac{a}{g} \right)^{1/2}. \quad (37.45)$$

Since $\dot{q}_c(t)$, which is an eigenfunction of the operator $\delta^2 \mathcal{S} / \delta q(t_1) \delta q(t_2)$ has a node at the turning point x_0 / \sqrt{g} , there exists one eigenfunction with negative eigenvalue. Therefore, the determinant of the operator $\delta^2 \mathcal{S} / \delta q(t_1) \delta q(t_2)$ from which the eigenvalue 0 has been removed is negative.

Collecting all factors, one obtains

$$\begin{aligned} \text{Im} \operatorname{tr} e^{-\beta H} &\sim \frac{1}{2} \frac{\beta}{\sqrt{2\pi}} e^{-\beta/2} \sqrt{\frac{a}{g}} \left[\det M_0 (-\det' M)^{-1} \right]^{1/2} e^{-a/g} \\ &\text{for } g \rightarrow 0, \quad \beta \rightarrow \infty, \end{aligned} \quad (37.46)$$

with M_0 and M given by

$$M(t_1, t_2) = \delta^2 \mathcal{S} / \delta q(t_1) \delta q(t_2) \Big|_{q=q_c}, \quad M_0(t_1, t_2) = \left[-(d_{t_1})^2 + 1 \right] \delta(t_1 - t_2), \quad (37.47)$$

and \det' means determinant in the subspace orthogonal to \dot{q}_c (see also Section 37.4).

One infers the imaginary part of the ‘ground state’ energy E_0 of the metastable state:

$$\text{Im } E_0 = \frac{1}{2} \sqrt{\frac{a}{2\pi g}} \left[\det M_0 (-\det' M)^{-1} \right]^{1/2} e^{-a/g}. \quad (37.48)$$

An expansion around the saddle point then generates an expansion in powers of g .

37.6 Evaluation of the determinant: The shifting method

To determine $\det' M$, we calculate the Gaussian integral explicitly using the *shifting method* [381], a calculation which can always be done in one-dimensional systems and, more generally, in classically integrable systems. The main drawback of the method is that it involves a dangerous change of variables, and the final result is at first sight undefined. On the other hand, it makes a rather straightforward evaluation of the determinant possible. The idea behind the calculation is that, if we know the solutions of the classical equation of motion for arbitrary boundary conditions, we can construct a canonical transformation that maps any Hamiltonian system onto a standard one (here we choose a free Hamiltonian). For details, see Section A3.2.

In Section A37.2, for a comparison, we describe the calculation with the use of the WKB method (solving the Schrödinger equation for $\hbar \rightarrow 0$).

37.6.1 The shifting method

For reasons that will become apparent later, we first calculate the general matrix element

$$\langle x' | e^{-\beta H} | x \rangle = \int_{q(-\beta/2)=x'}^{q(\beta/2)=x} [dq(t)] \exp [-\mathcal{S}(q)] \quad (37.49)$$

(we have used the quantum bra-ket notation). We denote by $q_c(t)$ a classical solution satisfying the boundary conditions $q_c(-\beta/2) = x'$ and $q_c(\beta/2) = x$, and by

$$\mathcal{S}_c(x', x; \beta) = \int_{-\beta/2}^{\beta/2} \left[\frac{1}{2} \dot{q}_c^2(t) + g^{-1} V(q_c(t) \sqrt{g}) \right] dt, \quad (37.50)$$

the corresponding classical action. Setting

$$q(t) = q_c(t) + r(t), \Rightarrow r(\pm\beta/2) = 0, \quad (37.51)$$

we obtain, at leading order, the path integral

$$\begin{aligned} \langle x' | e^{-\beta H} | x \rangle &\sim e^{-\mathcal{S}_c} \int_{r(-\beta/2)=0}^{r(\beta/2)=0} [dr(t)] \exp [-\Sigma(r)], \quad \text{with} \\ \Sigma(r) &= \int_{-\beta/2}^{\beta/2} \frac{1}{2} \left[\dot{r}^2(t) + V''(q_c \sqrt{g}) r^2(t) \right] dt. \end{aligned} \quad (37.52)$$

We then set

$$V''(\sqrt{g} q_c(t)) = \ddot{\kappa}(t)/\kappa(t). \quad (37.53)$$

We know at least one solution $\kappa(t)$. Differentiating the equation of motion (37.42), we obtain

$$\frac{d^2}{dt^2} \dot{q}_c(t) = V''(q_c(t) \sqrt{g}) \dot{q}_c(t). \quad (37.54)$$

If $\dot{q}_c(t)$ does not vanish on the classical trajectory, we can choose $\kappa(t) = \dot{q}_c(t)$. Otherwise, we look for a linear combination of the two independent solutions of equation (37.53), $\dot{q}_c(t)$ and

$$\dot{q}_c(t) \int^t \frac{d\tau}{[\dot{q}_c(\tau)]^2},$$

that does not vanish on the classical trajectory.

The action in expression (37.52) then can be written as

$$\int_{-\beta/2}^{\beta/2} \frac{1}{2} \left(\dot{r}^2(t) + \frac{\dot{\kappa}(t)}{\kappa(t)} r^2(t) \right) dt = \int_{-\beta/2}^{\beta/2} \frac{1}{2} \left(\dot{r}(t) - \frac{\dot{\kappa}(t)}{\kappa(t)} r(t) \right)^2 dt. \quad (37.55)$$

This can be verified by integrating by parts the cross term in the expansion of the square in the right-hand side and using the boundary conditions $r(\pm\beta/2) = 0$. Then, after the linear change of variable, $r(t) \mapsto \sigma(t)$, with

$$\dot{r}(t) - \frac{\dot{\kappa}(t)}{\kappa(t)} r(t) = \dot{\sigma}(t), \quad \sigma(-\beta/2) = 0, \quad (37.56)$$

the action for a time-dependent harmonic oscillator transforms into a free action:

$$\int_{-\beta/2}^{\beta/2} \frac{1}{2} \left(\dot{r}^2(t) + \frac{\dot{\kappa}(t)}{\kappa(t)} r^2(t) \right) dt = \int_{-\beta/2}^{\beta/2} \frac{1}{2} \dot{\sigma}^2(t) dt. \quad (37.57)$$

The transformation (37.56) has the form of a Langevin equation (34.1), $r(t)$ corresponding to the trajectory and $\dot{\sigma}(t)$ to the noise. Therefore, the same difficulty as in naive continuum derivations of the Fokker–Planck (FP) equation is encountered. Indeed, integrating equation (37.56), one obtains

$$r(t) = \kappa(t) \int_{-\beta/2}^t d\tau \frac{\dot{\sigma}(\tau)}{\kappa(\tau)} = \sigma(t) + \kappa(t) \int_{-\beta/2}^t d\tau \sigma(\tau) \frac{\dot{\kappa}(\tau)}{\kappa^2(\tau)}. \quad (37.58)$$

The Jacobian J of this transformation is formally the determinant of the kernel (see Section 35.2.1),

$$J = \frac{\delta r(t_2)}{\delta \sigma(t_1)} = \det \left[\delta(t_1 - t_2) + \theta(t_2 - t_1) \kappa(t_2) \frac{\dot{\kappa}(t_1)}{\kappa^2(t_1)} \right], \quad (37.59)$$

where $\theta(t)$ is the step function ($\theta(t) = 0$ for $t < 0$, $\theta(t) = 1$ for $t > 0$).

Expanding

$$\ln \det(1 + M) = \text{tr} \ln(1 + M) = \text{tr} M - \frac{1}{2} \text{tr} M^2 + \dots, \quad (37.60)$$

one verifies that only the first term does not vanish, but has the ambiguous form

$$\ln J = \theta(0) \int_{-\beta/2}^{\beta/2} dt \frac{\dot{\kappa}(t)}{\kappa(t)} = \theta(0) \ln [\kappa(\beta/2)/\kappa(-\beta/2)]. \quad (37.61)$$

For reasons we have discussed in Section 3.3.1 (commutation of derivative and expectation value, which is required to justify the identity (37.55) within the path integral), the suitable prescription is $\theta(0) = \frac{1}{2}$, and the Jacobian becomes

$$J = \sqrt{\frac{\kappa(\beta/2)}{\kappa(-\beta/2)}}. \quad (37.62)$$

In the path integral, we still have to impose the boundary condition:

$$0 = r(\beta/2) = \kappa(\beta/2) \int_{-\beta/2}^{\beta/2} dt \frac{\dot{\sigma}(t)}{\kappa(t)}. \quad (37.63)$$

This condition can be implemented by introducing a δ -function for which we use the Fourier representation

$$\delta(r(\beta/2)) = \frac{1}{\kappa(\beta/2)} \int \frac{d\lambda}{2\pi} \exp \left(i\lambda \int_{-\beta/2}^{\beta/2} dt \frac{\dot{\sigma}(t)}{\kappa(t)} \right).$$

The complete expression then reads

$$\langle x' | e^{-\beta H} | x \rangle \sim e^{-S_c} \int_{\sigma(-\beta/2)=0} [d\sigma(t)] \frac{d\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} e^{-S(\sigma, \lambda)}, \quad (37.64)$$

with

$$S(\sigma, \lambda) = \int_{-\beta/2}^{\beta/2} dt \left(\frac{1}{2} \dot{\sigma}^2(t) - i\lambda \frac{\dot{\sigma}(t)}{\kappa(t)} \right). \quad (37.65)$$

To eliminate the term linear in $\dot{\sigma}(t)$ in equation (37.65), we shift $\dot{\sigma}(t)$,

$$\sigma(t) \mapsto \varsigma(t), \text{ with } \dot{\sigma}(t) = i \frac{\lambda}{\kappa(t)} + \dot{\varsigma}(t). \quad (37.66)$$

After the shift, the path integral becomes

$$\begin{aligned} \langle x' | e^{-\beta H} | x \rangle &\sim e^{-S_c} \int_{\varsigma(-\beta/2)=0} [\mathrm{d}\varsigma(t)] \frac{\mathrm{d}\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} \\ &\times \exp \left[-\frac{1}{2} \lambda^2 \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} - \frac{1}{2} \int_{-\beta/2}^{\beta/2} \dot{\varsigma}^2(t) \mathrm{d}t \right]. \end{aligned} \quad (37.67)$$

The integration over λ yields

$$\langle x' | e^{-\beta H} | x \rangle \sim e^{-S_c(x',x;\beta)} \left[\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} \right]^{-1/2} \mathcal{N}(\beta). \quad (37.68)$$

The constant $\mathcal{N}(\beta)$ does not depend on x and x' , and is proportional to the matrix element $\langle 0 | e^{-\beta H_0} | 0 \rangle$, in which H_0 is the free Hamiltonian:

$$\langle x' | e^{-\beta H_0} | x \rangle = (2\pi\beta)^{-1/2} e^{-(x'-x)^2/2\beta}. \quad (37.69)$$

To determine $\mathcal{N}(\beta)$, we set $H = H_0$ in equation (37.68) and note that in this case $\kappa(t)$ is a constant. The final result is

$$\langle x' | e^{-\beta H} | x \rangle \sim e^{-S_c(x',x;\beta)} \left[2\pi\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} \right]^{-1/2}. \quad (37.70)$$

We leave as an exercise to show that the result (37.70) is formally independent of the particular linear combination of the two solutions of equation (37.53) that has been chosen. To obtain a more explicit expression, we then substitute for example $\kappa(t) = \dot{q}_c(t)$. We integrate the equation of motion (37.42), taking into account the boundary conditions,

$$\frac{1}{2} \dot{q}_c^2(t) = g^{-1} [V(q_c(t)\sqrt{g}) + E], \quad (37.71)$$

and, therefore,

$$\beta = \int_{x'\sqrt{g}}^{x\sqrt{g}} \frac{\mathrm{d}q}{[2(E + V(q))]^{1/2}}. \quad (37.72)$$

Differentiating equation (37.72) with respect to β , we obtain the equation

$$1 = - \int_{x'\sqrt{g}}^{x\sqrt{g}} \frac{\mathrm{d}q}{[2(E + V(q))]^{3/2}} \frac{\partial E}{\partial \beta}, \quad (37.73)$$

which can be written as

$$\frac{\partial E}{\partial \beta} = - \left[\int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} \right]^{-1}. \quad (37.74)$$

The result (37.70) can then also be expressed as

$$\langle x' | e^{-\beta H} | x \rangle \sim e^{-\mathcal{S}_c(x', x; \beta)} \frac{1}{\sqrt{2\pi \dot{q}_c(\beta/2) \dot{q}_c(-\beta/2)}} \left(-\frac{\partial E}{\partial \beta} \right)^{1/2}. \quad (37.75)$$

We leave as an exercise to verify the identity

$$\kappa(\beta/2) \kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{dt}{\kappa^2(t)} = \left(-\frac{\partial^2 \mathcal{S}_c}{\partial x \partial x'} \right)^{-1}. \quad (37.76)$$

Substituting equation (37.76) into equation (37.70), one then obtains Van Vleck's formula [382] (Section A37.2) in imaginary time:

$$\langle x' | e^{-\beta H} | x \rangle \sim \left(-\frac{1}{2\pi} \frac{\partial^2 \mathcal{S}_c}{\partial x \partial x'} \right)^{1/2} \exp[-\mathcal{S}_c(x', x; \beta)]. \quad (37.77)$$

Several component paths. The calculation of the instanton contribution by the shifting method can be generalized to $\nu > 1$ component vectors \mathbf{q} , provided one can find a non-singular $\nu \times \nu$ matrix \mathbf{K} , solution of the equation

$$\ddot{K}_{ij}(t) = \sum_k \frac{\partial V(\mathbf{q}_c(t))}{\partial q_i \partial q_k} K_{kj}(t). \quad (37.78)$$

The change of variables (37.56) then takes the form

$$\dot{\mathbf{r}}(t) - \dot{\mathbf{K}}(t) \mathbf{K}^{-1}(t) \mathbf{r}(t) = \dot{\boldsymbol{\sigma}}(t). \quad (37.79)$$

The matrix \mathbf{K} can be chosen in such a way that $\dot{\mathbf{K}} \mathbf{K}^{-1}$ is symmetric. It is easy to verify that all arguments can then be repeated and, finally, one obtains an expression similar to equation (37.75):

$$\begin{aligned} \langle \mathbf{x}' | e^{-\beta H} | \mathbf{x} \rangle &\sim \left\{ (2\pi)^d \det \left[\mathbf{K}(\beta/2) \mathbf{K}(-\beta/2) \int_{-\beta/2}^{\beta/2} dt (\mathbf{K}^T)^{-1} \mathbf{K}^{-1} \right] \right\}^{-1/2} \\ &\times \exp[-\mathcal{S}_c(\mathbf{x}', \mathbf{x}; \beta)]. \end{aligned} \quad (37.80)$$

This expression is again equivalent to Van Vleck's formula (see Section A37.2), and can be derived in the same conditions, that is, if the classical equations of motion can be solved for arbitrary initial and final conditions. For more than one degree of freedom, this is no longer the generic situation, and it corresponds only to the special class of integrable Hamiltonians. Simple examples are provided by $O(N)$ -symmetric potentials.

37.6.2 The partition function

In order to calculate $\text{tr } e^{-\beta H}$, we now impose periodic boundary conditions. Then,

$$[\dot{q}_c(\beta/2) \dot{q}_c(-\beta/2)]^{-1/2} = \left\{ \frac{2}{g} [V(x\sqrt{g}) + E] \right\}^{-1/2}. \quad (37.81)$$

Integrating over x , we obtain the trace. Using equation (37.72), we find

$$\int dx \left[\frac{2}{g} (V(x\sqrt{g}) + E) \right]^{-1/2} = \beta. \quad (37.82)$$

Collecting all factors, we obtain the more explicit expression

$$\text{Im} \operatorname{tr} e^{-\beta H} \sim \frac{\beta}{2i} \left(-\frac{\partial E}{\partial \beta} \frac{1}{2\pi g} \right)^{1/2} e^{-A(\beta)/g}, \quad (37.83)$$

where $E(\beta)$ and $A(\beta)$ are defined by

$$\beta = 2 \int_{x_-}^{x_+} \frac{dx}{[2(E(\beta) + V(x))]^{1/2}}, \quad (37.84)$$

$$A(\beta) = 2 \int_{x_-}^{x_+} dx [2(E(\beta) + V(x))]^{1/2} - \beta E(\beta). \quad (37.85)$$

The quantities x_+ and x_- are the zeros of $E(\beta) + V(x)$. Note the useful relation

$$\partial A / \partial \beta = -E(\beta). \quad (37.86)$$

It is clear that, at least for β large enough, $E(\beta)$ is a negative increasing function of β . Therefore, $-\partial E / \partial \beta$ is negative, and the result

$$\text{Im} \operatorname{tr} e^{-\beta H} \sim -\frac{\beta}{2} \left(\frac{\partial E}{\partial \beta} \frac{1}{2\pi g} \right)^{1/2} e^{-A(\beta)/g} \quad \text{for } g \rightarrow 0 \quad (37.87)$$

is real, as expected. This completes the calculation for finite β .

Remark. At β finite, the calculation is valid only above some critical value β_c . Indeed, when β decreases, x_+ and x_- approach a common value x_0 , which corresponds to a maximum of $V(x)$:

$$\begin{cases} V(x) \sim V_0 - \frac{1}{2}\omega^2(x - x_0)^2 + O[(x - x_0)^3], \\ V_0 > 0. \end{cases} \quad (37.88)$$

Let us parametrize E , x_+ and x_- ,

$$E = -V_0 + \frac{1}{2}\omega^2\varepsilon^2, \quad x_\pm = x_0 \pm \varepsilon. \quad (37.89)$$

Then,

$$\beta = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{dx}{[\omega^2\varepsilon^2 - \omega^2(x - x_0)^2]^{1/2}}, \Rightarrow \lim_{\varepsilon \rightarrow 0} \beta = \beta_c = 2\pi/\omega. \quad (37.90)$$

For $\beta \leq \beta_c$, no instanton solution can be found, and, by contrast, it is the perturbative expansion around the classical extremum $x = x_0$ of the potential that becomes relevant.

37.7 Zero temperature limit: The ground state

Rewriting equation (37.84) as

$$\beta = 2 \int_{x_-}^{x_+} \left\{ [2(V(x) + E)]^{-1/2} - (x^2 + 2E)^{-1/2} + (x^2 + 2E)^{-1/2} \right\} dx, \quad (37.91)$$

we can explicitly evaluate the last term, and neglect E in the difference between the first two terms. This leads to

$$E(\beta) \sim -2C e^{-\beta}, \quad \text{with } C = x_+^2 \exp \left[2 \int_0^{x_+} \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} \right) dx \right], \quad (37.92)$$

where now x_+ is the zero of the potential. Equation (37.85) has a large β expansion of the form

$$A(\beta) = a - 2C e^{-\beta} + O(e^{-2\beta}), \quad \text{with } a = 2 \int_0^{x_+} \sqrt{2V(x)}. \quad (37.93)$$

Substituting into equation (37.87), one obtains at leading order,

$$\text{Im } e^{-\beta E_0(g)} \underset{g \rightarrow 0}{\sim} \frac{\beta}{2} e^{-\beta/2} \left(\frac{C}{\pi g} \right)^{1/2} e^{-a/g}, \quad (37.94)$$

and thus,

$$\text{Im } E_0(g) \underset{g \rightarrow 0}{\sim} -\frac{1}{2} \left(\frac{C}{\pi g} \right)^{1/2} e^{-a/g}. \quad (37.95)$$

Here we have calculated only the imaginary part of the would-be ground state energy. To derive the imaginary part of the excited levels, we have to keep the correction of order $e^{-\beta}$ in $A(\beta)$ for β large. We then expand $\exp[-g^{-1}A(\beta)]$ in powers of $e^{-\beta}$. The coefficient of $e^{-N\beta}$ in the expansion yields the imaginary part of the N th level, at leading order.

Two remarks

(i) We have assumed that we have only one instanton solution corresponding to a given zero of the potential. If we find other instanton solutions corresponding to other zeroes of the potential, we have to look for the solution of minimal action, which gives the largest contribution in the small-coupling limit.

(ii) In Section 37.1, we have argued that the imaginary part of the energy levels which we evaluate is the inverse lifetime of a state whose wave function is originally concentrated near the bottom of the metastable minimum of the potential. This interpretation is not problematic for potentials which are either unbounded, or have a continuous spectrum, in which case the complex energy level corresponds to a resonance in the potential. For potentials which have a pure discrete spectrum (and all eigenvalues are real), the situation appears more puzzling. First, the energy of the initial state, which is large compared to the energy of true ground state, corresponds in the semi-classical limit to an almost continuous spectrum outside the well. Moreover, in the semi-classical limit, the lifetime of the metastable state is very long. For times that are not too long, the decay process is exponential, and ignores effects coming from the shape of the potential outside of the barrier. Eventually inverse tunnelling will occur, and the decay law will be modified.

A37 Exact Jacobian. WKB method.

We give here, *without proof, explicit exact expressions*, beyond the leading-order approximation used in the chapter, for the Jacobians generated by the method of collective coordinates [383]. The generalization to quantum field theory (QFT) is simple.

We then discuss semi-classical calculations, using the WKB method, in the framework of the Schrödinger equation, under the assumption that the classical equations of motion can be solved for arbitrary boundary conditions (see Section A3.2).

A37.1 The exact Jacobian

We denote by $\mathbf{q}(t)$, $t \in \mathbb{R}$, the N -component path over which one integrates, and $\mathbf{q}_c(t)$ the instanton solution. In terms of complete set of collective coordinates τ_i , the Jacobian can be written as

$$\mathcal{J} = \det \mathbf{J}(\mathbf{q}) / \det^{1/2} \mathbf{J}(\mathbf{q}_c), \quad (A37.1)$$

where $\mathbf{J}(\mathbf{q})$ is the matrix with elements

$$J_{ij}(\mathbf{q}) = \int dt \frac{\partial \mathbf{q}}{\partial \tau^i} \cdot \frac{\partial \mathbf{q}_c}{\partial \tau^j}. \quad (A37.2)$$

A37.1.1 Example: Time translation

We first assume that $q(t)$ has one component, and the instanton solution breaks the symmetry of the action under time translation. Then, the function $J(q)$ in equation (A37.2) reduces to the expression

$$J(q) = \int dt \dot{q}(t) \dot{q}_c(t).$$

Moreover, one integrates over all paths $q(t)$ with the constraint

$$\int dt (q(t) - q_c(t)) \dot{q}_c(t) = 0.$$

Setting $q(t) = q_c(t) + r(t)$, after an integration by parts, one obtains (assuming the boundary terms cancel),

$$J(q) = J(q_c) - \int dt \ddot{q}_c(t) r(t).$$

A37.1.2 Time translation and $O(N)$ internal rotations

We now consider a path integral where the integrand is both invariant under time translation and internal $O(N)$ group transformations. We assume that the general instanton solution takes the form

$$\mathbf{q}_c(t) = \mathbf{u} q_c(t + t_0), \quad (A37.3)$$

where \mathbf{u} is a time-independent unit vector: $\mathbf{u}^2 = 1$, a form that breaks both time-translation and $O(N)$ invariance. Parametrizing the sphere by collective coordinates τ_i , we then set

$$\mathbf{q}(t) = q_L(t + t_0) \mathbf{u}(\boldsymbol{\tau}) + \mathbf{q}_T(t + t_0), \quad (A37.4)$$

where

$$\mathbf{u} \cdot \mathbf{q}_T(t) = 0. \quad (A37.5)$$

The Jacobian can then be written as (g_{ij} is the metric on the sphere S_{N-1})

$$\begin{aligned} \mathcal{J} = & (\det g_{ij})^{1/2} \left[\int dt \mathbf{q}_c^2(t) \right]^{(1-N)/2} \left[\int dt \dot{\mathbf{q}}_c^2(t) \right]^{-1/2} \left[\int dt \mathbf{q}(t) \cdot \mathbf{q}_c(t) \right]^{N-2} \\ & \times \int dt dt' [\dot{\mathbf{q}}(t) \cdot \dot{\mathbf{q}}_c(t) \mathbf{q}(t') \cdot \mathbf{q}_c(t') - \dot{q}_c(t) \dot{q}_c(t') \mathbf{q}_{\text{T}}(t) \cdot \mathbf{q}_{\text{T}}(t')] . \end{aligned} \quad (A37.6)$$

A37.2 The WKB method

We reintroduce the quantity \hbar , which we often elsewhere set equal to 1, to make the expansion parameter explicit. We explicitly write the Schrödinger equation (in real time) for the evolution operator as

$$H \left(\frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}}, \mathbf{x}; t \right) U(\mathbf{x}, \mathbf{x}'; t) = i\hbar \frac{\partial U}{\partial t}(\mathbf{x}, \mathbf{x}'; t), \quad (A37.7)$$

with the definition $U(\mathbf{x}, \mathbf{x}'; t) = \langle \mathbf{x} | \mathbf{U}(t) | \mathbf{x}' \rangle$, and the boundary condition

$$\mathbf{U}(t = T') = \mathbf{1} . \quad (A37.8)$$

We assume that, in equation (A37.7), the Hamiltonian is Hermitian, and results from the quantization of a classical Hamiltonian. Thus, we now face the problem of how to associate a quantum operator with a real classical Hamiltonian $H(\mathbf{p}, \mathbf{q}, ; t)$ (see Chapter 3). As an ansatz, we set

$$U(\mathbf{x}', \mathbf{x}; t) = G(\mathbf{x}, \mathbf{x}'; t) e^{iA(\mathbf{x}, \mathbf{x}'; t)/\hbar} [1 + O(\hbar)] . \quad (A37.9)$$

Introducing the ansatz into equation (A37.7), and keeping the two first terms in \hbar , we obtain two equations. The first equation involves only the classical Hamiltonian. It is the Hamilton–Jacobi equation for the classical action on the classical trajectory,

$$H \left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t \right) = -\frac{\partial A}{\partial t} . \quad (A37.10)$$

With the boundary conditions implied by the condition (A37.8), it determines A completely. The derivation of the second equation involves some more work. First, we note that

$$H \left(\frac{\hbar}{i} \frac{\partial}{\partial \mathbf{x}}, \mathbf{x}; t \right) G = GH - i\hbar \sum_i \frac{\partial H}{\partial p_i} \frac{\partial G}{\partial x_i} + O(\hbar^2) . \quad (A37.11)$$

Here again, only the classical Hamiltonian is needed. The term containing $\partial H / \partial p_i$ is already multiplied by a factor \hbar , thus we can replace the operator p_i by $\partial A / \partial x_i$. For the first term, we now use the identity

$$\begin{aligned} e^{-iA/\hbar} H e^{iA/\hbar} = & H \left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t \right) - \frac{i\hbar}{2} \sum_j \frac{\partial^2 H}{\partial p_j \partial q_j} \left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t \right) \\ & - \frac{i\hbar}{2} \sum_{j,k} \frac{\partial^2 H}{\partial p_j \partial p_k} \left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t \right) \frac{\partial^2 A}{\partial x_j \partial x_k} + O(\hbar^2) . \end{aligned} \quad (A37.12)$$

The second term in the right-hand side comes from commuting all the derivatives completely on the right. It relies on the assumption that the quantum Hamiltonian is Hermitian. Indeed, let us first assume that we have symmetrized all monomials:

$$p^n q^m \rightarrow \frac{1}{2} (p^n q^m + q^m p^n). \quad (A37.13)$$

Then, a contribution to this term arises each time an operator p of $p^n q^m$ acts on q^m , and the factor $\frac{1}{2}$ comes from the symmetrization. If we choose another Hermitian quantization procedure, we can start commuting all operators p and q until the Hamiltonian is again a sum of terms (A37.13). Each commutation introduces a factor $i\hbar$. Since the difference between the two expressions is Hermitian, it can only involve $(i\hbar)^2$, which can be neglected at this order.

The third term in expression (A37.12) arises from two derivatives acting on the action. The factor $1/2$ is a counting factor. We then obtain the equation for G ,

$$\sum_i \frac{\partial H}{\partial p_i} \frac{\partial G}{\partial x_i} + \frac{1}{2} \left(\sum_j \frac{\partial^2 H}{\partial p_j \partial q_j} + \sum_{j,k} \frac{\partial^2 H}{\partial p_j \partial p_k} \frac{\partial^2 A}{\partial x_j \partial x_k} \right) G = -\frac{\partial G}{\partial t}. \quad (A37.14)$$

We introduce the matrix notation

$$M_{ij} = \frac{\partial^2 A}{\partial x'_i \partial x_j}, \quad H_{ij} = \frac{\partial^2 H}{\partial p_i \partial p_j}, \quad \tilde{H}_{ij} = \frac{\partial^2 H}{\partial p_i \partial q_j}. \quad (A37.15)$$

We now differentiate equation (A37.10) with respect to x'_i and x_j , successively. We find,

$$\sum_k M_{ik} \frac{\partial H}{\partial p_k} = -\frac{\partial^2 A}{\partial t \partial x'_i}, \quad (A37.16)$$

and

$$\sum_k \frac{\partial \mathbf{M}}{\partial x_k} \frac{\partial H}{\partial p_k} + \mathbf{M} \mathbf{H} \mathbf{M} + \mathbf{M} \tilde{\mathbf{H}} = -\frac{\partial \mathbf{M}}{\partial t}. \quad (A37.17)$$

All multiplications are meant in a matrix sense. We now multiply equation (A37.17) by \mathbf{M}^{-1} on the left and take the trace:

$$\sum_k \frac{\partial H}{\partial p_k} \text{tr } \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial x_k} + \text{tr} (\mathbf{H} \mathbf{M} + \tilde{\mathbf{H}}) = -\text{tr } \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial t}. \quad (A37.18)$$

We note that

$$\partial \ln \det \mathbf{M} = \partial \text{tr} \ln \mathbf{M} = \text{tr} \partial \mathbf{M} \mathbf{M}^{-1}. \quad (A37.19)$$

The equation (A37.18) can then be rewritten as

$$\frac{\partial H}{\partial p_i} \frac{\partial}{\partial x_i} \ln \det \mathbf{M} + \text{tr} \tilde{\mathbf{H}} + \text{tr} \mathbf{M} \mathbf{H} = -\frac{\partial}{\partial t} \ln \det \mathbf{M}, \quad (A37.20)$$

while equation (A37.14) can be written as

$$\frac{\partial H}{\partial p_i} \frac{\partial}{\partial x_i} \ln G + \frac{1}{2} \text{tr} \tilde{\mathbf{H}} + \frac{1}{2} \text{tr} \mathbf{M} \mathbf{H} = -\frac{\partial}{\partial t} \ln G. \quad (A37.21)$$

Comparing the two equations, we conclude that a solution to equation (A37.21) is

$$\ln G = \frac{1}{2} \ln \det \mathbf{M} + \text{const.} \quad (A37.22)$$

Taking into account the boundary conditions, we obtain Van Vleck's formula [382],

$$\langle \mathbf{x} | U(T, T') | \mathbf{x}' \rangle \sim \frac{1}{(2\pi i\hbar)^{n/2}} \left(-\det \frac{\partial^2 A}{\partial x_i \partial x_j} \right)^{1/2} e^{iA(\mathbf{x}, \mathbf{x}'; t)/\hbar}. \quad (A37.23)$$

It is straightforward to derive from this equation the corresponding expression for imaginary time.

38 Metastable vacua in quantum field theory (QFT)

In this chapter, we generalize to quantum field theory (QFT) [373, 374] the methods to evaluate barrier penetration effects in the semi-classical limit, as described in Chapter 37. In quantum mechanics (QM), we have shown that barrier penetration is associated with classical motion in imaginary time. Therefore, we consider QFT here in its Euclidean formulation.

In the representation of QFT in terms of field integrals, in the semi-classical limit, barrier penetration is related to finite action solutions (instantons) of the classical field equations [384]. We first characterize such solutions. We then explain how to evaluate the instanton contributions at leading order, the main new problem arising from ultraviolet (UV) divergences [375].

We have argued that the lifetime of metastable states is related to the imaginary part of the ‘ground state’ energy. However, for later purpose, it is useful to calculate the imaginary part not only of the vacuum amplitude, but also of correlation functions. In the case of the vacuum amplitude, we find that the instanton contribution is proportional to the space–time volume. Therefore, dividing by the volume, we obtain the probability per unit time and unit volume of a metastable pseudo-vacuum to decay.

We first discuss a scalar QFT with a ϕ^4 interaction, generalization of the quartic anharmonic oscillator considered in Sections 37.1–37.4, in two and three dimensions, dimensions in which the theory is super-renormalizable. We then consider more general scalar QFTs, of a form analogous to the quantum models discussed in Section 37.5 [380].

We calculate instanton contributions, at leading order, explicitly in the ϕ^4 theory in dimension 4, the dimension in which the theory is renormalizable. Several new problems arise. With help of Sobolev’s inequalities, we prove in Section A38.2 that the massive field equation has no instanton solution in dimension 4, and that the relevant instanton is a solution of the massless field equation. Therefore, we first study the massless ϕ^4 QFT, and comment at the end about the massive theory. The price to pay for such a simplification is the appearance of some subtle infrared (IR) problems. In the leading order calculation, in addition to the mass renormalization already met in the super-renormalizable case, the one-loop coupling constant renormalization has to be taken into account. This feature, together with the scale invariance of the classical theory leads to the appearance of an effective coupling constant at the scale of the instanton and, therefore, the calculation of the contribution of the instanton depends on global renormalization group (RG) properties of the theory.

Finally, Section 38.8 is devoted to a brief discussion of a speculative cosmological application of these results.

In the appendix, we discuss virial theorems, Sobolev inequalities relevant to the properties of the classical solutions of the ϕ^4 field theory, RG properties and conformal invariance relevant to the instanton calculations in the $\phi_{d=4}^4$ field theory.

38.1 The ϕ^4 QFT for negative coupling

We consider the d -dimensional QFT for a scalar field ϕ corresponding, in the classical approximation, to the Euclidean action

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{4!} g m^{4-d} \phi^4(x) \right], \quad (38.1)$$

m being the mass, and g the dimensionless coupling constant (the power of m that appears in front of the interaction term ϕ^4 takes care of the dimension).

The complete n -point correlation function has the field-integral representation

$$Z^{(n)}(x_1, \dots, x_n) = \int [d\phi(x)] \phi(x_1) \phi(x_2) \cdots \phi(x_n) \exp[-\mathcal{S}(\phi)]. \quad (38.2)$$

We normalize the field integral with respect to the vacuum amplitude (partition function) at $g = 0$ to avoid introducing a non-trivial g dependence through the normalization. Following the method described in Section 37.3, we assume that we start from positive values of g , and proceed by analytic continuation to define the field integral for g negative. The imaginary part of correlation functions is given by the difference between the continuations above and below the negative g -axis. For g small, only non-constant saddle points contribute to the imaginary part. Therefore, we look for instanton configurations, corresponding to non-constant finite action solutions of the Euclidean field equations, and evaluate their contributions.

38.1.1 Instantons: Classical solutions and classical action

The instanton solutions. The field equation corresponding to the action (38.1) is

$$(-\nabla^2 + m^2) \phi_c(x) + \frac{1}{6} g m^{4-d} \phi_c^3(x) = 0. \quad (38.3)$$

We set (g is negative)

$$\phi_c(x) = (-6/g)^{1/2} m^{d/2-1} \varphi(mx). \quad (38.4)$$

In terms of φ , the classical action (38.1) reads

$$\mathcal{S}(\varphi) = -\frac{6}{g} \int d^d x \left[\frac{1}{2} (\nabla \varphi(x))^2 + \frac{1}{2} \varphi^2(x) - \frac{1}{4} \varphi^4(x) \right]. \quad (38.5)$$

The function φ satisfies the parameter-free-field equation

$$(-\nabla^2 + 1) \varphi(x) - \varphi^3(x) = 0. \quad (38.6)$$

It can be shown (for details, see Section A38.2) that the solution with the smallest action is spherically symmetric [385]. Therefore, we choose an arbitrary origin x_0 , and set

$$r = |x - x_0|, \quad \varphi(x) = f(r). \quad (38.7)$$

The function f satisfies the non-linear differential equation,

$$\left[-\left(\frac{d}{dr} \right)^2 - \frac{d-1}{r} \frac{d}{dr} + 1 \right] f(r) - f^3(r) = 0. \quad (38.8)$$

Interpreting r as a (real) time, we note that the equation describes the motion of a particle in a potential $U(f)$, with

$$U(f) = -\frac{1}{2} f^2 + \frac{1}{4} f^4, \quad (38.9)$$

submitted, in addition, to a viscous damping force, due to the first derivative.

The finite-action solutions must satisfy the boundary condition

$$f(r) \rightarrow 0, \quad \text{for } r \rightarrow \infty. \quad (38.10)$$

Equation (38.8) shows that, if $f(r)$ goes to 0 for $r \rightarrow +\infty$, it goes exponentially. The equation has solutions even in r , which are thus determined by the value of f at the origin. For a generic value of $f(0)$, the corresponding solution, for $r \rightarrow \infty$, tends towards a minimum of the potential $f = \pm 1$. The condition (38.10) is only satisfied for a discrete set of initial values of $f(0)$. Moreover, it can be shown that the minimal action solution corresponds to the function for which $|f(0)|$ is minimal in the set, and which only vanishes at infinity. The values [387],

$$\text{for } d = 2, \quad f(0) = 2.20620086465074607(1), \quad A = 35.10268957367896(1), \quad (38.11)$$

$$\text{for } d = 3, \quad f(0) = 4.3373876799769943(1), \quad A = 113.38350781527714(1) \quad (38.12)$$

correspond to the suitable numerical solution of the field equation (38.8). Moreover, due to translation symmetry, equation (38.3) has a family of degenerate saddle points $\phi_c(x)$ depending on d parameters $x_{0\mu}$ (equation (38.7)).

The instanton action. Since g is dimensionless, the corresponding classical action has the general form

$$\mathcal{S}(\phi_c) \equiv \mathcal{S}(\varphi) = S/g, \quad \text{with } S = -A, \quad (38.13)$$

and the scaling arguments of Section A38.1 [386] lead to the relations

$$A = \frac{6}{d} \int [\nabla \varphi(x)]^2 d^d x = \frac{3}{2} \int \varphi^4(x) d^d x = \frac{6}{4-d} \int \varphi^2(x) d^d x, \quad (38.14)$$

which show that A is positive. We note that these relations can only be satisfied for $d < 4$, and that the dimension 4 is singular (see Section A38.2).

38.1.2 The Gaussian integration for $d < 4$

To perform the Gaussian integration in the neighbourhood of the saddle point, we have to examine the spectrum of the differential operator, second functional derivative of \mathcal{S} , given by (in quantum bra–ket matrix-element notation)

$$\langle x | \mathbf{M} | x' \rangle \equiv \left. \frac{\delta^2 \mathcal{S}}{\delta \phi(x) \delta \phi(x')} \right|_{\phi=\phi_c} = [(-\nabla_x^2 + m^2) + \frac{1}{2} g m^{4-d} \phi_c^2(x)] \delta^{(d)}(x - x'), \quad (38.15)$$

$$= [(-\nabla_x^2 + m^2) - 3m^2 \varphi^2(mx)] \delta^{(d)}(x - x'). \quad (38.16)$$

Differentiating the equation of motion (38.3) with respect to x_μ , we infer that, as expected, the d partial derivatives $\partial_\mu \phi_c(x)$ ($\partial_\mu \equiv \partial/\partial_\mu$) are eigenvectors of \mathbf{M} with vanishing eigenvalue:

$$(-\nabla^2 + m^2) \partial_\mu \phi_c(x) + \frac{1}{2} g m^{4-d} \phi_c^2(x) \partial_\mu \phi_c(x) = 0 \iff \mathbf{M} \partial_\mu \phi_c = 0. \quad (38.17)$$

As in QM (Section 37.4.1), to sum over all saddle points, we have to factorize the field integration into an integration over the d collective coordinates $x_{0\mu}$ (which has to be done exactly), and an integration over the remaining field modes [388]. For this purpose, we can, for example, apply the identity (37.25) to each variable $x_{0,\mu}$.

The Jacobian. As the result (37.30) shows, this factorization leads to the determinant of \mathbf{M} in the subspace orthogonal to the zero eigenvalue sector, and to a Jacobian J , which, at leading order, is

$$J = \prod_{\mu=1}^d \|\partial_\mu \phi_c(x)\| = \left[-\frac{6}{dg} \int d^d x (\nabla \varphi(x))^2 \right]^{d/2} = \left(\frac{-A}{g} \right)^{d/2}, \quad (38.18)$$

where the invariance under rotation (38.7), and the first relation (38.14) have been used. Moreover, a factor $(2\pi)^{-1/2}$ from the Gaussian integration, is generated for each variable. Note one important feature of this expression: each translation symmetry broken by the instanton solution (each component of x_0) has generated a factor $(-g)^{-1/2}$.

Wave function arguments, of the kind used for the Schrödinger equation, show that $\partial_\mu \phi_c$ is not one ground state of \mathbf{M} . One state has a negative eigenvalue and, therefore, the final result is real as expected. In Section A38.2, we give a proof of this property using Sobolev inequalities.

Correlation functions. In expression (38.2), one can replace, at leading order, the field $\phi(x)$ by $\phi_c(x)$ in the product $\prod_{i=1}^n \phi(x_i)$. Collecting all factors, one finds,

$$\text{Im } Z^{(n)}(x_1, \dots, x_n) = \frac{1}{2i} \left(\frac{A}{2\pi} \right)^{d/2} \Omega \frac{e^{A/g}}{(-g)^{(d+n)/2}} F_n(x_1, \dots, x_n), \quad (38.19)$$

with

$$F_n(x_1, \dots, x_n) = m^{d+n(d-2)/2} 6^{n/2} \int d^d x_0 \prod_{i=1}^n f(m(x_i - x_0)), \quad (38.20)$$

and

$$\langle x | \mathbf{M}_0 | x' \rangle = (-\nabla_x^2 + m^2) \delta^{(d)}(x - x'), \quad (38.21a)$$

$$\Omega^{-2} = \det' \mathbf{M} \mathbf{M}_0^{-1} \Big|_{m=1} = \lim_{\epsilon \rightarrow 0} \epsilon^{-d} \det [(\mathbf{M} + \epsilon) \mathbf{M}_0^{-1}] \Big|_{m=1}. \quad (38.21b)$$

While, for the vacuum amplitude, the integration over x_0 generates a factor proportional to the volume, for non-trivial correlation functions the integration restores translation invariance.

Discussion. A few comments concerning expression (38.19) are in order here. We have obtained a result for the complete correlation functions, improperly normalized, for convenience, with respect to the free QFT. However, because $\phi_c(x)$ is proportional to $1/\sqrt{-g}$, the imaginary part of the n -point function increases with n for g small. This shows that, at leading order, the correlation functions, normalized with respect to the partition function corresponding to the complete action (38.1), have the same behaviour as those renormalized with respect to the free QFT.

Moreover, for the same reason, when we consider a complete n -point function, the imaginary part coming from disconnected parts is subleading by at least one power of g . If we denote by $W^{(n)}(x_1, \dots, x_n)$ the connected n -point function, we thus find, at leading order,

$$\text{Im } W^{(n)} \sim \text{Im } Z^{(n)},$$

a result that is consistent with the observation that the explicit expression (38.19) is indeed connected.

Vertex functions are derived from connected correlation functions by first subtracting the reducible contributions, which involve functions with a smaller number of arguments, and which are, therefore, negligible at leading order, and then by amputating the remaining part. Again, for the same reason, only the real part of the propagator matters; therefore, to amputate expression (38.19), we must simply multiply it by the product of the inverse free propagators corresponding to each external line. Introducing \tilde{f} , the Fourier transform of f , and denoting the n -point vertex function in the Fourier representation by $\tilde{\Gamma}^{(n)}$, one obtains

$$\frac{\text{Im } \tilde{\Gamma}^{(n)}(p_1, \dots, p_n)}{m^{d-n(d/2+1)}} \sim -\frac{1}{2i} \left(\frac{A}{2\pi}\right)^{d/2} \Omega \frac{e^{A/g}}{(-g)^{(d+n)/2}} \prod_{i=1}^n \sqrt{6} \tilde{f}\left(\frac{p_i}{m}\right) (p_i^2 + m^2). \quad (38.22)$$

The structure, at leading order, of the imaginary part of the n -point vertex function is particularly simple. In particular, it only depends on the square of the momenta p_i , and not of their scalar products.

Up to this point, the discussions of the ϕ^4 QFT and of the anharmonic oscillator have been remarkably similar. Now comes one significant difference: the determinant of the operator \mathbf{M} is actually UV divergent, and one has to deal with this new problem.

38.1.3 UV divergences and renormalization for $d < 4$

A regularization is required to define the ϕ^4 theory in dimensions $d < 4$, and then, a mass counter-term must be added to the classical action. After the introduction of higher derivatives and a momentum cut-off Λ (Section 8.4.2), the regularized action \mathcal{S}_Λ takes the form:

$$\mathcal{S}_\Lambda(\phi) = \int d^d x \left[\frac{1}{2} \phi(x) (-\nabla^2 + \nabla^4/\Lambda^2 + r) \phi(x) + \frac{1}{4!} g \phi^4(x) \right], \quad (38.23)$$

where $r = m^2 + \delta m^2(\Lambda)$, $\delta m^2(\Lambda)$ being the mass counter-term (Section 9.2.1).

At large cut-off Λ , the additional term

$$\frac{1}{\Lambda^2} \int \phi(x) \nabla^4 \phi(x) d^d x$$

modifies the equation of motion but, when $\Lambda \rightarrow \infty$, the modification vanishes like $1/\Lambda^2$. On the other hand, the counter-term increases with the cut-off, but is proportional to at least one power of g . Hence, because we take the small g limit before taking the large cut-off limit, the counter-term does not contribute to the classical equation of motion.

If we then calculate the contribution of the counter-term to the classical action, we find that the one-loop counter-term, which is proportional to g , gives a contribution of order 1 in g , because $\phi_c(x)$ is proportional to $1/\sqrt{-g}$. Therefore, it generates an additional multiplicative factor.

The operator \mathbf{M} (equation (38.15)) in the regularized theory is given by

$$\langle x | \mathbf{M} | x' \rangle = [(-\nabla^2 + \nabla^4/\Lambda^2 + m^2) + \frac{1}{2} g m^{4-d} \phi_c^2(x)] \delta^{(d)}(x - x'). \quad (38.24)$$

We expand its determinant in powers of $\phi_c^2(x)$, using the identity $\text{tr} \ln = \ln \det$,

$$\begin{aligned} \ln \det \mathbf{M} &= \ln \det (-\nabla^2 + \nabla^4/\Lambda^2 + m^2) \\ &\quad - \sum_{k=1}^{\infty} \frac{1}{k} \text{tr} \left[\frac{1}{2} g m^{4-d} \phi_c^2(x) (\nabla^2 - \nabla^4/\Lambda^2 - m^2)^{-1} \right]^k. \end{aligned} \quad (38.25)$$

The first term is cancelled by the free determinant $\det \mathbf{M}_0$. All terms for $k \geq 2$, are UV finite in two and three dimensions. Finally, the $k = 1$ term is

$$\tfrac{1}{2} \text{tr } gm^{4-d} \phi_c^2(x) (-\nabla^2 + \nabla^4/\Lambda^2 + m)^{-1} = \tfrac{1}{2} gm^{4-d} D \int d^d x \phi_c^2(x), \quad (38.26)$$

in which D is the regularized free propagator at coinciding arguments,

$$D = \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 + p^4/\Lambda^2 + m^2}. \quad (38.27)$$

The determinant of the operator \mathbf{M} thus contains a factor that diverges at large cut-off as

$$\exp \left[-\tfrac{1}{4} gm^{4-d} D \int d^d x \phi_c^2(x) \right]. \quad (38.28)$$

This factor exactly cancels the infinite factor coming from the mass counter-term, in such a way that the final expression for the imaginary part is finite. The fact that really we have to calculate $\det' \mathbf{M}$ does not change the argument, because UV divergences are insensitive to the omission of a finite number of eigenvalues of \mathbf{M} , as the second form (38.21b) explicitly shows [375].

The decay of the false vacuum. The special case $n = 0$ corresponds to the imaginary part of the vacuum amplitude. We have expanded perturbation theory around the minimum $\phi = 0$ of the potential. The perturbative ground state corresponds to a wave functional concentrated around small fields. However, because we have expanded around a relative minimum of the potential, this state is actually metastable. We have calculated its decay rate due to barrier penetration. We note that the integral over $x_{0\mu}$ in equation (38.20) yields a space–time volume factor. To obtain a finite decay amplitude, we have to divide the result by this volume factor. We thus obtain the probability per unit time and *unit volume* for the metastable state ('false vacuum') of the theory to decay. Some implications of such a result are discussed in a slightly more general context in Section 38.8.

38.2 General potentials: Instanton contributions

We now extend the analysis of Section 37.5 to analogous scalar field theories using the techniques developed in Section 38.1 [380]. We consider a Euclidean action of the form

$$\mathcal{S}(\phi) = \int d^d x \left[\tfrac{1}{2} (\nabla \phi(x))^2 + g^{-1} V(\phi(x)\sqrt{g}) \right], \quad (38.29)$$

in which the polynomial potential $V(\phi)$ has one stable and one metastable minimum, and is of the type discussed in Section 37.5. Assuming that, at some initial time, the quantum state corresponds to a field concentrated around the metastable minimum of the potential, the 'false' vacuum, we want to evaluate, in the semi-classical limit, the probability for the false vacuum to decay into the true vacuum of the theory. The calculation, at leading order, again involves the determination of an instanton solution, a factorization of the integral over collective coordinates, and a remaining Gaussian integration around the instanton.

38.2.1 Calculation of the instanton contribution

We define the field in such a way that the metastable minimum corresponds to $\phi = 0$. The discussion of the existence of an instanton solution is similar to the one given in Section 38.1. A theorem establishes, under mild assumptions, that spherically symmetric solutions give the minimal action [385]. Therefore, we look for such a solution and set

$$r = |x - x_0|, \quad f(r) = \sqrt{g}\phi_c(x).$$

The classical equation of motion reduces to

$$\frac{d^2 f}{dr^2} + \frac{d-1}{r} \frac{df}{dr} = V'(f(r)). \quad (38.30)$$

This is the equation governing the motion of a particle in a potential $-V(f)$ and submitted to a viscous damping force. We denote by f_+ the absolute minimum of the potential. The solution depends on its value at the origin $f(0)$. If we choose $f(0)$ too close to f_+ , $f(r) - f(0)$ remains small until r becomes very large. When r is large, the damping force is small, the energy is almost conserved, and the particle overshoots. If $f(0)$ is too close to 0, the particle loses too much energy and, therefore, undershoots the asymptotic value $f(r)$ then corresponding to the maximum of $V(f)$. Thus, somewhere in between, we expect to find values $f(0)$, which correspond to solutions that vanish at infinity and, therefore, have finite actions.

The virial theorem, derived in Section A38.1, implies that the corresponding action is positive:

$$\mathcal{S}(\phi_c) = S/g, \quad (38.31)$$

with

$$S = \frac{1}{d} \int (\nabla f(x))^2 d^d x > 0. \quad (38.32)$$

Moreover, we also derive in Section A38.1 that the operator \mathbf{M} with kernel

$$\langle x | \mathbf{M} | x' \rangle = \delta^2 \mathcal{S} / \delta \phi(x) \delta \phi(x') \Big|_{\phi=\phi_c},$$

has one and only one negative eigenvalue.

Again, we factorize the field integration measure in an integration over x_0 , and an integration over the other field modes. This generates a Jacobian J which, as we have shown in Section 38.1, at leading order is given by

$$J = \prod_{\mu=1}^d \|\partial_\mu \phi_c\| = \left[\frac{1}{d} \int d^d x (\nabla \phi_c(x))^2 \right]^{d/2} = \left(\frac{S}{g} \right)^{d/2}, \quad (38.33)$$

where equations (38.31, 38.32) are used.

The remaining details of the calculation can be borrowed from the ϕ^4 example, and lead to an explicit expression for the imaginary part of the n -point correlation function.

Renormalization: A few remarks. If the QFT is super-renormalizable or renormalizable, the renormalized theory can be generated in the following way: after regularizing the theory, we proceed by induction, adding the counter-terms that render the theory finite order by order in a loop expansion, that is, here an expansion in powers of g . The renormalized action $\mathcal{S}_r(\phi)$ takes the form

$$\mathcal{S}_r(\phi) = \mathcal{S}_0(\phi\sqrt{g})/g + \mathcal{S}_1(\phi\sqrt{g}) + \dots + g^{L-1} \mathcal{S}_L(\phi\sqrt{g}) + \dots$$

For the instanton calculation, at leading order, only the one-loop counter-terms are needed. To evaluate them, we expand the generating functional of vertex (1PI) functions (or effective potential) $\Gamma(\phi)$ up to one-loop order using the regularized action and calculate its divergent part.

In Chapter 7, we have derived the one-loop expression (equation (7.93))

$$\Gamma_{\text{1 loop}}(\phi) = \mathcal{S}(\phi) + \frac{1}{2} \text{tr} \ln \frac{\delta^2 \mathcal{S}}{\delta \phi(x) \delta \phi(x')}. \quad (38.34)$$

To render vertex functions finite, we have to subtract to the regularized action the divergent part of the one-loop contribution of $\Gamma(\phi)$:

$$\mathcal{S}_1(\phi \sqrt{g}) = -\frac{1}{2} \left(\text{tr} \ln \frac{\delta^2 \mathcal{S}}{\delta \phi \delta \phi} \right)_{\text{div.}}.$$

When evaluated for $\phi = \phi_c$, this contribution exactly cancels the divergence in the determinant coming from the Gaussian integration around the saddle point. This argument can be generalized to arbitrary orders.

38.3 The ϕ^4 QFT in dimension 4

In dimension 4, the ϕ^4 QFT is just renormalizable, and, as we show in Section A38.2, only the massless field equations have instanton solutions. This leads to a set of new problems which we now examine. We first consider the massless theory, which is simpler, although it has some subtle IR problems. In particular, the barrier penetration is rather peculiar, since it is not generated by the potential but only by the kinetic term of the action.

We explain the leading order calculation of instanton contribution for the one-component ϕ^4 theory, but the extension to the $O(N)$ -symmetric model is simple, and the explicit expressions can be found in the literature [375].

The classical Euclidean action of the massless theory ϕ^4 theory can be written as

$$\mathcal{S}(\phi) = \int d^4x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{4} g \phi^4(x) \right], \quad (38.35)$$

and the corresponding field equation reads

$$-\nabla^2 \phi(x) + g \phi^3(x) = 0. \quad (38.36)$$

Note the different normalization of the coupling constant. To return to the convention used elsewhere, one has to substitute $g \mapsto g/6$.

We know that the solution of minimal action is spherically symmetric, and thus we set

$$\phi(x) = \frac{1}{\sqrt{-g}} f(r), \quad \text{with } r = |x - x_0|. \quad (38.37)$$

We then obtain the differential equation

$$-\left[\left(\frac{d}{dr} \right)^2 + \frac{3}{r} \frac{d}{dr} \right] f(r) = f^3(r). \quad (38.38)$$

The classical action is scale invariant (actually, the QFT is conformal invariant, see Section A38.4). If $\phi(x)$ is an instanton solution to the field equation, then the rescaled functions $\psi(x) = \phi(x/\lambda)/\lambda$ (λ is a scale parameter) are also solutions.

This property suggests the following parametrization,

$$f(r) = e^{-t} h(t), \quad \text{with } r = e^t, \quad (38.39)$$

which transforms equation (38.38) into

$$\ddot{h}(t) = h(t) - h^3(t). \quad (38.40)$$

We recognize the equation of motion of the anharmonic oscillator in Chapter 37, whose solution is (equation (37.16)),

$$h_c(t) = \pm \frac{\sqrt{2}}{\cosh(t - t_0)}. \quad (38.41)$$

The solution $\phi_c(x)$ of equation (38.36) is then

$$f(r) = \pm \frac{2\sqrt{2}\lambda}{1 + \lambda^2 r^2}, \quad (38.42a)$$

$$\Rightarrow \phi_c(x) = \pm \frac{1}{\sqrt{-g}} \frac{2\sqrt{2}\lambda}{1 + \lambda^2 (x - x_0)^2}, \quad (38.42b)$$

where we have defined $\lambda = e^{-t_0}$. The corresponding classical action $\mathcal{S}(\phi_c)$ is

$$\mathcal{S}(\phi_c) = S/g, \quad \text{with } S = -8\pi^2/3. \quad (38.43)$$

With the usual normalization of g , one finds $S = -16\pi^2$.

Because the classical theory is scale invariant, the instanton solution now depends on a scale parameter λ , in addition to the four translation parameters $x_{0\mu}$. Therefore, we have to introduce five collective coordinates to calculate the instanton contribution.

38.4 Instanton contributions at leading order

General strategy. The second functional derivative of the action at the saddle point is

$$\langle x | \mathbf{M} | x' \rangle = \left. \frac{\delta^2 \mathcal{S}}{\delta \phi(x) \delta \phi(x')} \right|_{\phi=\phi_c} = \left[-\nabla^2 - \frac{24\lambda^2}{(1 + \lambda^2 x^2)^2} \right] \delta^{(4)}(x - x'). \quad (38.44)$$

To determine the eigenvalues of the operator \mathbf{M} , one has to solve a four-dimensional Schrödinger equation with a spherically symmetric potential. We note, at this stage, two serious problems. The operator \mathbf{M} has formally, as expected, five eigenvectors, $\nabla \phi_c(x)$ and $(d/d\lambda)\phi_c(x)$ with eigenvalue 0, but the last of these eigenvectors is not normalizable with the natural measure of the problem

$$\int \left(\frac{d}{d\lambda} \phi_c(x) \right)^2 d^4x = \infty. \quad (38.45)$$

This is an IR problem which arises, because the theory is massless.

Moreover, the mass counter-term, which has to be added to the action, and has the form

$$\frac{1}{2} \delta m_0^2 \int d^4x \phi_c^2(x) = \infty, \quad (38.46)$$

is also IR divergent, and this IR divergence is expected to cancel an IR divergence coming from $\det \mathbf{M}$.

Thus, in general, we need some kind of IR regularization. In the particular case of dimensional regularization, this problem is postponed to the two-loop order.

These problems will be solved in several steps. First, we realize that we do not need the eigenvalues of \mathbf{M} , but only the determinant $\det' \mathbf{M} \mathbf{M}_0^{-1}$ (equations (38.21)). We can multiply \mathbf{M} and \mathbf{M}_0 by the same operator. A specific choice which makes full use of the scale invariance of the classical theory, then transforms \mathbf{M} into an operator whose eigenvalues can be calculated analytically. Because the calculations are somewhat tedious, we only indicate here the main steps, without giving all details.

The transformation. We extend the transformation (38.39) to arbitrary fields, setting

$$\phi(x) = e^{-t} h(t, \hat{n}), \quad \text{with } t = \ln |x|, \quad \hat{n}_n = x^\mu / |x|. \quad (38.47)$$

The classical action $\mathcal{S}(\phi) = \tilde{\mathcal{S}}(h)$ then becomes

$$\tilde{\mathcal{S}}(h) = \int dt d\Omega \left\{ \frac{1}{2} \left[\dot{h}(t, \hat{n}) - h(t, \hat{n}) \right]^2 + h(t, \hat{n}) \mathbf{L}^2 h(t, \hat{n}) + \frac{1}{4} g h^4(t, \hat{n}) \right\}. \quad (38.48)$$

The symbol $\int d\Omega$ means integration over the angular variables \hat{n} , and \mathbf{L}^2 is the square of the angular momentum operator with eigenvalues $l(l+2)$ and degeneracy $(l+1)^2$. The expression (38.48) can be rewritten as

$$\tilde{\mathcal{S}}(h) = \int dt d\Omega \left\{ \frac{1}{2} \left[(\dot{h}(t, \hat{n}))^2 + h(t, \hat{n}) (\mathbf{L}^2 + 1) h(t, \hat{n}) \right] + \frac{1}{4} g h^4(t, \hat{n}) \right\}, \quad (38.49)$$

because the integral of the cross term $\dot{h}h$ vanishes due to the boundary conditions.

With the parametrization

$$\lambda = e^{-t_0}, \quad \mathbf{x}_0 = e^{t_0} \mathbf{v},$$

the classical solution (38.42b) transforms into

$$h_c(t) = \pm \frac{2(-2/g)^{1/2}}{e^{(t-t_0)} - 2\mathbf{v} \cdot \mathbf{n} + e^{-(t-t_0)}(\mathbf{v}^2 + 1)}. \quad (38.50)$$

We note that, in these new variables, translations take a complicated form, unlike dilatation which simply corresponds to a translation of the variable t .

The second derivative of the classical action at the saddle point now takes the form (for $t_0 = x_{0\mu} = 0$)

$$\mathbf{M} = \frac{\delta^2 \mathcal{S}}{\delta h_c \delta h_c} = - \left(\frac{d}{dt} \right)^2 + \mathbf{L}^2 + 1 - \frac{6}{\cosh^2 t}. \quad (38.51)$$

The natural measure associated to this Hamiltonian problem is

$$\int dt d\Omega,$$

which, in the original variables, is

$$\int \frac{d^4 x}{\mathbf{x}^2}.$$

This measure is not translation invariant, and thus the Jacobian resulting from the introduction of collective coordinates, and the determinant depend individually on $x_{0\mu}$. However, the product of the corresponding contributions to the final result should not, thus we perform the calculation for $x_{0\mu} = 0$.

38.4.1 The Jacobian

With the new measure, $d\phi_c/d\lambda$ is normalizable. Indeed,

$$J_1 = \left[\int \frac{d^4x}{\mathbf{x}^2} \left(\frac{d}{d\lambda} \phi_c(x) \right)^2 \right]^{1/2}, \quad (38.52)$$

$$= \left[\frac{16\pi^2}{(-g)} \int_0^\infty r dr \frac{(1 - \lambda^2 r^2)^2}{(1 + \lambda^2 r^2)^4} \right]^{1/2}. \quad (38.53)$$

This leads to a first factor:

$$J_1 = \frac{1}{\lambda} \sqrt{\frac{8}{3}} \frac{\pi}{\sqrt{-g}}. \quad (38.54)$$

The second Jacobian J_2 is generated by the collective coordinates $x_{0\mu}$:

$$J_2 = \left[\frac{1}{4} \int \frac{d^4x}{\mathbf{x}^2} \sum_{\mu=1}^4 (\partial_\mu \phi_c(x))^2 \right]^2, \quad (38.55)$$

$$= \frac{1}{g^2} \left[16\pi^2 \int_0^\infty \frac{r^3 dr \lambda^6}{(1 + \lambda^2 r^2)^4} \right]^2 = \frac{\lambda^4}{g^2} \times \frac{16}{9}\pi^4, \quad (38.56)$$

where rotation invariance has been used. The complete Jacobian J is thus

$$J = J_1 J_2 = \frac{\lambda^3}{(-g)^{5/2}} \pi^5 \times \frac{32\sqrt{2}}{9\sqrt{3}}. \quad (38.57)$$

38.4.2 The determinant

In the angular momentum basis, for each value l of the angular momentum, the component of the operator (38.51) reads

$$M_l = - \left(\frac{d}{dt} \right)^2 + (1 + l)^2 - \frac{6}{\cosh^2 t}. \quad (38.58)$$

Using equation (37.31), it is possible to calculate the determinant of M_l . One finds

$$\det(M_l + \varepsilon)(M_{0l} + \varepsilon)^{-1} = \frac{\sqrt{\varepsilon + (l+1)^2} - 1}{\sqrt{\varepsilon + (l+1)^2} + 2} \frac{\sqrt{\varepsilon + (l+1)^2} - 2}{\sqrt{\varepsilon + (l+1)^2} + 1}, \quad (38.59)$$

in which M_{0l} is the operator of the corresponding free theory. As we know, this determinant is UV divergent and has to be renormalized. However, we first calculate formally the unrenormalized determinant:

$$l \geq 2 : \quad \det M_l M_{0l}^{-1} = \frac{l(l-1)}{(l+2)(l+3)}, \quad (38.60)$$

$$l = 1 : \quad \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \det(M_1 + \varepsilon)(M_{01} + \varepsilon)^{-1} = \frac{1}{48}, \quad (38.61)$$

$$l = 0 : \quad \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \det(M_{l=0} + \varepsilon)(M_{0l=0} + \varepsilon)^{-1} = -\frac{1}{12}. \quad (38.62)$$

As expected the determinant is negative, and we obtain the formal expression

$$\det' \mathbf{M} \mathbf{M}_0^{-1} = -\frac{1}{12} \times \left(\frac{1}{48} \right)^4 \times \prod_{l=2}^{\infty} \left[\frac{l(l-1)}{(l+2)(l+3)} \right]^{(l+1)^2}. \quad (38.63)$$

Renormalization. In these variables, the UV divergences appear as divergences of the infinite product on l . As an intermediate step, we use a maximum value L of l as a cut-off. From the general analysis, we know the UV divergent part of $\ln \det \mathbf{M}$ is completely contained in the two first terms of the expansion in powers of ϕ_c^2 . Therefore, we proceed in the following way: the determinant of the operator

$$\mathbf{M}(s) = -\left(\frac{d}{dt} \right)^2 - \frac{s(s+1)}{\cosh^2 t}, \quad (38.64)$$

is exactly known:

$$\det [\mathbf{M}(s) + z] [\mathbf{M}_0 + z]^{-1} = \frac{\Gamma(1 + \sqrt{z}) \Gamma(\sqrt{z})}{\Gamma(1 + s + \sqrt{z}) \Gamma(\sqrt{z} - s)}. \quad (38.65)$$

Setting:

$$s(s+1) = 6\gamma, \quad (38.66)$$

one expands $\ln \det \mathbf{M}(s)$ in powers of γ . One infers from this expansion, the expansion up to second order of $\ln \det \mathbf{M}$ in powers of the potential $-6/\cosh^2 t$ in the representation (38.59). One then subtracts these two terms from $\ln \det \mathbf{M}$ as obtained from the representation (38.63).

One verifies that, indeed, the large L limit of the subtracted quantity,

$$\begin{aligned} \{\det' \mathbf{M} \mathbf{M}_0^{-1}\}_{\text{ren.}}^{-1/2} &= \lim_{L \rightarrow +\infty} i2\sqrt{3} \times (48)^2 \prod_{l=2}^L \left[\frac{(l+2)(l+3)}{(l-1)} \right]^{(l+1)^2/2} \prod_{l=0}^L e^{-3(l+1)} \\ &\times \prod_{l=0}^L e^{-18(l+1)^2} \left[\sum_{k=l+1}^{\infty} \frac{1}{k^2} - \frac{1}{l+1} - \frac{1}{2(l+1)^2} \right], \end{aligned} \quad (38.67)$$

is finite. We set:

$$\{\det' \mathbf{M} \mathbf{M}_0^{-1}\}_{\text{ren.}}^{-1/2} = iC_1. \quad (38.68)$$

Taking into account the Jacobians, the factor $(2\pi)^{-1/2}$ for each collective mode, the factor $(2i)^{-1}$, and a factor 2 for the two saddle points, one obtains a first factor C_2 of the form

$$C_2 = \frac{\lambda^3}{(-g)^{5/2}} \times \pi^5 \times \frac{32\sqrt{2}}{9\sqrt{3}} \times \frac{C_1}{(2\pi)^{5/2}}, \quad (38.69)$$

which we write as

$$C_2 = C_3 \lambda^3 / (-g)^{5/2}. \quad (38.70)$$

We then have to add to the classical action the two terms we have subtracted above from $\ln \det \mathbf{M}$. However, we can now write them in the normal space representation, regularized as we have regularized the perturbative correlation functions, and take into account the one-loop counter-terms. The first term in the expansion in powers of ϕ_c^2 is exactly cancelled by the mass counter-term, as we have already discussed. The second term in the expansion, which is the one-loop contribution to the four-point function, is logarithmically divergent. In Section 38.5, we calculate explicitly the finite difference between this term and the coupling constant counter-term that cancels the divergence.

38.5 Coupling constant renormalization

The terms we want to calculate involve the renormalized four-point function. We first choose a renormalization scheme for the field based on minimal subtraction (MS) after dimensional regularization. The renormalization constants have been calculated in Section 10.5. Note the different normalization of the coupling constant. The contribution $\delta\mathcal{S}_2$, which we have to add to the action, coming from the subtraction of $\ln \det \mathbf{M}$ and the one-loop coupling renormalization constant, is

$$\delta\mathcal{S}_2 = \frac{9}{4} \frac{N_d}{\varepsilon} g^2 \int \phi_c^4(x) d^4x - \frac{9}{4} g^2 \text{tr} \left[\phi_c^2(-\nabla^2)^{-1} \phi_c^2(-\nabla^2)^{-1} \right], \quad (38.71)$$

in which N_d is the usual loop factor,

$$N_d = 2(4\pi)^{-d/2}/\Gamma(d/2), \quad (38.72)$$

and $d = 4 - \varepsilon$. The expression can be rewritten as

$$\begin{aligned} \delta\mathcal{S}_2 &= -\frac{9}{4} g^2 \int d^4x d^4y \frac{d^4p}{(2\pi)^4} e^{ip(x-y)} \phi_c^2(x) \phi_c^2(y) \\ &\times \lim_{d \rightarrow 4} \left(\int \frac{d^dq}{(2\pi)^d} \frac{\mu^\varepsilon}{q^2(p-q)^2} - \frac{N_d}{\varepsilon} \right), \end{aligned} \quad (38.73)$$

in which μ is the renormalization scale.

The integral over \mathbf{q} has been performed in Section 10.1 (see equation (10.6)),

$$\int \frac{d^dq}{(2\pi)^d} \frac{1}{q^2(p-q)^2} - \frac{N_d}{\varepsilon} = \frac{1}{8\pi^2} \left(\frac{1}{2} - \ln p \right) + O(\varepsilon). \quad (38.74)$$

We also introduce the Fourier transform of the function $f^2(r)$ (for $\lambda = 1$) ($f(r)$ being defined by equation (38.42a)),

$$v(p) = \frac{1}{(2\pi)^4} \int d^4x \frac{8 e^{ipx}}{(1+x^2)^2}. \quad (38.75)$$

The solution $\phi_c(x)$ depends on the scale λ . Rescaling the variables x , y , and p , we can then write the complete expression more explicitly as

$$\delta\mathcal{S}_2 = -\frac{9\pi^2}{2} \int d^4p v^2(p) \left[\frac{1}{2} - \ln(\lambda p/\mu) \right]. \quad (38.76)$$

From the definition of $v(p)$, we derive after a short calculation,

$$\int d^4p v^2(p) = \frac{2}{(3\pi^2)}, \quad (38.77)$$

$$\int d^4p \ln p v^2(p) = \frac{2}{3\pi^2} \left(\ln 2 + \gamma + \frac{1}{6} \right), \quad (38.78)$$

in which γ is Euler's constant: $\gamma = -\psi(1) = 0.577215\dots$. We then obtain

$$\delta\mathcal{S}_2 = 3 \ln \lambda/\mu - \ln C_4, \quad (38.79)$$

with

$$\ln C_4 = 1 - 3 \ln 2 - 3\gamma. \quad (38.80)$$

We note that the right-hand side of equation (38.79) now depends on the scale parameter λ . The interpretation of this result is the following: the coupling constant renormalization breaks the scale invariance of the classical theory and, therefore, the scale parameter λ remains in the expression. Moreover, the term proportional to $\ln \lambda$ together with the contribution from the classical action can be rewritten as

$$\frac{8\pi^2}{3g} - 3\ln \lambda/\mu = \frac{8\pi^2}{3g(\lambda)} + O(g), \quad (38.81)$$

in which $g(\lambda)$ is the effective coupling at the scale λ , solution of the RG equation

$$\frac{dg(\lambda)}{d\ln \lambda} = \beta[g(\lambda)], \quad (38.82)$$

with

$$\beta(g) = \frac{9}{8\pi^2} g^2 + O(g^3). \quad (38.83)$$

This property is expected. The renormalization of the perturbative expansion renders the instanton contribution, before integration over dilatation, finite. Therefore, this contribution should satisfy a RG equation, and the coupling constant g can be present only in the combination $g(\lambda)$, since λ fixes the scale in the calculation (for details, see Section A38.3).

38.6 The imaginary part of the n -point function

The complete contribution to the imaginary part of the n -point function then takes the form

$$\begin{aligned} \text{Im } Z^{(n)}(x_1, \dots, x_n) \\ \underset{g \rightarrow 0_-}{\sim} C_5 \int d^4 x_0 \int_0^\infty \frac{d\lambda}{\lambda} \lambda^4 \prod_{i=1}^n \frac{2\sqrt{2}\lambda}{1 + \lambda^2 (x_i - x_0)^2} \frac{e^{8\pi^2/3g(\lambda)}}{(-g)^{(n+5)/2}}, \end{aligned} \quad (38.84)$$

where we have set:

$$C_5 = C_3 C_4.$$

To calculate the Fourier transform of the expression (38.84), we introduce

$$u(p) = 2\sqrt{2} \int e^{ipx} \frac{d^4 x}{1+x^2}. \quad (38.85)$$

Then, after factorizing the δ -function of momentum conservation, one obtains

$$\text{Im } \tilde{Z}^{(n)}(p_1, \dots, p_n) \sim \frac{C_5}{(-g)^{(n+5)/2}} \int_0^\infty d\lambda \lambda^{3-3n} e^{8\pi^2/3g(\lambda)} \prod_{i=1}^n u(p_i/\lambda). \quad (38.86)$$

The corresponding expression for vertex functions is

$$\text{Im } \tilde{\Gamma}^{(n)}(p_1, \dots, p_n) \sim \frac{C_5}{(-g)^{(n+5)/2}} \int_0^\infty \frac{d\lambda}{\lambda} \lambda^{4-n} e^{8\pi^2/3g(\lambda)} \prod_{i=1}^n (p_i^2/\lambda^2) u(p_i/\lambda). \quad (38.87)$$

One verifies that $p^2 u(p)$ goes to a constant for $|p|$ small.

In contrast to the super-renormalizable case, because the theory is only renormalizable, the final result is not totally explicit, since it involves a final integration over dilatations whose convergence is not obvious. Let us now discuss this point.

The small instanton contribution. Small instantons correspond to λ large. For λ large, the integral behaves like

$$\int^{\infty} d\lambda \lambda^{3-n} e^{8\pi^2/3g(\lambda)} \quad (38.88)$$

and, therefore, we have to examine the behaviour of $g(\lambda)$ for λ large. From equation (38.83), we see that the theory is UV asymptotically free, because for g is negative, that is, $g(\lambda)$ goes to zero for λ large. Thus, perturbation theory is applicable, and we can use the approximation (38.81). The argument remains true even if we take g slightly complex. Thus, the integral has the form

$$\int^{\infty} d\lambda \lambda^{-n}. \quad (38.89)$$

We note that the power behaviour in λ depends explicitly on the coefficient of the g^2 term of the $\beta(g)$ -function. Without the contribution coming from $g(\lambda)$, the integral (38.89) would have a UV divergence similar to the one found in the corresponding perturbative expansion. Due to the additional power of λ coming from $g(\lambda)$, only the vacuum amplitude is divergent.

The convergence of the dilatation integral is thus better than expected: indeed, the renormalization constants are now themselves given by divergent series and are complex for g negative. Their imaginary part contributes directly to the imaginary part of $\tilde{\Gamma}^{(n)}(p_1, \dots, p_n)$ for $n \leq 4$.

In the ϕ^6 QFT in dimension 3, for example, these contributions cancel the divergences coming from the integral over λ . By contrast, here the integrals over λ are finite at this order. In particular, this implies that, in the MS scheme, the imaginary parts of the renormalization constants vanish at leading order. In another renormalization scheme (*e.g.* fixed-momentum subtraction), these imaginary parts are finite at leading order.

The large instanton contribution. We now examine the convergence of the λ integral for λ small. The behaviour of $g(\lambda)$ is unknown. On the other hand, it is easy to verify that the factors $u(p_i/\lambda)$ decrease exponentially for λ small. Thus, if the behaviour of $g(\lambda)$ does not cancel this decrease, the integrals converge, and it is justified to replace $g(\lambda)$ by the expansion (38.81). For the vacuum amplitude, this argument does not apply, and so the result is unknown.

This analysis shows that, although this calculation seems to be a simple formal extension of the calculation for lower dimensions, coupling constant renormalization introduces a set of new problems, which are not all completely under control. The fact that the theory is massless only makes matters worse. Consideration of the massive theory improves the situation in this respect, but the instanton calculation becomes more involved.

38.7 The massive theory

In Section A38.2, we show that the massive field equation has no instanton solutions, and that the minimum of the action is obtained from the massless theory. To study the massive theory, we thus start from the instanton solution of the massless theory with its scale parameter λ . However, a problem appears: as explained in Section 38.4, the integral of ϕ_c^2 is IR divergent.

Thus, it is necessary to modify the field configuration at large distances, by connecting it smoothly to the solution of the massive free equation with mass m . The idea is to define a configuration which, up to a distance R , $\lambda R \gg 1$, $mR \ll 1$, is $\lambda\phi_c(\lambda x)$, and for $|x| > R$ is proportional to the free massive solution. An analogous problem is met in Chapter 42 in the case of multi-instanton configurations. Although the theory is no longer scale invariant, λ has to be kept as a collective coordinate. The mass term then acts as an IR cut-off, and restrict the domain of integration in λ to values large compared to m . The classical action has the form

$$\mathcal{S}_m(\phi_c) = -\frac{1}{g} \left(\frac{8\pi^2}{3} + 8\pi^2 \frac{m^2}{\lambda^2} \ln \frac{\lambda}{m} \right), \quad \text{for } \lambda \gg m, \quad (38.90)$$

where the $\ln m$ term is directly related to the initial IR divergence of the ϕ^2 integral.

The remaining part of the calculation closely follows the calculation for the massless case, and the reader is referred to the literature for details [389].

In the massless theory, the instanton contribution to the vacuum energy could not be evaluated without some knowledge of the non-perturbative IR behaviour of the RG β -function. In the massive theory the problem is absent, because the λ integral is cut at a scale $m/\sqrt{-g}$. For correlation functions, the integral is cut by the largest between momenta and $m/\sqrt{-g}$. This implies that the limits $m \rightarrow 0$ and $g \rightarrow 0$ do not commute.

38.8 Cosmology: The decay of the false vacuum

In previous sections, we have determined the probability for a ‘false vacuum’ of a QFT to decay through barrier penetration. It had been speculated [390] that such a phenomenon could be linked to the dynamics of the early Universe. When the Universe started to cool down, some symmetries started to be spontaneously broken. However, some region might have been trapped in the wrong phase. The false vacuum must eventually decay in the true vacuum, but if the process is slow enough, it might have occurred at a much later time when the Universe was already cool. This kind of physics speculation can be studied by an instanton approach [391].

According to the previous discussion, if the Universe is in the wrong vacuum, there is some probability at each point in space for some bubble of true vacuum to be created, and if the bubble is large enough, it becomes favourable for it to expand, eventually absorbing the whole space. To discuss what happens once a bubble has been created, it is useful to consider first the analogous problem for a quantum particle.

Quantum particle. In the example of a quantum particle, a semi-classical description of the decay process is the following: a particle is sitting in the well of the potential corresponding to the metastable minimum. At a given time, it makes a quantum jump and reappears outside of the barrier at the point where the potential has the same value as in the bottom of the well, with zero velocity (by energy conservation). Then its further trajectory can be entirely described by classical mechanics.

QFT. We apply the same ideas to the field theoretical model we discuss in this section. At time 0, the system makes a quantum jump. According to the previous discussion, the value of the field at time 0 is then (with the choice $x_{0\mu} = 0$)

$$\phi(t=0, \mathbf{x}) = \phi_c(x_d = 0, \mathbf{x}), \quad (\mathbf{x} = x_1, \dots, x_{d-1}), \quad (38.91)$$

and its time derivative vanishes,

$$\left. \frac{\partial}{\partial t} \phi(t, \mathbf{x}) \right|_{t=0} = 0. \quad (38.92)$$

At a later time, $\phi(t, \mathbf{x})$ then obeys the *real-time* field equation,

$$[\nabla_x^2 - (\partial/\partial t)^2] \phi(t, \mathbf{x}) = \frac{1}{\sqrt{g}} V'(\sqrt{g\phi}(t, \mathbf{x})). \quad (38.93)$$

The first equation (38.91) tells us that the same function describes the form of the instanton in Euclidean space, and its shape in ordinary $(d-1)$ space when it materializes. We now consider the continuation in real time of the solution of the Euclidean field equation $\phi_c[(\mathbf{x}^2 - t^2)^{1/2}]$ (since $\phi_c(r)$ is an even function, the sign in front of the square root is irrelevant). It satisfies the conditions (38.91, 38.92), and clearly obeys the field equation (38.93). Therefore, it is the solution of our problem for positive times.

Since the size of the bubble is given by microphysics, the interior of the bubble corresponds to small values of r on a macroscopic scale,

$$0 \leq \mathbf{x}^2 - t^2 = r^2 \ll 1.$$

Therefore, after a short time, the bubble starts expanding at almost the speed of light.

A38 Instantons: Additional remarks

Here we prove a few simple relations and inequalities concerning instanton solutions, which we have used in several places in the chapter, discuss the RG properties of instanton contributions, and the conformal invariance of the $\phi_{d=4}^4$ massless QFT.

A38.1 Virial theorem

We consider the general action, with polynomial potential,

$$\mathcal{S}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \mathcal{V}(\phi(x)) \right], \quad (A38.1)$$

and assume that the field equation has a finite action solution $\phi_c(x)$. If the action $\mathcal{S}(\phi_c)$ is finite, so is the action $\mathcal{S}(\phi_c, \lambda)$ for $\phi(x) = \phi_c(\lambda x)$, where λ is an arbitrary constant. If we now change variables in the action setting $\lambda x = x'$, we obtain [386]

$$\mathcal{S}(\phi_c, \lambda) = \lambda^{2-d} \int \frac{1}{2} (\nabla \phi_c(x))^2 d^d x + \lambda^{-d} \int \mathcal{V}(\phi_c(x)) d^d x. \quad (A38.2)$$

Since $\phi_c(x)$ satisfies the field equation, the variation of the action vanishes for $\lambda = 1$:

$$\frac{d}{d\lambda} \mathcal{S}(\phi_c, \lambda) \Big|_{\lambda=1} = 0 \Rightarrow (d-2) \int \frac{1}{2} (\nabla \phi_c(x))^2 d^d x + d \int \mathcal{V}(\phi_c(x)) d^d x = 0. \quad (A38.3)$$

The classical action $\mathcal{S}(\phi_c)$ can thus be expressed in terms of the kinetic term only,

$$\mathcal{S}(\phi_c) = \frac{1}{d} \int (\nabla \phi_c(x))^2 d^d x > 0, \quad (A38.4)$$

a form that shows that $\mathcal{S}(\phi_c)$ is always positive.

It is also interesting to calculate the second derivative of $\mathcal{S}(\phi_c, \lambda)$,

$$\frac{d^2}{(d\lambda)^2} \mathcal{S}(\phi_c, \lambda) \Big|_{\lambda=1} = (2-d) \int (\nabla \phi_c(x))^2 d^d x. \quad (A38.5)$$

For $d \geq 2$, the solution is not a local minimum of the action, and the operator defined by

$$\langle x | \mathbf{M} | x' \rangle = \frac{\delta^2 \mathcal{S}}{\delta \phi(x) \delta \phi(x')} \Big|_{\phi=\phi_c},$$

has at least one negative eigenvalue.

Moreover, a general theorem [385] states that $\phi_c(x)$ corresponds to an absolute minimum of $\mathcal{S}(\phi)$ at fixed integral of the potential $\int d^d x \mathcal{V}(\phi_c(x))$. If \mathbf{M} has two negative eigenvalues, one can then find a linear combination of the corresponding two eigenvectors which, added to $\phi_c(x)$, leaves at first-order the integral of the potential unchanged, and decreases $\mathcal{S}(\phi)$. This contradicts the theorem. Thus \mathbf{M} has at most one negative eigenvalue. Since the equation (A38.5) shows that \mathbf{M} has at least one negative eigenvalue, it has one and only one.

Special potentials: Other relation. We consider potentials of the special form ($N > 2$),

$$\mathcal{V}(\phi) = \frac{1}{2}m^2\phi^2 + g\phi^N. \quad (A38.6)$$

If the action $\mathcal{S}(\phi_c)$ is finite, so is $\mathcal{S}(\phi_c/\lambda)$. Again, if ϕ_c is a solution, the derivative of $\mathcal{S}(\phi_c/\lambda)$ for $\lambda = 1$ vanishes. This yields

$$\frac{d\mathcal{S}(\phi_c/\lambda)}{d\lambda} \Big|_{\lambda=1} = -2 \int \frac{1}{2} \left[(\nabla\phi_c(x))^2 + m^2\phi_c^2(x) \right] d^d x - Ng \int \phi_c^N(x) d^d x = 0. \quad (A38.7)$$

Combining the relations (A38.3) and (A38.7), for $2d - N(d - 2) \neq 0$, one infers

$$\mathcal{S}(\phi_c) = \frac{N-2}{2d-N(d-2)} m^2 \int \phi_c^2(x) d^d x. \quad (A38.8)$$

Since equation (A38.4) implies $\mathcal{S}(\phi_c) > 0$, equation (A38.8) can only be satisfied for $2 < N < 2d/(d-2)$. Therefore, the existence of instanton solutions implies that the QFT must be super-renormalizable.

In the marginal renormalizable case $N = 2d/(d-2)$ (($N = 3, d = 6$), ($N = 4, d = 4$), ($N = 6, d = 3$)), one derives

$$m^2 \int \phi_c^2(x) d^d x = 0,$$

which is only satisfied for $m = 0$. The massive field equations have also no instanton solution, only the massless equations do. We examine this problem more thoroughly in Section A38.2, in the example $N = 4$.

A38.2 Sobolev inequalities

We consider the functional

$$R(\varphi) = \frac{\left\{ \int d^d x \left[(\nabla\varphi(x))^2 + \varphi^2(x) \right] \right\}^2}{\int \varphi^4(x) d^d x}. \quad (A38.9)$$

For dimensions $d \leq 4$, Sobolev inequalities imply [392]

$$R(\varphi) \geq R > 0. \quad (A38.10)$$

In addition, for $d < 4$, there exists a spherically symmetric, zero-free function $\varphi_c(x)$, such that $R(\varphi_c) = R$, and which is a solution of the variational equation

$$\frac{\delta R}{\delta \varphi(x)} \Big|_{\varphi=\varphi_c} = 0. \quad (A38.11)$$

Dimension smaller than 4. The equation (A38.11) has the explicit form

$$(-\nabla^2 + 1) \varphi(x) - \varphi^3(x) K = 0, \quad (A38.12)$$

in which we have defined

$$K = \int d^d x \left[(\nabla\varphi_c(x))^2 + \varphi_c^2(x) \right] / \int \varphi_c^4(x) d^d x. \quad (A38.13)$$

This equation is, up to a rescaling of $\varphi_c(x)$, the equation of motion (38.6). Both equations become identical if we choose the scale of $\varphi_c(x)$, which is otherwise arbitrary, such that

$$K = 1 \Rightarrow f(x) = \varphi_c(x). \quad (A38.14)$$

For each instanton solution, we have derived the identities (38.14). Combining them with $K = 1$, we obtain

$$A = \frac{3}{2} \int d^d x f^4(x) = \frac{3}{2} R(\varphi_c). \quad (A38.15)$$

The smallest action solution thus corresponds to the minimum of $R(\varphi)$:

$$A = 3R/2, \quad (A38.16)$$

and the solution $f(x)$ we are looking for is given by

$$f(x) = \varphi_c(x), \quad \text{for } K = 1. \quad (A38.17)$$

The introduction of the functional $R(\varphi)$ has the following advantage: the action (38.1) is obviously not bounded from below. But, if the fields $\phi(x)$ are restricted to be solutions of the equation of motion with finite action, then the action can be related to the functional $R(\varphi)$, which is bounded from below for all fields.

We then derive the property that the operator \mathbf{M} has one and only one negative eigenvalue from the form of $R(\varphi)$, and the assumption that φ_c corresponds to an absolute minimum of R . The operator defined by the kernel

$$\frac{\delta^2 R}{\delta \varphi(x) \delta \varphi(x')} \Big|_{\varphi=\varphi_c}, \quad (A38.18)$$

is positive. Calculating the second functional derivative of R explicitly, one obtains

$$\frac{\delta^2 R}{\delta \varphi(x) \delta \varphi(x')} \Big|_{\varphi=\varphi_c} = 4 \left\{ [-\nabla^2 + 1 - 3\varphi_c^2(x)] \delta^{(d)}(x - x') + \frac{2\varphi_c^3(x)\varphi_c^3(x')}{\int \varphi_c^4(y) d^d y} \right\}. \quad (A38.19)$$

We have again set $K = 1$. We will now express \mathbf{M} in terms of $f(x)$, or $\varphi_c(x)$, for $m = 1$,

$$\langle x | \mathbf{M} | x' \rangle = [-\nabla^2 + 1 - 3\varphi_c^2(x)] \delta^{(d)}(x - x'). \quad (A38.20)$$

Therefore, we have derived the relation

$$\langle x | \mathbf{M} | x' \rangle = \frac{1}{4} \frac{\delta^2 R}{\delta \varphi(x) \delta \varphi(x')} \Big|_{\varphi=\varphi_c} - 2 \left(\frac{\varphi_c^3(x') \varphi_c^3(x)}{\int \varphi_c^4(y) d^d y} \right). \quad (A38.21)$$

(i) Since $R(\varphi)$ is invariant in the change $\varphi_c(x)$ in $\lambda \varphi_c(x)$, φ_c is an eigenvector of $\delta^2 R / (\delta \varphi_c)^2$ with eigenvalue 0, thus

$$\int \varphi_c(x') \varphi_c(x) \langle x | \mathbf{M} | x' \rangle d^d x d^d x' = -2 \int \varphi_c^4(x) < 0. \quad (A38.22)$$

The operator \mathbf{M} has at least one negative eigenvalue.

(ii) Since \mathbf{M} is the sum of a positive operator and a projector of rank 1, it can have at most one negative eigenvalue.

Indeed, if it had two negative eigenvalues, one could find a linear combination of the corresponding two eigenvectors which would decrease \mathbf{M} at an average of the projector fixed. This would imply that the kernel $\delta^2 R / \delta\varphi(x)\delta\varphi(x')|_{\varphi=\varphi_c}$ does not define a positive operator.

We conclude that \mathbf{M} has *one and only one* negative eigenvalue.

Dimension 4. Let us calculate $R[\varphi(\lambda x)]$ for $d \leq 4$. Then, changing λx in x' in the various integrals, we obtain

$$R[\varphi(\lambda x)] = \frac{\left\{ \int d^d x \left[\lambda^{2-d} (\nabla \varphi(x))^2 + \lambda^{-d} \varphi^2(x) \right] \right\}^2}{\lambda^{-d} \int \varphi^4(x) d^d x}. \quad (A38.23)$$

We can now write

$$R = \min_{\{\varphi(x)\}} \min_{\lambda} R[\varphi(\lambda x)]. \quad (A38.24)$$

The minimum in λ of expression (A38.23) is obtained for

$$\lambda = \left[\frac{d}{(4-d)} \frac{\int \varphi^2(x) d^d x}{\int (\nabla \varphi(x))^2 d^d x} \right]^{1/2}, \quad (A38.25)$$

and equation (A38.24) becomes

$$R = \min_{\{\varphi(x)\}} \frac{16}{d^{d/2} (4-d)^{(4-d)/2}} \frac{\left(\int (\nabla \varphi(x))^2 d^d x \right)^{d/2} \left(\int \varphi^2(x) d^d x \right)^{2-d/2}}{\int \varphi^4(x) d^d x}. \quad (A38.26)$$

For $d = 4$, we note that the solution is $\lambda = \infty$, and expression (A38.26) is just the equivalent of expression (A38.9) for the massless ϕ^4 QFT. Since, for $d = 4$, the massless ϕ^4 is scale invariant, the contribution of the mass term can be arbitrarily decreased by a rescaling of the variable x .

We can draw two interesting conclusions from this analysis: the minimal value of $R(\varphi)$ is the same in four dimensions for the massive and the massless theory. The same applies to the ϕ^4 action.

The minimum of $R(\varphi)$ is obtained from a solution of the massless field equation. The massive field equation has no solution.

These remarks explain a number of peculiarities of the ϕ^4 QFT in four dimensions, which we have discussed in Section 38.3.

A38.3 Instantons and RG equations

We now briefly describe a few RG properties of the instanton contributions in the $\phi_{d=4}^4$ QFT (see Section 38.3).

The instanton contribution to the n -point vertex function can be written as

$$\text{Im } \tilde{\Gamma}^{(n)}(p_i; \mu, g) = \int_0^\infty \frac{d\lambda}{\lambda} F^{(n)}(p_i; \mu, g, \lambda), \quad (A38.27)$$

in which μ represents the subtraction scale, and λ the dilatation parameter. The counter-terms that renormalize the perturbative expansion, also render $F^{(n)}$ finite for reasons we have explained. Therefore, $F^{(n)}$ satisfies the RG equation,

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \eta(g) \right] F^{(n)}(p_i; \mu, g, \lambda) = 0. \quad (A38.28)$$

Integrated by the method of characteristics, the equation yields

$$F^{(n)}(p_i; \mu; g; \lambda) = Z^{-n/2}(\tau) F^{(n)}(p_i; \mu\tau, g(\tau), \lambda), \quad (A38.29)$$

with the definitions:

$$\begin{aligned} \ln \tau &= \int_g^{g(\tau)} \frac{dg'}{\beta(g')}, \\ \ln Z &= \int_g^{g(\tau)} \frac{\eta(g')}{\beta(g')} \end{aligned} \quad (A38.30)$$

The coupling constant $g(\tau)$ is the effective coupling constant at the scale τ .

Dimensional analysis implies the scaling relation

$$F^{(n)}(p_i; \mu; g; \lambda) = \lambda^{4-n} F^{(n)}(p_i/\lambda, \mu/\lambda, g, 1). \quad (A38.31)$$

Applied to the right-hand side of equation (A38.29), this identity yields

$$F^{(n)}(p_i; \mu; g; \lambda) = Z^{-n/2}(\tau) \lambda^{4-n} F^{(n)}\left(\frac{p_i}{\lambda}, \frac{\mu\tau}{\lambda}, g(\tau)\right). \quad (A38.32)$$

Finally, the choice

$$\tau = \lambda/\mu$$

leads to the relation

$$F^{(n)}(p_i; \mu; g; \lambda) = [Z(\lambda/\mu)]^{-n/2} \lambda^{4-n} F^{(n)}[p_i/\lambda; g(\lambda/\mu)]. \quad (A38.33)$$

A38.4 Conformal invariance

In Section 38.3, the scale invariance of the classical $\phi_{d=4}^4$ field theory has made it possible to obtain an analytic instanton solution. Moreover, by introducing the special coordinates (t, n_μ) , we have been able to use the results obtained for the anharmonic oscillator in Chapter 37 and calculate explicitly the instanton contribution at leading order. Actually, the scale invariant classical $\phi_{d=4}^4$ QFT is also conformal invariant (see Section A13.3). This property, which also holds for other scale invariant field theories like gauge theories, can be used more directly to calculate the instanton contribution [384]. The conformal group is isomorphic to $SO(5, 1)$. It is expected that the minimal action solution will be invariant under $SO(5)$, the maximal compact subgroup of $SO(5, 1)$. A stereographic mapping of \mathbb{R}^4 onto the sphere S_4 can be used to simplify the $SO(5)$ transformations. One sets:

$$\xi^\mu = \frac{2x^\mu}{1 + \mathbf{x}^2}, \quad \xi^5 = \frac{1 - \mathbf{x}^2}{1 + \mathbf{x}^2} \quad \Rightarrow \quad \sum_{a=1}^5 \xi^a \xi^a = 1. \quad (A38.34)$$

Correspondingly, one can introduce a field that has simple transformation properties under $SO(5)$. In the $\phi_{d=4}^4$ QFT, the conformal transformation properties of the field ϕ lead to set:

$$\phi(x) = \frac{1}{1 + \mathbf{x}^2} \psi(x). \quad (A38.35)$$

We then express the classical action (38.35) in terms of these new variables by performing the transformations in two steps: first we keep the variables x^μ , but now considered as coordinates on S_4 , and only perform the substitution (A38.35).

The metric $g_{\mu\nu}$ on S_4 in the coordinates x^μ is

$$g_{\mu\nu} = 4 \frac{\delta_{\mu\nu}}{(1 + \mathbf{x}^2)^2}. \quad (A38.36)$$

The invariant measure on the sphere involves the square root of the determinant of the metric \mathbf{g} (see Section 28.3.1):

$$\sqrt{\det \mathbf{g}} = \frac{16}{(1 + \mathbf{x}^2)^4}. \quad (A38.37)$$

Finally, after an integration by parts, the kinetic term can be rewritten as

$$\int d^4x (\nabla\phi(x))^2 = \int d^4x \left[\frac{(\nabla\psi(x))^2}{(1 + \mathbf{x}^2)^2} + \frac{8\psi^2(x)}{(1 + \mathbf{x}^2)^4} \right]. \quad (A38.38)$$

The classical action then reads

$$\mathcal{S}(\psi) = \int d^4x \sqrt{\det \mathbf{g}} \left(\frac{1}{8} \sum_{\mu,\nu} g^{\mu\nu}(x) \partial_\mu \psi(x) \partial_\nu \psi(x) + \frac{1}{4} \psi^2(x) + \frac{1}{64} g \psi^4(x) \right). \quad (A38.39)$$

In this covariant form (see Section 28.3), the change of coordinates (A38.34) is straightforward and hardly necessary. One solution of minimal action is a constant:

$$\psi^2(x) = -1/8g. \quad (A38.40)$$

The classical action is then proportional to the surface of S_4 which is $8\pi^2/3$. The operator \mathbf{M} , second functional derivative of the action, is given by (see Section 3.5)

$$\mathbf{M} = \frac{1}{4}\mathbf{L}^2 - 1, \quad (A38.41)$$

in which \mathbf{L} is the angular momentum in five space dimensions. The eigenvalues of \mathbf{L}^2 are $l(l+3)$ with the degeneracy

$$\delta_l = \frac{1}{6} \frac{(2l+3)\Gamma(l+3)}{\Gamma(l+1)}. \quad (A38.42)$$

The form of \mathbf{M} shows that it has 0 as eigenvalue, corresponding to $l=1$, with degeneracy 5, in agreement with the considerations of Section 38.4. We leave it up to the reader, as an exercise, to verify other results.

39 Degenerate classical minima and instantons

In this chapter, we study a situation in which instantons play an important role: quantum theories corresponding to classical actions that have non-continuously connected degenerate minima. The simplest examples are provided by one-dimensional quantum systems where the potential has degenerate minima. Classically, the states of minimum energy correspond to a particle sitting at any of the minima of the potential. In the case of symmetric minima, the position of the particle breaks (spontaneously) the symmetry of the system. By contrast, in quantum mechanics (QM) the modulus of the ground-state wave function is expected to be large near all the minima of the potential, as a consequence of *barrier penetration effects*. We illustrate this phenomenon with two typical examples: the double-well potential [393], and the cosine potential, whose periodic structure is closer to field theory examples.

In the context of stochastic dynamics, in Section 39.3, we relate instantons to Arrhenius law. The proof of the existence of instantons relies on an inequality related to supersymmetric structures, and which generalizes to some field theory examples.

In field theory, the problem is more subtle as the study of phase transitions shows. However, the presence of instantons again indicates that the classical minima are connected by quantum tunnelling, and that the symmetry between them is not spontaneously broken [394]. Examples of such a situation are provided, in two dimensions, by the $CP(N-1)$ models and, in four dimensions, by $SU(2)$ gauge theories.

39.1 The quartic double-well potential

We first discuss the Hamiltonian of the quartic double-well potential,

$$H = -\frac{1}{2} \left(\frac{d}{dq} \right)^2 + V(q\sqrt{g})/g, \quad g > 0, \quad (39.1)$$

with (g is a loop expansion parameter),

$$V(q) = \frac{1}{2}q^2(1-q)^2. \quad (39.2)$$

The Hamiltonian commutes with the operator P , which acts on wave functions as

$$P\psi(q) = \psi(g^{-1/2} - q) \Rightarrow P^2 = 1, \quad [H, P] = 0.$$

Correspondingly, the potential V has two degenerate minima located at $q = 0$ and at $q = 1/\sqrt{g}$. The symmetry is not essential for the existence of instanton solutions. It is a simplifying feature, which, moreover, is present in several examples of physical interest.

39.1.1 The structure of the ground state

Due to the symmetry of the potential, one can generate a perturbative expansion starting from each of the minima of the potential, and one finds the same expansion to all orders. Therefore, one could conclude that the quantum Hamiltonian has a doubly degenerate ground state, corresponding to two eigenfunctions concentrated, respectively, around each of the classical minima of the potential.

However, due to barrier penetration, the true eigenstates are eigenstates of the reflection operator P , the ground state being an even state.

The reflection symmetry cannot be spontaneously broken in QM in the case of regular potentials: correlation functions constructed with a Hamiltonian of this type have, from the point of view of phase transitions, the properties of correlation functions of the one-dimensional Ising model (see Section 14.1).

The partition function. Since one expands in g small first, in the large β limit, the partition function $\text{tr } e^{-\beta H}$ is dominated by the two lowest eigenvalues E_+ and E_- :

$$\begin{aligned} \text{tr } e^{-\beta H} &\sim e^{-\beta E_+} + e^{-\beta E_-} \sim 2e^{-\beta(E_+ + E_-)/2} \cosh[\beta(E_+ - E_-)/2] \\ \text{for } g \rightarrow 0, \quad \beta \rightarrow \infty. \end{aligned} \quad (39.3)$$

To all orders in g , the partition function only depends on the half sum $\frac{1}{2}(E_+ + E_-)$, and is only sensitive to the non-perturbative difference between the eigenvalues E_+ and E_- at order $(E_+ - E_-)^2$:

$$\begin{aligned} -\frac{1}{\beta} \ln \text{tr } e^{-\beta H} &= \frac{1}{2}(E_+ + E_-) - \frac{1}{\beta} \ln 2 + O[e^{-\beta}, \beta(E_+ - E_-)^2] \\ \text{for } g \rightarrow 0, \quad \beta \rightarrow \infty. \end{aligned} \quad (39.4)$$

By contrast, the difference $(E_+ - E_-)$ dominates the twisted partition function $\text{tr } P e^{-\beta H}$ (see Section 14.2). Indeed, in the same limits $g \rightarrow 0$, and $\beta \rightarrow \infty$, one finds

$$\text{tr } P e^{-\beta H} \sim e^{-\beta E_+} - e^{-\beta E_-} \sim -2 \sinh[\beta(E_+ - E_-)/2] e^{-\beta(E_+ + E_-)/2}. \quad (39.5)$$

$$\sim -\beta e^{-\beta/2}(E_+ - E_-) [1 + O(g, e^{-\beta})]. \quad (39.6)$$

Actually, it is convenient to consider the ratio between the quantities (39.3) and (39.5),

$$\text{tr } P e^{-\beta H} / \text{tr } e^{-\beta H} \sim -\frac{1}{2}\beta(E_+ - E_-) [1 + O(e^{-\beta}, (E_+ - E_-)^2)]. \quad (39.7)$$

The ratio makes it possible to distinguish between a situation in which the ground state is degenerate, and the symmetry spontaneously broken, and a situation in which quantum fluctuations restore the symmetry and lift the degeneracy between the two lowest lying states. Since the ratio vanishes to all orders in perturbation theory, one has to look for non-perturbative effects: they are due here to instantons.

39.1.2 Instanton contributions

The partition function is given by the path integral

$$\mathcal{Z}(\beta) = \int [dq(t)] \exp[-\mathcal{S}(q)], \quad (39.8)$$

where

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} [\frac{1}{2}\dot{q}^2(t) + V(q(t)\sqrt{g})/g] dt, \quad (39.9)$$

and the paths satisfy periodic boundary conditions: $q(-\beta/2) = q(\beta/2)$. The path integral representation of the twisted partition function $\text{tr } P e^{-\beta H}$ only differs by the boundary conditions, which are now $q(-\beta/2) + q(\beta/2) = g^{-1/2}$.

For $g \rightarrow 0$, the path integral (39.8) is dominated by the saddle points corresponding to the constant functions $q(t) \equiv 0$ or $q(t) \equiv g^{-1/2}$, and this leads to the usual perturbative expansion.

However, these paths do not contribute to the path integral representation of $\text{tr } P e^{-\beta H}$ because they do not satisfy the different boundary conditions. This is not too surprising, since the difference $E_+ - E_-$ vanishes to all orders in an expansion in powers of g . Therefore, we have to look for non-constant solutions of the equation of motion, which have a finite action in the infinite β limit. The boundary conditions then impose

$$q(\mp\infty) = 0 \quad \text{and} \quad q(\pm\infty) = g^{-1/2}. \quad (39.10)$$

The non-degeneracy of the ground state depends on quantum tunnelling, and the corresponding existence of instanton solutions connecting the two minima of the potential.

In the infinite β limit, in terms of $u(t) = \sqrt{g}q_c(t)$, the Euclidean equation of motion yields

$$-\ddot{u}(t) + V'(u(t)) = 0 \Rightarrow \frac{1}{2}\dot{u}^2(t) = V(u(t)),$$

in which the boundary conditions (39.10) have been taken into account. The equation has two one-parameter family of solutions with finite classical action, which we call instanton and anti-instanton when it is necessary to distinguish between them. There are given by

$$q_c^\pm(t) = u^\pm(t)/\sqrt{g}, \quad \text{with} \quad u^\pm(t) = \frac{1}{1 + e^{\mp(t-t_0)}}, \quad (39.11)$$

and, therefore,

$$\mathcal{S}(q_c) = \frac{1}{g} \int dt \left[\frac{1}{2}\dot{u}^2(t) + V(u(t)) \right] = \frac{1}{6g}. \quad (39.12)$$

Large β expansion. The methods of Section 37.6.1 can be adapted to the present problem. For β large but finite, from equations analogous to equations (37.84) and (37.85), one infers the expansions of the classical energy and action:

$$E(\beta) = -2e^{-\beta} + O(e^{-2\beta}), \quad (39.13)$$

$$g\mathcal{S}(q_c) = \frac{1}{6} - 2e^{-\beta} + O(e^{-2\beta}). \quad (39.14)$$

The determinant resulting from the integration around the saddle point can also be evaluated by the method explained in Section 37.5. The only noticeable modification stems from the property that $\dot{q}_c(t)$ has no zero: it corresponds to the ground state of the differential operator $\partial^2\mathcal{S}/\delta q(t)\delta q(t')|_{q=q_c}$, which, therefore, is a positive operator. The final result is real, as expected. We can use the expression (37.83) to obtain it, except that no $1/2i$ factor appears here, and one has to multiply by a factor 2, since the two solutions q_c^+ and q_c^- give identical contributions:

$$\text{tr } P e^{-\beta H} \sim \frac{2}{\sqrt{\pi g}} \beta e^{-\beta/2} e^{-1/6g} (1 + O(g)), \quad \text{for } g \rightarrow 0, \quad \beta \rightarrow \infty. \quad (39.15)$$

From equation (39.7), one then infers the asymptotic behaviour of $E_+ - E_-$ for g small,

$$E_+ - E_- \underset{g \rightarrow 0}{=} -\frac{2}{\sqrt{\pi g}} e^{-1/6g} (1 + O(g)). \quad (39.16)$$

The difference is exponentially small in $1/g$, a result that, as expected, for $g \rightarrow 0$, vanishes to all orders in g .

39.2 The periodic cosine potential

We now consider the slightly more complicated problem of the Hamiltonian

$$H = -\frac{1}{2} (\mathrm{d}/\mathrm{d}q)^2 + V(q\sqrt{g})/g, \quad (39.17)$$

with the periodic potential,

$$V(q) = 1 - \cos q. \quad (39.18)$$

Since the potential is periodic, the action has an infinite number of degenerate classical minima. Starting from any of these minima one obtains, to all orders in powers of g , the same perturbative spectrum and, therefore, the quantum Hamiltonian seems to have an infinite number of degenerate ground states. Actually, we know that the spectrum of the Hamiltonian H is continuous and has, at least for g small enough, a band structure: this property, for g small, again is due to barrier penetration.

39.2.1 The structure of the ground state

We introduce the unitary translation operator T , which acting on a wave function $\psi(q)$, translates it by one period,

$$T\psi(q) = \psi(q + 2\pi/\sqrt{g}) \Rightarrow [T, H] = 0. \quad (39.19)$$

Since T commutes with the Hamiltonian, both operators can be diagonalized simultaneously. The eigenvalues of T are pure phases. Each eigenfunction $\psi_N(\varphi, g, q)$ of H , which denote by $|N, \varphi, g\rangle$, is then characterized by a phase $e^{i\varphi}$, eigenvalue of T :

$$H|N, \varphi, g\rangle = \mathcal{E}_N(\varphi, g)|N, \varphi, g\rangle, \quad T|N, \varphi, g\rangle = e^{i\varphi}|N, \varphi, g\rangle. \quad (39.20)$$

For $g = 0$, H has the spectrum of the harmonic oscillator, $\mathcal{E}_N(\varphi, 0) = (N + \frac{1}{2})$. Moreover, to all orders in an expansion in powers of g , $\mathcal{E}_N(\varphi, g)$ is independent of φ : $\mathcal{E}_N(\varphi, g) \equiv \mathcal{E}_N(g)$. However, beyond the perturbative expansion, due to barrier penetration, for $g \neq 0$, $\mathcal{E}_N(\varphi, g)$ depends on φ and, for g small enough, to each value of N is associated a band.

Globally, the spectrum of H is periodic in φ . Moreover, because, for g small, the bands for different values of N do not overlap, in a band the energy eigenvalue itself is a periodic function of φ , which can be expanded in a Fourier series:

$$\mathcal{E}_N(\varphi, g) = \sum_{l=-\infty}^{+\infty} \mathcal{E}_N(l, g) e^{il\varphi}, \quad \mathcal{E}_N(l, g) = \mathcal{E}_N(-l, g). \quad (39.21)$$

All coefficients $\mathcal{E}_N(l, g)$, except $\mathcal{E}_N(0, g)$, vanish to all orders in an expansion in g .

We now consider the partition function, which is here $\mathrm{tr}' e^{-\beta H}$. The notation tr' has the following meaning: since the diagonal matrix elements of $e^{-\beta H}$ in configuration space are periodic functions, we only integrate over one period.

For g small, the large β limit selects the lowest band, and we obtain (see equation (A39.8)):

$$\mathrm{tr}' e^{-\beta H} \underset{\beta \rightarrow \infty}{\sim} \frac{1}{2\pi} \int_0^{2\pi} d\varphi e^{-\beta \mathcal{E}_0(\varphi, g)}. \quad (39.22)$$

Like in the case of the double-well potential, we note that it is difficult to determine the dependence on φ of the energy levels by calculating the partition function. By contrast, we can consider (see equation (A39.8)),

$$\text{tr}' T e^{-\beta H} \underset{\beta \rightarrow \infty}{\sim} \frac{1}{2\pi} \int_0^{2\pi} d\varphi e^{i\varphi} e^{-\beta \mathcal{E}_0(\varphi, g)}. \quad (39.23)$$

To simplify the notation, we focus on the lowest band $N = 0$, and define

$$\mathcal{E}_0(\varphi, g) \equiv E(\varphi, g), \quad \mathcal{E}_0(l, g) \equiv E_l(g).$$

For g small, $E(\varphi, g) - E_0(g)$ vanishes faster than any power of g . Therefore, we can expand equation (39.23), for $g \rightarrow 0$ and then $\beta \rightarrow \infty$, as

$$\text{tr}' T e^{-\beta H} \sim e^{-\beta E_0(g)} \int \frac{d\varphi}{2\pi} e^{i\varphi} [1 - \beta(E(\varphi, g) - E_0(g)) + \dots]. \quad (39.24)$$

The integration over φ selects $E_1(g)$. Therefore,

$$\text{tr}' T e^{-\beta H} \sim -\beta e^{-\beta E_0(g)} E_1(g), \quad g \rightarrow 0, \quad \beta \rightarrow \infty. \quad (39.25)$$

This equation can be more conveniently rewritten as

$$\text{tr}' T e^{-\beta H} / \text{tr}' e^{-\beta H} \sim -\beta E_1(g). \quad (39.26)$$

As explained previously, if E_1 does not vanish this implies that the translation symmetry is not spontaneously broken.

Remark. To evaluate the other Fourier series coefficients E_2, E_3, \dots , for g small, the most convenient method is to consider $\text{tr}' T^k e^{-\beta H}$ for $k = 2, 3, \dots$ (see Chapter 42).

39.2.2 The instanton contributions

The path integral representations of the partition function $\text{tr}' e^{-\beta H}$ and of $\text{tr}' T e^{-\beta H}$ again only differ by the boundary conditions. The operator T has the effect of translating the argument q in the matrix element $\langle q' | \text{tr}' e^{-\beta H} | q \rangle$ before taking the trace:

$$\text{tr}' T e^{-\beta H} = \int_{q(\beta/2)=q(-\beta/2)+2\pi/\sqrt{g}} [dq(t)] \exp[-\mathcal{S}(q)], \quad (39.27)$$

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} [\frac{1}{2}\dot{q}^2(t) + V(q(t)\sqrt{g})/g] dt. \quad (39.28)$$

We recall that $q(-\beta/2)$ only varies over one period of the potential. For β large and g small, due to the boundary conditions, the path integral is dominated by instanton configurations which connect two consecutive minima of the potential. The explicit solution of the equation of motion is

$$q_c(t) = \frac{4}{\sqrt{g}} \tan^{-1} e^{(t-t_0)}, \quad (39.29)$$

and the corresponding classical action, in the infinite β limit, is

$$\mathcal{S}(q_c) = 8/g. \quad (39.30)$$

For all potentials for which the minima can be exchanged by a reflection, the analogue of expression (37.92) is

$$E(\beta) \sim -e^{-\beta} \frac{x_0^2}{2} \exp \left[2 \int_0^{x_0/2} \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} \right) dx \right], \quad (39.31)$$

in which x_0 is the location of the other minimum. Applying equation (39.31) to the analogue of equation (37.83), one obtains

$$-\beta e^{-\beta/2} E_1(g) \underset{g \rightarrow 0}{\sim} \frac{4\beta e^{-\beta/2}}{\sqrt{\pi g}} e^{-8/g}, \quad (39.32)$$

or

$$E_1(g) \underset{g \rightarrow 0}{\sim} -\frac{4}{\sqrt{\pi g}} e^{-8/g}. \quad (39.33)$$

Without evaluating E_n for $n \geq 2$ explicitly, one verifies that the corresponding boundary conditions for $\text{tr } T^n e^{-\beta H}$, which are

$$q(\beta/2) = q(-\beta/2) + n \frac{2\pi}{\sqrt{g}},$$

then select an instanton solution which for β large has an action $8n/g$. Therefore, E_1 gives the dominant non-perturbative contribution for g small and

$$E(\varphi, g) = E_0(g) - \frac{8}{\sqrt{\pi g}} e^{-8/g} [1 + O(g)] \cos \varphi + O\left(e^{-16/g}\right). \quad (39.34)$$

Discussion. The two examples have illustrated that, as anticipated, in a theory in which, at the classical level, a discrete symmetry is spontaneously broken, because the classical potential has degenerate minima, the existence of instantons implies that quantum fluctuations restore the symmetry. However, we have also shown that, by contrast, spontaneous symmetry breaking of discrete symmetries is possible in higher dimensions. Analogous conclusions have been reached on the lattice in Sections 14.1 and 14.2.

Note that, in contrast to discrete symmetries, where quantum fluctuations lead to exponentially small effects in $1/\hbar$, or the equivalent coupling constant, in the case of continuous symmetries, the effects of quantum fluctuations show up already at first order in perturbation theory as a consequence of the Goldstone phenomenon (see Section 19.3).

While in theories in which the dynamical variables live in flat Euclidean space, instantons are associated with a degeneracy of the classical minimum of the potential, this is no longer necessarily the case when the space has curvature or is topologically non-trivial. An example is provided by the cosine potential with compactified space, the coordinate q representing a point on a circle of radius $2\pi/\sqrt{g}$. The Hamiltonian then corresponds to a $O(2)$ rotator in a potential (Section 3.4), or a one-dimensional classical spin chain in a magnetic field. The classical minimum is no longer degenerate, because all minima correspond to one point on the circle. The quantum ground state is equally unique, since the Hilbert space consists in strictly periodic eigenfunctions ($\varphi = 0$). The same instanton solutions still exist, which now start from and return to the same classical minimum, winding around the circle. They are stable because the circle is topologically non-trivial. They generate the same exponentially small corrections, which we have described previously.

Further insight into the problem can be gained by generalizing the Hamiltonian to the $O(N)$ symmetric rotator of Section 3.5 in a potential $1 - q_1$. The classical solutions are the same, but the degeneracy and the stability properties are different. For $N > 2$, the solutions which wind around the sphere have $(N - 2)$ directions of instability. Their contributions have to be discussed in the context of the large-order behaviour of perturbation theory (see Chapter 40).

39.3 Instantons and stochastic dynamics

We now describe the role of instantons in the context of stochastic dynamics. We consider the problem of evaluating the decay probability of a metastable state by thermal fluctuations. At first, one could think that this topic should be discussed in Chapter 38, simultaneously with the problem of decay by quantum fluctuations. We show here that, technically, the problem has a more direct relation with degenerate classical minima. At the end of the section, we also briefly examine the role of instantons when the equilibrium distribution has degenerate minima.

39.3.1 Random walk

We consider the Langevin equation (34.34) (see Sections 34.4–34.6),

$$\dot{\mathbf{q}}(t) = -\frac{1}{2}\Omega \nabla E(\mathbf{q}(t)) + \boldsymbol{\nu}(t), \quad (39.35)$$

with the Gaussian white noise distribution (34.3) defined by (Ω is a positive constant)

$$\langle \nu_i(t) \rangle_\nu = 0, \quad \langle \nu_i(t) \nu_j(t') \rangle_\nu = \Omega \delta_{ij} \delta(t - t'). \quad (39.36)$$

To the Langevin equation is associated the *Hermitian* Hamiltonian (34.39),

$$H = \frac{1}{2}\Omega \left[-\nabla^2 + \frac{1}{4} (\nabla E(\mathbf{q}))^2 - \frac{1}{2}\nabla^2 E(\mathbf{q}) \right]. \quad (39.37)$$

Observables can be calculated with a path integral with the functional measure $e^{-S(\mathbf{q})}[dq]$, where the corresponding dynamic action is given by (equation (34.61), with $S/\Omega \mapsto S$),

$$S(\mathbf{q}) = \frac{1}{2}\Omega^{-1} \int \left\{ \dot{\mathbf{q}}^2(t) + \frac{1}{4}\Omega^2 [\nabla E(\mathbf{q}(t))]^2 - \frac{1}{2}\Omega^2 \nabla^2 E(\mathbf{q}(t)) \right\} dt. \quad (39.38)$$

The classical limit here is replaced by the small Ω and thus low temperature limit. At leading order in a semi-classical analysis, the term $\Omega^2 \nabla^2 E$ can be omitted. Therefore, the classical minima of the action (39.38) correspond to all points where $\nabla E(\mathbf{q})$ vanishes, thus all critical points (extrema or saddle points) of the function $E(\mathbf{q})$. If more than one critical point can be found, the classical equations may have instanton solutions.

Examples. We consider an analytic function $E(\mathbf{q})$ which has only a relative minimum at $\mathbf{q} = 0$:

$$E(\mathbf{q}) = \frac{1}{2}\omega^2 \mathbf{q}^2 + O(|\mathbf{q}|^3).$$

Then the function $E(\mathbf{q})$ must also have elsewhere a saddle point or a relative maximum. Physically, we then know that if we put a particle at time 0 at the relative minimum $\mathbf{q} = 0$, then after some time the particle will escape from the well, as a result of thermal fluctuations, as described by the Langevin equation. The problem is to evaluate, in the small Ω limit, the escape probability per unit time, or the average escape time τ .

A class of examples corresponds to functions such that the distribution $e^{-E(\mathbf{q})}$ is not normalizable, like in one dimension

$$E(q) = q^2 - 2q^3/3. \quad (39.39)$$

We then know that τ is the inverse of the smallest eigenvalue of the Hamiltonian (39.37) (see Section 34.3.1), and this eigenvalue is strictly positive.

However, in all examples to all orders in a perturbative expansion in powers of Ω starting from the saddle point $\mathbf{q} = 0$, the function $e^{-E(\mathbf{q})/2}$ is the formal ground state eigenvector associated with the eigenvalue 0. It follows that the calculation of the eigenvalue is not perturbative. We now show that the instantons connecting the critical points of the function $E(\mathbf{q})$ provide a solution to the problem [395].

Instantons. An instanton connects the minimum at $\mathbf{q} = 0$ to another critical point \mathbf{q}_0 where ∇E vanishes. For an instanton solution \mathbf{q}_c the following inequality holds,

$$\int_{-\infty}^{+\infty} dt [\dot{\mathbf{q}}_c(t) \pm \frac{1}{2}\Omega \nabla E(\mathbf{q}_c(t))]^2 \geq 0 \quad (39.40)$$

and, therefore,

$$\mathcal{S}(\mathbf{q}_c) \geq |Q(\mathbf{q}_c)|, \quad (39.41)$$

with

$$Q(\mathbf{q}_c) = \frac{1}{2} \int_{-\infty}^{+\infty} dt \dot{\mathbf{q}}_c(t) \cdot \nabla E(\mathbf{q}_c(t)) = \frac{1}{2}(E(\mathbf{q}_0) - E(0)). \quad (39.42)$$

We conclude that the action satisfies [396]

$$\mathcal{S}(\mathbf{q}_c) \geq \frac{1}{2}|E(\mathbf{q}_0) - E(0)|. \quad (39.43)$$

The equality corresponds to a local minimum of the action and \mathbf{q}_c then is a solution of a first-order differential equation,

$$\dot{\mathbf{q}}(t) = \pm \frac{1}{2}\Omega \nabla E(\mathbf{q}(t)). \quad (39.44)$$

However, this is not the end of the story. Indeed, in Section 34.5, we have shown that the degeneracy between the minima and maxima of the function $E(\mathbf{q})$ is lifted by the first quantum correction. Therefore, the two minima of the action are not really degenerate, and no instanton can connect them. What really happens is that we have to only consider closed trajectories passing through the origin. If we consider a finite time interval β , we can find such trajectories. In the infinite β limit, they decompose into a succession of instantons and anti-instantons. The limit of the classical action is an even multiple of the instanton action. The leading contribution thus is (Arrhenius law)

$$\mathcal{S}(\mathbf{q}_c) = E(\mathbf{q}_0) - E(0). \quad (39.45)$$

We conclude, quite generally, that if the function $E(\mathbf{q})$ has a relative minimum where $E = E_{\min}$, separated from a lower minimum (possibly $E = -\infty$) by a local maximum $E = E_{\max}$, then the eigenvalue corresponding to an eigenfunction concentrated around the first minimum is of the order $e^{-\Delta E}$, in which ΔE is the variation of the function E :

$$\Delta E = E_{\max} - E_{\min}.$$

The time $\tau = O(e^{\Delta E})$ characterizes the exponential decay of the probability of finding $\mathbf{q}(t)$ near the origin when the initial conditions at $t = 0$ are $\mathbf{q}(t = 0) = 0$.

Finally, in order to complete the calculation of the eigenvalue, it is necessary to use multi-instanton techniques of the kind explained in Chapter 42.

Note that the inequality (39.40), which we have used, corresponds to a structure that is typical for *supersymmetric* theories (*cf.*, the action (34.74)).

Degenerate minima. A new problem arises when the function $E(\mathbf{q})$ has a degenerate minimum. Let us assume that the corresponding distribution is normalizable. Then the ground state eigenvalue vanishes. The interesting question is how to calculate the difference between the two first eigenvalues, difference which vanishes to all orders in perturbation theory. This is the problem we have solved in Section 39.1 for the double-well potential. However, here the set-up is slightly different because, if $E(\mathbf{q})$ is regular, as we always assume, the two minima are necessarily separated by a maximum or a saddle point and, therefore, $(\nabla E)^2$ has at least three minima. A one-dimensional example is

$$E(q) = q^2(1 - q)^2 \quad (39.46)$$

and, therefore,

$$E'^2(q) = 4(1 - 2q)^2q^2(1 - q)^2. \quad (39.47)$$

This time we look for instanton solutions connecting $q = 0$ to $q = 1$. However, in the infinite time limit, only instantons which go from 0 to $1/2$ or $1/2$ to 1 survive. From the analysis of the previous problem, we guess that the relevant configurations correspond to gluing together two instantons. Therefore, the difference between the two leading eigenvalues, which is also the second eigenvalue ϵ_1 , is again of the form

$$\ln \epsilon_1 = -\ln \tau \sim -(E_{\max} - E_{\min}). \quad (39.48)$$

in which E_{\min} and E_{\max} are respectively the values of the function $E(q)$ at the degenerate minima and at the maximum which connects them.

39.3.2 Quantum field theory (QFT)

We consider a dynamics governed by a purely dissipative Langevin equation (equation (35.35)), which formally converges towards an equilibrium distribution corresponding to the field integral of a d -dimensional Euclidean QFT. Moreover, the field theory, with the Euclidean action

$$\mathcal{A}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{6} g \phi^3(x) \right], \quad (39.49)$$

has a metastable minimum at $\phi(x) = 0$. We know that a quantum state concentrated around the minimum will decay due to quantum fluctuations, and we have calculated the rate by instanton methods.

We now want to evaluate the decay probability due to thermal fluctuations. The relevant dynamic action, which is now $(d + 1)$ dimensional, at leading order, reads:

$$\mathcal{S}(\phi) = \frac{1}{2} \int d^d x dt \left\{ \frac{\dot{\phi}^2(t, x)}{\Omega} + \frac{1}{4} \Omega \left[-\nabla_x^2 \phi(t, x) + m^2 \phi(t, x) + \frac{1}{2} g \phi^2(t, x) \right]^2 \right\}. \quad (39.50)$$

Formally, the discussion follows the same line as in the case of a finite number of degrees of freedom. The problem is to identify the minimum and the maximum of the action.

The minimum is easy to find: $\phi \equiv 0$. The maximum requires some more thought. It does not correspond to a constant field configuration: $\phi(x) = -2m^2/g$. Indeed, it is sufficient that some part of the field starts passing the barrier. Instead, the relevant maximum of the action corresponds to a static instanton configuration. The arguments of Section 39.2.2, then lead to the estimate $\ln \tau \sim \exp(\mathcal{S}_{\text{inst.}} - \mathcal{S}(\phi = 0))$, where $\mathcal{S}_{\text{inst.}}$ is the instanton action.

39.4 Instantons in stable boson field theories: General remarks

We now briefly study the existence of instantons in stable theories, connecting, for example, degenerate classical minima. The most interesting examples correspond, unfortunately, to scale invariant classical theories. The evaluation of the instanton contributions at leading order, which formally follows the method described in Chapter 38, leads to difficulties due both to ultraviolet (UV) and infrared (IR) divergences. Some of them are examined in Chapter 38. Since, for the two examples we consider in Sections 39.5 and 39.6, they have not been satisfactorily solved, we restrict ourselves here to semi-classical considerations.

We begin with a few general remarks about the possible existence of instantons in stable field theories.

Scalar field theories. First, we assume that the action for a multicomponent scalar boson field ϕ^i has the form

$$\mathcal{S}(\phi) = \int [K(\phi(x)) + V(\phi(x))] d^d x, \quad (39.51)$$

with

$$K(\phi(x)) = \frac{1}{2} \sum_{i,j} g_{ij}(\phi(x)) \nabla \phi^i(x) \nabla \phi^j(x),$$

in which $g_{ij}(\phi)$ a positive matrix (positive definite almost everywhere) and

$$\min_{\{\phi\}} V(\phi) = 0. \quad (39.52)$$

Equation (A38.3) immediately generalizes to

$$(2-d) \int K(\phi(x)) d^d x = d \int V(\phi(x)) d^d x.$$

This equation has no solution for $d > 2$. For $d = 2$, it has solutions only if

$$V(\phi_c(x)) = 0. \quad (39.53)$$

The condition (39.52) then implies that $\phi_c(x)$ is for all x a minimum of the potential,

$$\frac{\partial V(\phi_c)}{\partial \phi} = 0,$$

and, therefore, $\phi_c(x)$ is a solution of the field equations

$$\frac{\delta}{\delta \phi^i(y)} \int K(\phi(x)) d^2 x = 0.$$

These two equations are in general incompatible, except if $V(\phi)$ vanishes identically. In the latter case, the action (39.51) corresponds to a two-dimensional model on a Riemannian manifold. A particular class of such models based on homogeneous spaces has been discussed in Chapters 19 and 29. Among them, the $CP(N-1)$ models are known to admit instanton solutions, which are described in Section 39.5.

Gauge theories. If, in addition to scalar fields, the theory contains gauge fields A_μ^a (see Chapter 22), the gauge invariant action has the form (D_μ is the covariant derivative)

$$\mathcal{S}(\phi, \mathbf{A}) = \mathcal{S}(\mathbf{A}) + \Sigma(\phi, \mathbf{A}) + \int d^d x V(\phi(x)), \quad (39.54)$$

with $V(\phi) \geq 0$, and

$$\mathcal{S}(\mathbf{A}) = \sum_{a,\mu,\nu} \frac{1}{4g} \int d^d x F_{\mu\nu}^a(x) F_{\mu\nu}^a(x), \text{ with } g > 0, \quad \Sigma(\phi, \mathbf{A}) = \frac{1}{2} \int d^d x \sum_{\mu,i} (D_\mu \phi_i(x))^2.$$

Assuming the existence of a finite action solution $\{\phi^c, \mathbf{A}_\mu^c\}$ (in which \mathbf{A}_μ^c is not a pure gauge), one calculates the action for $\lambda \mathbf{A}_\mu^c(\lambda x)$ and $\phi^c(\lambda x)$. After the change of variables $\lambda x \mapsto x$, one obtains

$$\mathcal{S}(\phi^c, \mathbf{A}^c; \lambda) = \lambda^{4-d} \mathcal{S}(\mathbf{A}^c) + \lambda^{2-d} \Sigma(\phi^c, \mathbf{A}^c) + \lambda^{-d} \int V(\phi^c(x)) d^d x. \quad (39.55)$$

Stationarity at $\lambda = 1$ implies

$$(4-d)\mathcal{S}(\mathbf{A}^c) + (2-d)\Sigma(\phi^c, \mathbf{A}^c) - d \int V(\phi^c(x)) d^d x = 0. \quad (39.56)$$

We see that no solution can exist for $d > 4$, since a sum of negative terms cannot vanish.

For $d = 4$, we find two conditions:

$$V(\phi^c(x)) = 0, \quad (39.57a)$$

$$D_\mu \phi^c(x) = 0. \quad (39.57b)$$

Applying these conditions to the field equations, we conclude that \mathbf{A}_μ^c is the solution of the pure gauge field equations. As we show in Section 39.6, instantons can indeed be found in pure gauge theories. Equation (39.57b), which now is an equation for ϕ^c , then leads to the integrability conditions:

$$[D_\mu, D_\nu] = F_{\mu\nu} \Rightarrow \sum_{a,j} (F_{\mu\nu}^a(x))^c t_{ij}^a \phi_j^c(x) = 0, \quad (39.58)$$

in which the matrices t^a are the generators of the Lie algebra. The conditions (39.58) together with the equation (39.57a) show that, in general, the system has only the trivial solution $\phi^c(x) = 0$.

39.5 Instantons in $CP(N-1)$ models

The preceding considerations can be illustrated by the two-dimensional $CP(N-1)$ models. We mainly describe the nature of the instanton solutions and refer the reader to the literature for a detailed discussion [397].

We consider a set of N complex fields φ_α , subject to the condition

$$\bar{\varphi}(x) \cdot \varphi(x) = 1. \quad (39.59)$$

Moreover, two vectors φ and φ' are equivalent if they are related by the $U(1)$ gauge transformation,

$$\varphi'(x) = e^{i\Lambda(x)} \varphi(x), \quad (39.60)$$

where $\Lambda(x)$ is an arbitrary real function. These conditions characterize the manifold $CP(N-1)$ (for $(N-1)$ -dimensional Complex Projective), which is isomorphic to the complex Grassmannian manifold $U(N)/U(1)/U(N-1)$, one of the symmetric spaces exhibited in Section A29.4.3. One form of the unique classical action is

$$\mathcal{S}(\varphi, A) = \frac{1}{g} \sum_{\mu} \int d^2x \overline{D_\mu \varphi}(x) \cdot D_\mu \varphi(x), \quad g > 0, \quad (39.61)$$

in which D_μ is the covariant derivative:

$$D_\mu = \partial_\mu + iA_\mu. \quad (39.62)$$

The gauge field A_μ implements the invariance of the action under the $U(1)$ gauge transformations (39.60). Since the action contains no kinetic term for the gauge field A_μ , A_μ is an auxiliary field that can be integrated out. The integral is Gaussian, and the result is obtained by replacing in the action A_μ by the solution of the A_μ -field equation, $\delta\mathcal{S}/\delta A_\mu(x) = 0$. Using equation (39.59), one finds

$$A_\mu(x) = i\bar{\varphi}(x) \cdot \partial_\mu \varphi(x). \quad (39.63)$$

After this substitution $\bar{\varphi}(x) \cdot \partial_\mu \varphi(x)$ plays the role of a composite gauge field.

Instantons. A proof of the existence of locally stable non-trivial minima of the action follows from the inequality (note the analogy with equation (39.40))

$$\sum_{\mu} \int d^2x \left| D_\mu \varphi(x) \mp i \sum_{\nu} \epsilon_{\mu\nu} D_\nu \varphi(x) \right|^2 \geq 0, \quad (39.64)$$

($\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$ and $\epsilon_{12} = 1$). Expanding the expression, one obtains

$$\mathcal{S}(\varphi) \geq |Q(\varphi)|/g, \quad (39.65)$$

with

$$Q(\varphi) = -i \sum_{\mu, \nu} \epsilon_{\mu\nu} \int d^2x D_\mu \varphi(x) \cdot \overline{D_\nu \varphi}(x) = i \sum_{\mu, \nu} \int d^2x \epsilon_{\mu\nu} D_\nu D_\mu \varphi(x) \cdot \bar{\varphi}(x), \quad (39.66)$$

after an integration by parts. Then, in the representation (39.61)

$$i \sum_{\mu, \nu} \epsilon_{\mu\nu} D_\nu D_\mu = \frac{1}{2} i \sum_{\mu, \nu} \epsilon_{\mu\nu} [D_\nu, D_\mu] = -\frac{1}{2} \sum_{\mu, \nu} \epsilon_{\mu\nu} F_{\mu\nu}, \quad (39.67)$$

where

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$$

is the curvature.

Therefore, using equation (39.59), one finds

$$Q(\varphi) = -\frac{1}{2} \sum_{\mu,\nu} \int d^2x \epsilon_{\mu\nu} F_{\mu\nu}(x). \quad (39.68)$$

The integrand is proportional to the two-dimensional Abelian chiral anomaly (see the general expression (23.92), or equation (A30.3)), and thus is a total divergence:

$$\frac{1}{2} \sum_{\mu,\nu} \epsilon_{\mu\nu} F_{\mu\nu}(x) = \sum_{\mu,\nu} \partial_\mu \epsilon_{\mu\nu} A_\nu(x).$$

Substituting this form into equation (39.68), and integrating in a large disk of radius R , one infers

$$Q(\varphi) = - \lim_{R \rightarrow \infty} \oint_{|x|=R} d\mathbf{x} \cdot \mathbf{A}(x).$$

$Q(\varphi)$ thus only depends on the behaviour of the classical solution for $|x|$ large and is a *topological charge*. Finiteness of the action demands that at large distances $D_\mu \varphi$ vanishes. Equation (39.67) then implies that $F_{\mu\nu}$ vanishes, and thus A_μ is a pure gauge (and φ a gauge transform of a constant vector),

$$A_\mu(x) = \partial_\mu \Lambda(x) \Rightarrow Q(\varphi) = - \lim_{R \rightarrow \infty} \oint_{|x|=R} d\mathbf{x} \cdot \nabla \Lambda(x). \quad (39.69)$$

The topological charge measures the variation of the angle $\Lambda(x)$ on a large circle, which necessarily is a multiple of 2π because φ is regular. One is thus led to the consideration of the homotopy classes of mappings from $U(1)$, that is, S_1 (the circle) to S_1 , which are characterized by an integer n , the *winding number*, and

$$Q(\varphi) = 2\pi n \Rightarrow S(\varphi) \geq 2\pi|n|/g. \quad (39.70)$$

The equality $S(\varphi) = 2\pi|n|/g$ corresponds to a local minimum, and implies that the classical solutions satisfy the first-order partial differential (self-duality) equations

$$D_\mu \varphi(x) = \pm i \sum_\nu \epsilon_{\mu\nu} D_\nu \varphi(x). \quad (39.71)$$

It can be shown that, in the variable $z = x_1 + ix_2$, the solutions of the equations (39.71) are proportional to holomorphic or anti-holomorphic (depending on the sign) vectors (this reflects the conformal invariance of the classical field theory). Using then equation (39.59) and a gauge transformation (39.60), one can cast the holomorphic solution into the form

$$\varphi_\alpha(z) = P_\alpha(z) / [P(z) \cdot \bar{P}(z)]^{1/2}, \quad (39.72)$$

where the $P_\alpha(z)$ are polynomials in z without common roots. The anti-holomorphic solution corresponds to interchanging φ and $\bar{\varphi}$.

The semi-classical vacuum. In contrast to the method used in Sections 39.1 and 39.2, the existence of instantons has been discussed here without reference to the structure of the classical vacuum. To find an interpretation of instantons in gauge theories, it is convenient to express the results in the temporal gauge.

39.5.1 The semi-classical vacuum: Temporal gauge

In the temporal gauge, classical minima of the potential correspond to fields $\varphi(x_1)$, where x_1 is only the space variable, gauge transforms of a constant vector:

$$\varphi(x_1) = e^{i\Lambda(x_1)} \mathbf{v}, \quad \bar{\mathbf{v}} \cdot \mathbf{v} = 1.$$

If the vacuum state is invariant under space reflection, then $\varphi(+\infty) = \varphi(-\infty)$, and thus

$$\Lambda(+\infty) - \Lambda(-\infty) = 2n\pi \quad n \in \mathbb{Z}.$$

The integer n is a topological number that classifies the degenerate classical minima, and the semi-classical vacuum thus has a periodic structure. This analysis is consistent with Gauss's law (Section 22.3), which only implies that states are invariant under infinitesimal gauge transformations and, therefore, under gauge transformations of the class $n = 0$, which are continuously connected to the identity.

We now consider a large rectangle with an extension R in the space direction and T in the Euclidean time direction, and by a smooth gauge transformation continue the instanton solution to the temporal gauge. Then, the variation of the pure gauge comes entirely from the sides with $x_2 = 0$ and $x_2 = T$. One finds for $R \rightarrow \infty$,

$$\Lambda(+\infty, 0) - \Lambda(-\infty, 0) - [\Lambda(+\infty, T) - \Lambda(-\infty, T)] = 2n\pi.$$

Therefore, instantons interpolate between different classical minima. One can project onto a proper quantum eigenstate, the θ -vacuum, corresponding to an angle θ , by adding a topological term to the classical action,

$$\mathcal{S}(\varphi) \mapsto \mathcal{S}(\varphi) + i \frac{\theta}{4\pi} \sum_{\mu, \nu} \int d^2x \epsilon_{\mu\nu} F_{\mu\nu}(x),$$

in analogy with expressions (23.40), or as in the example of the cosine potential (42.61, 42.60).

The $CP(1)$ model. The $CP(1)$ model is locally isomorphic to the $O(3)$ non-linear σ -model, with the identification

$$\phi(x) = \sum_{\alpha, \beta} \bar{\varphi}_\alpha(x) \boldsymbol{\sigma}_{\alpha\beta} \varphi_\beta(x), \tag{39.73}$$

where σ_i are the three Pauli matrices. In the $O(3)$ σ -model, the $CP(1)$ minimal instanton solution becomes the stereographic mapping of the sphere S_2 onto the plane [398]:

$$P_1(z) = z, \quad P_2(z) = 1 \Rightarrow \phi_1 = \frac{z + \bar{z}}{1 + \bar{z}z}, \quad \phi_2 = i \frac{z - \bar{z}}{1 + \bar{z}z}, \quad \phi_3 = \frac{1 - \bar{z}z}{1 + \bar{z}z}.$$

39.6 Instantons in the $SU(2)$ gauge theory

In four dimensions, non-Abelian gauge theories provide an example of instantons related to the vacuum structure of quantum chromodynamics (QCD, see Chapter 22) [399]. According to the analysis of Section 39.4, we can consider only pure gauge theories. Actually, it is sufficient to consider the gauge group $SU(2)$, since a general theorem states that for a Lie group containing $SU(2)$ as a subgroup the instantons are those of the $SU(2)$ subgroup.

In $SO(3)$ notation, the gauge field \mathbf{A}_μ is a vector and the gauge action reads

$$\mathcal{S}(\mathbf{A}) = \frac{1}{4g} \sum_{\mu,\nu} \int [\mathbf{F}_{\mu\nu}(x)]^2 d^4x, \quad (39.74)$$

with

$$\mathbf{F}_{\mu\nu}(x) = \partial_\mu \mathbf{A}_\nu(x) - \partial_\nu \mathbf{A}_\mu(x) + \mathbf{A}_\mu(x) \times \mathbf{A}_\nu(x). \quad (39.75)$$

Generalizing the arguments used for the $CP(N-1)$ model, one derives the existence, and some properties, of instantons in this theory.

The dual of the tensor $\mathbf{F}_{\mu\nu}$ is defined by

$$\tilde{\mathbf{F}}_{\mu\nu}(x) = \frac{1}{2} \sum_{\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} \mathbf{F}_{\rho\sigma}(x). \quad (39.76)$$

Then the inequality [396]

$$\sum_{\mu,\nu} \int d^4x \left[\mathbf{F}_{\mu\nu}(x) \pm \tilde{\mathbf{F}}_{\mu\nu}(x) \right]^2 \geq 0, \quad (39.77)$$

implies,

$$\mathcal{S}(\mathbf{A}_\mu) \geq |Q(\mathbf{A}_\mu)|/4g, \quad (39.78)$$

where $Q(\mathbf{A})$ is an expression one also meets in Section 23.6.3 (equation (23.95), here written in $SO(3)$ notation) in the calculation of the axial anomaly

$$Q(\mathbf{A}) = \int d^4x \sum_{\mu,\nu} \mathbf{F}_{\mu\nu}(x) \cdot \tilde{\mathbf{F}}_{\mu\nu}(x). \quad (39.79)$$

There it is shown that the quantity $\sum_{\mu,\nu} \mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu}$ is a pure divergence (equation (23.96)). Indeed, one verifies that

$$\sum_{\mu,\nu} \mathbf{F}_{\mu\nu}(x) \cdot \tilde{\mathbf{F}}_{\mu\nu}(x) = \nabla \cdot \mathbf{V}(x), \quad (39.80)$$

with

$$V_\mu(x) = 2 \sum_{\nu,\rho,\sigma} \epsilon_{\mu\nu\rho\sigma} [\mathbf{A}_\nu(x) \cdot \partial_\rho \mathbf{A}_\sigma(x) + \frac{1}{3} \mathbf{A}_\nu(x) \cdot (\mathbf{A}_\rho(x) \times \mathbf{A}_\sigma(x))]. \quad (39.81)$$

The integral thus only depends on the behaviour of the gauge field at large distances, and its values are quantized (equation (23.102)). Here again, as in the example of the $CP(N-1)$ model, the bound involves a topological charge, $Q(\mathbf{A}_\mu)$.

The finiteness of the action implies that the classical solution must asymptotically become a pure gauge, that is, with our conventions,

$$-\frac{1}{2} i \mathbf{A}_\mu(x) \cdot \boldsymbol{\sigma} = \mathbf{g}(x) \partial_\mu \mathbf{g}^{-1}(x) + O(|x|^{-2}) \quad |x| \rightarrow \infty, \quad (39.82)$$

in which $\boldsymbol{\sigma}$ are Pauli matrices and $\mathbf{g}(x)$ is an element of $SU(2)$.

Since $SU(2)$ is topologically equivalent to S_3 , we are now led to consider the homotopy classes of mappings from S_3 to S_3 , which are also classified by an integer, the *winding number*. The one to one mapping corresponds to an element of the form

$$\mathbf{g}(x) = \frac{x_4 + i\mathbf{x} \cdot \boldsymbol{\sigma}}{r}, \quad r = (x_4^2 + \mathbf{x}^2)^{1/2}, \quad (39.83)$$

and, thus,

$$A_m^i(x) \underset{r \rightarrow \infty}{\sim} 2 \left(x_4 \delta_{im} + \sum_k \epsilon_{imk} x_k \right) \frac{1}{r^2}, \quad \text{for } m \leq 3, \quad A_4^i(x) = -2 \frac{x_i}{r^2}. \quad (39.84)$$

It follows that

$$\int d^4x \sum_{\mu, \nu} \mathbf{F}_{\mu\nu}(x) \cdot \tilde{\mathbf{F}}_{\mu\nu}(x) = \int d\Omega \hat{\mathbf{n}} \cdot \mathbf{V}(x) = 32\pi^2, \quad (39.85)$$

where $d\Omega$ is the invariant measure on the sphere, and $\hat{\mathbf{n}}$ the unit vector normal to surface of the sphere.

Comparing this result with equation (23.102), we note that we have indeed found the minimal action solution. In general, we then expect

$$\int d^4x \mathbf{F}_{\mu\nu}(x) \cdot \tilde{\mathbf{F}}_{\mu\nu}(x) = 32\pi^2 n, \quad (39.86)$$

and therefore,

$$S(\mathbf{A}_\mu) \geq 8\pi^2 |n|/g. \quad (39.87)$$

The equality, which corresponds to a local minimum of the action, is obtained for fields satisfying the self-duality equations

$$\mathbf{F}_{\mu\nu}(x) = \pm \tilde{\mathbf{F}}_{\mu\nu}(x), \quad (39.88)$$

which are first-order partial differential equations. The one-instanton solution, which depends on an arbitrary scale parameter λ , is ($r = |\mathbf{x}|$)

$$A_m^i(x) = \frac{2}{r^2 + \lambda^2} \left(x_4 \delta_{im} + \sum_k \epsilon_{imk} x_k \right), \quad A_4^i(x) = -\frac{2x_i}{r^2 + \lambda^2}. \quad (39.89)$$

The semi-classical vacuum. In analogy with the analysis of the $CP(N-1)$ model, we quantize in the temporal gauge $\mathbf{A}_4 = 0$. The classical minima of the potential correspond to gauge field components \mathbf{A}_i , $i = 1, 2, 3$, which are pure gauge functions of the three space variables x_i :

$$-\frac{1}{2} i \mathbf{A}_m(x) \cdot \boldsymbol{\sigma} = \mathbf{g}(x_i) \partial_m \mathbf{g}^{-1}(x_i). \quad (39.90)$$

The structure of the classical minima is related to the homotopy classes of mappings of the group elements \mathbf{g} into compactified \mathbb{R}^3 (because $\mathbf{g}(x)$ goes to a constant for $|x| \rightarrow \infty$), that is, again of S_3 into S_3 , and thus the semi-classical vacuum has a periodic structure. One verifies that the gauge equivalent in the temporal gauge of the instanton solution (39.89) connects minima with different winding numbers.

Therefore, as in the example of the $CP(N-1)$ model, to project onto a θ vacuum, one can add the topological term to the classical action of gauge theories,

$$\mathcal{S}_\theta(\mathbf{A}) = \mathcal{S}(\mathbf{A}) + i \frac{\theta}{32\pi^2} \int d^4x \sum_{\mu,\nu} \mathbf{F}_{\mu\nu}(x) \cdot \tilde{\mathbf{F}}_{\mu\nu}(x), \quad (39.91)$$

and then integrate over all fields \mathbf{A}_μ without restriction. At least in the semi-classical approximation, the gauge theory thus depends on one additional parameter, the angle θ . For non-vanishing values of θ , the additional term violates CP (charge conjugation parity) conservation, and is at the origin of the *strong CP violation* problem: experimental bounds fix the value of θ with an unnatural precision of the order of 10^{-9} .

39.6.1 Fermions in an instanton background

In QCD (see Section 23.4.1), gauge fields are coupled to quarks $(\mathbf{Q}, \bar{\mathbf{Q}})$ with an action of the form (using here a $SU(3)$ notation)

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{Q}}, \mathbf{Q}) = - \int d^4x \left(\frac{1}{4g^2} \sum_{\mu,\nu} \text{tr} \mathbf{F}_{\mu\nu}^2(x) + \sum_{f=1}^{N_f} \bar{\mathbf{Q}}_f(x) (\mathcal{D} + m_f) \mathbf{Q}_f(x) \right),$$

where N_f is the number of quark flavours.

Then, first if the θ term in (39.91) contributes and one fermion field is massless, according to the analysis of Section 23.6.4, the Dirac operator has at least one vanishing eigenvalue, and the determinant resulting from the fermion integration vanishes. Then, the instantons do not contribute to the field integral, and the strong CP violation problem is solved. However, such a hypothesis seems to be inconsistent with experimental estimates of quark masses.

Second, as we have already discussed in Section 23.7.2, if the instantons contribute, they solve the $U(1)$ problem, that is, the absence of a Goldstone boson associated with the almost spontaneous breaking of the axial $U(1)$ current.

The Gaussian integration. In $CP(N-1)$ models, and in non-Abelian gauge theories, the classical theory is scale invariant. Therefore, instanton solutions depend on a scale parameter, which is a collective coordinate over which one has to integrate. This leads to difficult problems, as the analysis of the massless $\phi_{d=4}^4$ field theory reveals (see Chapter 38). Both theories are asymptotically free, and the problems come from the IR region, that is, from instantons of large size for which the semi-classical approximation is no longer legitimate, because the interaction increases with distance (see Chapters 24 and 25).

The role of instantons thus is not fully understood, a complete calculation being possible only with an IR cut-off, provided, for example, by a finite volume. One piece of information presently available concerns the $O(3)$ non-linear σ -model, whose instantons are derived from those of the $CP(1)$ -model. It has been rather indirectly argued, by mapping the σ -model onto a one-dimensional quantum spin chain, that instantons are only relevant for $\theta = \pi$, but then they drastically alter the physical picture.

A39 Trace formula for periodic potentials

We consider a Hamiltonian H corresponding to a real periodic analytic potential $V(x)$ with period τ :

$$V(x + \tau) = V(x). \quad (A39.1)$$

Then H commutes with the unitary translation operator T , which on wave functions acts like

$$T\psi(x) = \psi(x + \tau) \Rightarrow T^\dagger\psi(x) = \psi(x - \tau).$$

Both operators T and H can be diagonalized simultaneously (see Section 39.2). At φ fixed, the spectrum of H is discrete. We define

$$H\psi_n(\varphi, x) = E_n(\varphi)\psi_n(\varphi, x), \quad T\psi_n(\varphi, x) = e^{i\varphi}\psi_n(\varphi, x), \quad (A39.2)$$

with $\|\psi_n(\varphi, x)\| = 1$.

In an interval of size $N\tau$ with periodic boundary conditions, φ is quantized:

$$e^{iN\varphi} = 1 \Rightarrow \varphi = \varphi_p \equiv \frac{2\pi p}{N}, \quad 0 \leq p < N. \quad (A39.3)$$

We express the matrix elements of $T^k e^{-\beta H}$ ($k \in \mathbb{Z}$) in terms of eigenfunctions as

$$\langle x' | T^k e^{-\beta H} | x \rangle = \sum_{p,n} \psi_n^*(\varphi_p, x') e^{-\beta E_n(\varphi_p) + ik\varphi_p} \psi_n(\varphi_p, x). \quad (A39.4)$$

This implies for the diagonal elements,

$$\langle x | T^k e^{-\beta H} | x \rangle = \sum_{p,n} |\psi_n(\varphi_p, x)|^2 e^{-\beta E_n(\varphi_p) + ik\varphi_p}. \quad (A39.5)$$

Because in a translation of a period (equation (A39.2)), the eigenfunctions are multiplied by a phase, equation (A39.5) shows that the left-hand side is a periodic function of x with period τ . Therefore,

$$\int_0^\tau \langle x | T^k e^{-\beta H} | x \rangle dx = \sum_{n,p} e^{ik\varphi_p - \beta E_n(\varphi_p)} \frac{1}{N} \int_0^{N\tau} |\psi_n(\varphi_p, x)|^2 dx. \quad (A39.6)$$

Since the eigenfunctions $\psi_{\varphi_p, n}(x)$ are orthonormal over $N\tau$,

$$\int_0^\tau \langle x | T^k e^{-\beta H} | x \rangle dx = \frac{1}{N} \sum_{n,p} e^{ik\varphi_p - \beta E_n(\varphi_p)}. \quad (A39.7)$$

Taking the large N limit, one obtains the expression

$$\int_0^\tau \langle x | T^k e^{-\beta H} | x \rangle dx = \frac{1}{2\pi} \sum_n \int_0^{2\pi} e^{ik\varphi - \beta E_n(\varphi)} d\varphi. \quad (A39.8)$$

40 Large order behaviour of perturbation theory

In quantum field theory (QFT), the main *analytic tool* to calculate physical quantities is the perturbative expansion. Quite early, Dyson argued, with intuitive arguments [400], that perturbative expansions correspond to divergent series.

This problem started to be mathematically investigated, using the Schrödinger equation, in some models of quantum mechanics (QM) with polynomial potentials [401]. Dyson's intuition was confirmed and, moreover, the connection between barrier penetration effects in the semi-classical limit and large behaviour in perturbation theory in QM was realized. Later, it was proposed to study the problem within a path integral formulation [402]. However, a systematic study of the problem was triggered by Lipatov [384], using the field-integral representation of the $\phi_{d=4}^4$ field theory [403].

In Chapter 37, we have studied the analytic structure of the ground-state energy $E(g)$ of the quartic anharmonic oscillator. We have argued that $E(g)$ is analytic in a cut-plane. For g small, $E(g)$ can be calculated as a perturbative series in g :

$$E(g) = \sum_k E_k g^k. \quad (40.1)$$

On the cut, for g small and negative, its imaginary part $\text{Im } E(g)$ can be calculated by instanton methods. In this chapter, we show how the behaviour of $\text{Im } E(g)$ for $g \rightarrow 0_-$ is related to the behaviour of the coefficients E_k for k large [404, 405]. The method is then generalized to the class of potentials for which we have calculated instanton contributions. The same method can readily be applied to boson field theories, using the results of Chapter 38, while the extension to field theories involving fermions, like quantum electrodynamics (QED), requires additional considerations.

A general conclusion is that, in QFT, all perturbative series, expanded in terms of a loop-expansion parameter, are divergent series.

40.1 QM

We first examine two situations where we have found instantons. We then argue that, for other analytic potentials, complex solutions to the Euclidean equation of motion are also relevant [380].

40.1.1 Real instantons

The quartic anharmonic oscillator. We first consider the ground-state energy $E(g)$ of the Hamiltonian of the quartic anharmonic oscillator (37.1),

$$H = -\frac{1}{2} (\mathrm{d}/\mathrm{d}q)^2 + \frac{1}{2} q^2 + \frac{1}{4} g q^4.$$

Since $E(g)$ is analytic in the cut-plane and behaves like $g^{1/3}$ for $|g|$ large, it has the Cauchy representation

$$E(g) = \frac{1}{2} + \frac{g}{\pi} \int_{-\infty}^0 \frac{\text{Im } E(g') \mathrm{d}g'}{g'(g' - g)}. \quad (40.2)$$

Expanding the integrand in powers of g , one obtains the integral representation for the perturbative coefficients

$$E_k = \frac{1}{\pi} \int_{-\infty}^0 \frac{\text{Im } E(g) \, dg}{g^{k+1}}, \quad \text{for } k > 0. \quad (40.3)$$

When k , the order in the expansion, becomes large, due to the factor g^{-k} the dispersion integral (40.3) is dominated by the small negative g values.

In Section 37.4 (equation (37.37)), we have evaluated $\text{Im } E(g)$ for g small and negative. We use here the result to estimate the large k behaviour of E_k :

$$E_k \underset{k \rightarrow \infty}{\sim} \frac{1}{\pi} \int_{0^-}^{0^-} \left(\frac{8}{\pi} \right)^{1/2} \frac{1}{\sqrt{-g}} \frac{e^{4/3g}}{g^{k+1}} [1 + O(g)] \, dg. \quad (40.4)$$

The explicit integration yields

$$E_k = (-1)^{k+1} \sqrt{6/\pi^3} (3/4)^k \Gamma(k + 1/2) [1 + O(1/k)]. \quad (40.5)$$

This result confirms that the perturbative series is divergent for all values of $g > 0$, and *determines the nature of the divergence*. Successive corrections to the semi-classical result yield a series in powers of g which, integrated, generate a systematic expansion in powers of $1/k$.

General holomorphic potentials. The same argument is applicable to the situation described in Section 37.5. We can calculate the energy of the metastable state in power series of the coupling constant g by making a systematic expansion around the relative minimum of the potential. On the other hand, we can, as previously mentioned, derive from the knowledge of the imaginary part of the energy level for small coupling, an estimate of the behaviour of the perturbative coefficients at large order. We consider the action

$$\mathcal{S}(q) = \int dt [\frac{1}{2} \dot{q}^2(t) + g^{-1} V(q(t)\sqrt{g})], \quad (40.6)$$

where g is the loop expansion parameter. The analogue of the dispersion integral (40.3) is

$$E_k \sim \frac{1}{\pi} \int_0^\infty \frac{\text{Im } E(g) \, dg}{g^{k+1}}.$$

The behaviour of $\text{Im } E(g)$ for g small is given by the expression (37.95). Integrating near $g = 0$, one obtains (x_+ is a zero of $V(x)$),

$$E_k \sim -\frac{x_+}{2\pi^{3/2}} \exp \left[\int_0^{x_+} \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} \right) dx \right] A^{-(k+1/2)} \Gamma(k + 1/2), \quad (40.7)$$

where A is the instanton action

$$A = 2 \int_0^{x_+} \sqrt{2V(x)} \, dx. \quad (40.8)$$

We now see general features emerging: at large orders, the perturbative coefficients E_k behave like

$$E_k \underset{k \rightarrow \infty}{\sim} C k^{b-1} k! A^{-k}. \quad (40.9)$$

The universal factor $k!$, which is characteristic of a semi-classical or loop expansion, implies that the perturbation series is a divergent series. The factor A^{-k} , where A is the action of the classical solution, is common to all quantities calculated with the same Hamiltonian. The power k^b depends, in particular, on the number of continuous symmetries broken by the classical solution, but it is also on the specific quantity that is expanded. This can be verified by explicitly calculating the imaginary parts of the energy of the excited levels and using equation (40.3). The parameter b is in general a half integer. Finally, the constant multiplicative factor C depends in a more complicated way on all the specific features of the expanded quantity.

Discussion. In both examples, we have been able to derive the large-order behaviour of perturbation series from the decay rate, due to barrier penetration, of a metastable minimum of the potential. For the potentials considered in Section 37.5, the action A is positive and, therefore, all terms in the perturbative expansion have the same sign. The same property holds for the quartic anharmonic oscillator in the metastable case, that is, when g is negative. However, for g positive, in which case perturbation series has been expanded around an absolute minimum of the potential, we observe that the perturbative coefficients oscillate in sign. Also, we note that for $g > 0$, the instanton solution becomes purely imaginary. This suggests how one can derive the large behaviour in the generic stable case.

40.1.2 Complex instantons

So far, we have characterized the large-order behaviour of perturbation theory in two cases: in the generic case, in which we expand around a relative minimum of the potential, and in one special case in which we have expanded around an absolute minimum of the potential, but which by analytic continuation in the coupling constant can be transformed into a relative minimum. We now consider actions of the form (40.6), in which the potential $V(q)$ is still an entire function of q , and satisfies the condition

$$V(q) = \frac{1}{2}q^2 + O(q^3),$$

and we assume that perturbation theory is expanded around $q = 0$, the absolute minimum of the potential. Then, no real instanton solutions can be found. Following the example of the anharmonic oscillator, we thus assume that we can introduce parameters in the potential which make an analytic continuation to a metastable situation possible. We obtain the large-order behaviour from the expression (40.7). We then use the inverse analytic continuation to return to the initial situation. It is plausible that the analytic continuation of the expression (40.7) still gives the large behaviour of the initial expansion.

We can now formulate the rules of the large-order behaviour calculation directly in the initial theory. Complex instanton solutions, with, in general, complex (or exceptionally negative) action [380] are associated to the complex zeros (at finite or infinite distance) of the potential $V(q)$. These instantons are candidates to contribute to the large-order behaviour. In the expression (40.7), we see that the action(s) with the smallest modulus (when the action is complex, there will be at least two complex conjugate actions) gives the leading contribution to the large-order behaviour. Note that the difference we have found between the anharmonic oscillator and the metastable case is generic. In the stable case, the classical action is not real positive, and the perturbative coefficients at large order involve an order-dependent phase factor.

Such a property plays an essential role for the summability of divergent series (see Section 41.1.2).

40.1.3 Degenerate classical minima

The preceding discussion does not apply directly to the case of potentials with non-continuously connected degenerate minima (see Chapter 39), for example, the potentials $x^2(1-x)^2$, or $\cos(x)$. Let us indeed consider such a potential as the limit of a potential which has two minima with very close values of the potential. From the explicit form of the large-order behaviour (40.7), we note that the classical action (40.8) has a limit, which can be identified as being *twice the action of the instanton that connects the two minima of the potential*. This property generalizes to the field-theory examples. However, the calculation of the determinant generated by expanding to quadratic order around the saddle point leads to a new problem.

This is illustrated by the QM example: the integral in expression (40.7) diverges when x_+ is an extremum of the potential. This property can be understood in the following way. When the values at the two minima approach each other, the time spent by the instanton path close to the second minimum of the potential diverges. Therefore, fluctuations which tend to change this time leave the action almost stationary. Correspondingly, one eigenvalue of the operator $\delta^2\mathcal{S}/\delta q(t)\delta q(t')|_{q=q_c}$ goes to 0, and this explains the divergence of expression (40.7) in this case. It becomes necessary to let the separation time fluctuate and, therefore, to introduce an additional time collective coordinate. In this way, one can derive the correct answer [406–408]. Let us also note that here, like in the case of relative minima, the instanton action is positive. Summing the perturbative expansion becomes a more difficult problem, which we discuss in Chapter 42.

40.2 Scalar field theories: The example of the ϕ^4 field theory

In Chapter 38, we have shown how to evaluate the contribution of instantons to the decay rate of metastable states. These results can be applied to large-order behaviour estimates. In a general scalar boson field theory, if instanton solutions can be found, the same arguments applied to n -point correlation functions lead to [384, 380]

$$\left\{ Z^{(n)}(x_1, \dots, x_n) \right\}_{k \rightarrow \infty} \sum_{\substack{\text{dominant} \\ \text{saddle points}}} C_n(x_1, \dots, x_n) k^{b-1} S^{-k} k!, \quad (40.10)$$

in which

- (i) S is the instanton action which, in general, is complex;
- (ii) $b = \frac{1}{2}(n + \delta)$, δ being the number of symmetries broken by the classical solution;
- (iii) $C_n(x_1, \dots, x_n)$, which does not depend on k , contains the whole dependence in the arguments of the correlation function.

In the case of the ϕ^4 field theory, the discontinuity across the cut of the n -point function reads (equation (38.19), where the instanton action is $S = -A$)

$$\text{disc. } Z^{(n)}(x_1, \dots, x_n) \underset{g \rightarrow 0_-}{\sim} \left(\frac{-S}{2\pi} \right)^{d/2} \Omega \frac{e^{-S/g}}{(-g)^{(d+n)/2}} F_n(x_1, \dots, x_n), \quad (40.11)$$

with

$$\Omega = (\det M' M_0^{-1})_{\text{ren.}}^{-1/2},$$

and

$$F_n(x_1, \dots, x_n) = m^{d+n(d-2)/2} 6^{n/2} \int d^d x_0 \prod_{i=1}^n f(m(x_i - x_0)). \quad (40.12)$$

Table 40.1

The coefficients β_k of the coupling constant RG-function $\beta(g)$ divided by the large-order estimate, in the case of the $O(N)$ -symmetric $(\phi^2)_{d=3}^2$ field theory.

k	2	3	4	5	6	7
$N = 0$	3.53	1.55	1.185	1.022	0.967	0.951
$N = 1$	3.98	1.75	1.32	1.120	1.050	1.023
$N = 2$	4.82	2.09	1.53	1.29	1.20	1.15
$N = 3$	6.14	2.58	1.86	1.55	1.41	1.35

Using previous arguments, we can immediately translate this result into the large-order behaviour estimate for correlation functions,

$$\left\{ Z^{(n)}(x_1, \dots, x_n) \right\}_k = \frac{1}{2i\pi} \int_{-\infty}^0 \frac{dg}{g^{k+1}} \text{disc. } Z^{(n)}(x_1, \dots, x_n),$$

and, therefore,

$$\left\{ Z^{(n)}(x_1, \dots, x_n) \right\}_{k \rightarrow \infty} \sim \frac{1}{2i\pi} \frac{\Omega}{(2\pi)^{d/2}} F_n(x_1, \dots, x_n) (-1)^k \frac{\Gamma(k + (d+n)/2)}{(-S)^{n/2+k}}. \quad (40.13)$$

Example: the renormalization group β -function in the $(\phi^2)^2$ field theory in dimension 3. The large-order behaviour has been determined numerically by solving the field equations to determine the instanton action S , and then by evaluating the determinant [409]. The predictions of the asymptotic formulae can be compared with the available terms of the series (see Section 41.3.1). The agreement is rather good and strongly suggests that the large-order behaviour estimates are indeed correct (see Table 40.1).

40.3 The $(\phi^2)^2$ field theory in dimension 4 and $4 - \varepsilon$

As a by-product of the calculation of the instanton contribution in Sections 38.3–38.6, one can evaluate the semi-classical contribution to the large-order behaviour in the $(\phi^2)^2$ field theory in four dimensions. However, because the theory is exactly renormalizable, at order k , as a consequence of their large momenta properties, some diagrams grow themselves like $k!$, generating additional contributions to the large-order behaviour (*e.g.*, see Section 18.5). Moreover, infrared (IR) singularities in the massless theory also yield contributions of order $k!$, but with different signs.

40.3.1 Semi-classical contribution

The instanton contribution to the large-order behaviour for vertex functions is given by

$$\left\{ \tilde{\Gamma}^{(n)}(p_1, \dots, p_n) \right\}_k = \frac{1}{\pi} \int_{-\infty}^0 \frac{\text{Im } \tilde{\Gamma}^{(n)}(p_1, \dots, p_n)}{g^{k+1}} dg. \quad (40.14)$$

This yields a result of the form

$$\left\{ \tilde{\Gamma}^{(n)}(p_1, \dots, p_n) \right\}_{k \rightarrow \infty} \sim \tilde{C}_n(p_1, \dots, p_n) \int^{0-} \frac{e^{8\pi^2/3g}}{(-g)^{n+5/2}} \frac{dg}{g^{k+1}}. \quad (40.15)$$

After integration, one obtains

$$\left\{ \tilde{\Gamma}^{(n)}(p_1, \dots, p_n) \right\}_k \sim \tilde{C}_n(p_1, \dots, p_n) (-1)^k \left(\frac{3}{8\pi^2} \right)^{n+3+k} \Gamma(k + n/2 + 5/2). \quad (40.16)$$

From this expression, it is simple to derive the semi-classical contribution to the large order behaviour of various renormalization-group (RG) functions in, for example, the fixed momentum subtraction scheme. A comparison between large order behaviour and explicit calculations can be found in Table 40.2, in the case of the RG β -function.

Finally, note that, in the massive theory, the calculation is slightly modified because the integral over the collective dilatation coordinate is cut at a scale of order $m\sqrt{k}$ (see Section 38.7).

Table 40.2

The coefficients β_k of the RG β -function divided by the semi-classical asymptotic estimate, in the case of the $O(N)$ -symmetric $((\phi)^2)_{d=4}^2$ field theory.

k	2	3	4	5
$N = 1$	0.10	0.66	1.08	1.57
$N = 2$	0.06	0.49	0.87	1.32
$N = 3$	0.04	0.33	0.66	1.09

40.3.2 Ultraviolet (UV) and infrared (IR) (renormalons) contributions

So far, an implicit assumption in the large-order behaviour calculation has been that the singularities of correlation functions come entirely, in the neighbourhood of the origin, from barrier penetration effects. If this assumption is certainly correct in QM, if there is convincing evidence that it is valid for super-renormalizable theories, it is much more questionable for renormalizable theories, in the absence of a finite UV cut-off, or for massless renormalizable theories. We first explain the large momentum problem, and then the IR problem of massless theories [410].

UV singularities: Renormalons [411]. If the semi-classical analysis is valid for the cut-off regularized field theory, it becomes somewhat formal for the renormalized theory in the infinite cut-off limit. We have already seen that even in the straightforward calculation, non-trivial questions arise about the global RG properties of the theory. A direct investigation of the perturbative expansion raises new questions, and suggests that UV singularities yield additional contributions to the large-order behaviour.

Let us consider the $O(N)$ -symmetric $(\phi^2)^2$ field theory, where the field ϕ is an N -component vector, in dimension 4. The action has the form

$$\mathcal{S}(\phi) = \int d^4x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{m^2}{2} \phi^2(x) + \frac{g}{4} (\phi^2(x))^2 \right]. \quad (40.17)$$

In Section 18.5, we have shown that, at order $1/N$ in the large N expansion, the renormalized two-point function is given by a divergent integral, because the integrand has a pole, corresponding to the Landau ghost.

We briefly recall the argument. The $1/N$ contribution to the two-point function in the massive renormalized theory is

$$F_2(p) = \frac{2g}{(2\pi)^4} \int \frac{d^4 q}{[(p+q)^2 + m^2] [1 + NgB_r(q)]} - \text{subtractions}, \quad (40.18)$$

where the renormalized ‘bubble’ diagram is given by

$$B_r(p) = \frac{1}{(2\pi)^4} \int \frac{d^4 q}{[(p+q)^2 + m^2] (q^2 + m^2)} - \text{subtraction}. \quad (40.19)$$

For $|p| \rightarrow \infty$, $B_r(p)$ behaves like

$$B_r(p) \sim \frac{1}{8\pi^2} \ln(m/|p|). \quad (40.20)$$

Therefore, the sum of the bubble diagrams which appears in expression (40.18) has a singularity for g small (which justifies the large momentum approximation) and positive, at a momentum

$$|p| \sim m e^{8\pi^2/Ng}, \quad \text{for } g \rightarrow 0_+. \quad (40.21)$$

Since the theory is IR-free, and not UV asymptotically free, this singularity occurs for positive values of the coupling constant. Once this sum of bubbles is inserted into expression (40.18), it produces a cut for g small and positive. More precisely, after subtraction, and for q large, the integrand of F_2 at large momenta behaves like

$$\int_{|q| \gg 1} \frac{dq}{q^3} \left[1 + \frac{Ng}{8\pi^2} \ln(m/q) \right]^{-1} + \dots \quad (40.22)$$

The change of variables $t = \ln(q/m)$ transforms the expression (40.22) into

$$\int_0^\infty dt e^{-2t} \frac{1}{1 - Ngt/(8\pi^2)}. \quad (40.23)$$

This yields an imaginary contribution to the correlation functions for g small and positive of the form $\exp(-16\pi^2/Ng)$. Alternatively, by expanding expression (40.18) in powers of g , we obtain the contribution of individual diagrams containing bubble insertions. These diagrams behave like $(N/16\pi^2)^k k!$ at large order k . Therefore, in contrast to super-renormalizable theories in which an individual diagram behaves like an exponential of k , and the $k!$ comes from the number of diagrams, here, some individual diagrams give a $k!$ contribution, without the sign oscillations characteristic of the semi-classical result.

Further investigations show that, if a non-perturbative contribution exists, it should satisfy the homogeneous RG equations. For simplicity, we consider the example of a dimensionless ratio of correlation functions $R(p/m, g)$ without anomalous dimensions,

$$\left(m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g} \right) R(p/m, g) = 0. \quad (40.24)$$

The RG equation implies that the function $R(p/m, g)$ is actually a function of only one variable $s(g)p/m$, in which $s(g)$ then satisfies

$$\beta(g)s'(g) = s(g), \quad (40.25)$$

which after integration yields

$$s(g) \sim \exp \left[\int^g \frac{dg'}{\beta(g')} \right]. \quad (40.26)$$

For g small, $s(g)$ behaves like

$$\beta(g) = \beta_2 g^2 + O(g^3), \quad \text{with} \quad \beta_2 = \frac{N+8}{8\pi^2}, \quad (40.27)$$

$$s(g) \underset{g \rightarrow 0}{\propto} g^{-\beta_3/\beta_2^2} e^{-1/\beta_2 g}. \quad (40.28)$$

Since the correlation function depends only on the mass squared, only $s^2(g)$ enters the calculation, and the contribution to the large-order behaviour has the form

$$\int_0 \frac{s(g)}{g^{k+1}} dg \propto (\beta_2/2)^k \Gamma(k+1+2\beta_3/\beta_2^2), \quad (40.29)$$

a result which coincides, in the large N limit, with the contribution that we obtained from the set of bubble diagrams. This potential contribution has to be compared with the semi-classical result (40.16).

In fact, this problem is related to the question of the existence of the renormalized $(\phi^2)^2$ field theory in four dimensions. If the theory does not exist, then probably the sum of perturbation theory is complex for g positive, and these singular terms, sometimes called *renormalon* effects, are the small coupling evidence of this situation. More generally, the existence of renormalons shows that the perturbation series is not Borel summable and does not define unique correlation functions (see Section 41.1).

Massless renormalizable theories: IR renormalons [412]. Again, we illustrate the problem with the $(\phi^2)^2$ field theory in the large N limit. We now work in a massless theory with fixed cut-off Λ . We evaluate the contribution of the small momentum region to the mass renormalization constant. The bubble diagram (40.19) behaves like

$$I(p) \sim \frac{1}{8\pi^2} \ln(\Lambda/p).$$

The sum of bubbles yields a contribution to the mass renormalization proportional to

$$\int^\Lambda \frac{d^4 q}{q^2(1+N g I(q))} = \int \frac{d^4 q}{q^2(1+\frac{N}{8\pi^2} g \ln(\Lambda/q))}.$$

Expanded in powers of g , the integral yields a contribution of order $(-1)^k (N/16\pi^2)^k k!$ for large order k . This contribution has the sign oscillations of the semi-classical term. More generally, for finite N one finds $(-\beta_2/2)^k k!$. IR singularities yield an additional Borel summable contribution to the large-order behaviour.

For massless but asymptotically free theories, the role of the IR and UV regions are interchanged. UV renormalons are expected to yield additional singularities to the Borel transform on the real negative axis, while IR contributions destroy Borel summability. When these theories have real instantons like QCD or the $CP(N-1)$ models (see Section 39.5, 39.6), the Borel transform has also semi-classical singularities on the real positive axis.

Table 40.3

Sum of the successive terms of the ε -expansion of γ and η for $\varepsilon = 1$ and $N = 1$.

k	0	1	2	3	4	5
γ	1.000	1.1667	1.2438	1.1948	1.3384	0.8918
η	0.0...	0.0...	0.0185	0.0372	0.0289	0.0545

40.3.3 Wilson–Fisher’s ε -expansion

In the theory of critical phenomena, many universal physical quantities have been calculated as power series of $\varepsilon = 4 - d$, where d is the space dimension (Section 15.4). The physical dimensions correspond to $\varepsilon = 1, 2$, that is, three and two dimensions. As Table 40.3 illustrates, the ε -expansion generates divergent series. Divergent series can be used for small values of the argument. However, only a finite number of terms of the series can then be taken into account. The last added term gives an indication of the size of the irreducible error. Therefore, for the critical exponents γ and η we conclude from the series displayed in Table 40.3,

$$\gamma = 1.244 \pm 0.050, \quad \eta = 0.037 \pm 0.008,$$

where the errors are only indicative of the uncertainty about the value.

The precision of the estimates can only be improved if the ε expansion is Borel summable (see Section 41.1), and thus free of renormalon singularities.

In the minimal subtraction (MS) scheme, the RG functions have a simple form (see Section 10.4): only the RG β -function depends on ε with the explicit dependence,

$$\beta(g, \varepsilon) = -\varepsilon g + \beta(g, 0), \quad \text{with } \beta(g, 0) = \frac{N+8}{48\pi^2} g^2 + O(g^3).$$

The large-order behaviour of the ε -expansion can only be guessed because, as discussed previously, it vanishes at leading order. A calculation of the next order would be necessary, and this has not yet been done. Since, at leading order, the fixed point constant $g^*(\varepsilon)$ solution of $\beta(g, \varepsilon) = 0$ is

$$g^*(\varepsilon) = 48\pi^2 \varepsilon / (N+8) + O(\varepsilon^2),$$

except if for some unknown reason the accident of leading order persists, the ε -expansion is likely to involve a factor $(-3/(N+8))^k k!$ multiplied by an unknown power of k .

Finally, we note that, at leading order in the $1/N$ expansion for the Wilson–Fisher ε -expansion, and thus also for suitably defined RG functions like by the MS scheme, the renormalon singularities cancel. We conjecture on this basis, and on the basis of the numerical evidence presented in Chapter 41, that the ε -expansion of universal quantities is free of renormalon singularities, and can be Borel summable.

40.4 Field theories with fermions

In the case of boson field theories, we have related the large-order behaviour of perturbation theory to the decay of the false vacuum for, in general, non-physical values of the coupling constant. Therefore, we expect some modifications if we consider a system of self-interacting fermions, or of fermions interacting with bosons that themselves have no self-interaction. (The first case can be reduced to the second one by introducing an auxiliary boson field but additional difficulties arise.) Indeed, the Pauli principle renders the decay of a false vacuum more difficult, because several fermions cannot occupy the same state to create a classical field, and this effect is especially strong in low dimensions. Note that if the bosons have self-interactions, these interactions drive the decay of the vacuum, and fermions no longer play a role, at least at leading order.

Seen from the point of view of integrals, the difference between fermions and bosons is also immediately apparent. We have shown that the simple integral counting the number of Feynman diagrams, which is also the ϕ^4 field theory in $d = 0$ dimensions, already has the characteristic $k!$ behaviour at large orders.

By contrast, let us consider an example of a zero-dimensional fermion theory, the integral over a finite number of fermion degrees of freedom,

$$I(\lambda) = \int \prod_{i=1}^N (\mathrm{d}\bar{\xi}_i \mathrm{d}\xi_i) \exp \left[\sum_{i,j} \bar{\xi}_i D_{ij} \xi_j + \lambda \sum_{i,j,k,l} C_{ijkl} \bar{\xi}_i \bar{\xi}_j \xi_k \xi_l \right]. \quad (40.30)$$

The quantities ξ_i and $\bar{\xi}_i$ are anticommuting (Grassmann) variables and D_{ij} and C_{ijkl} are sets of real or complex numbers. Because we assume a finite number of anticommuting variables, the expansion of the exponential yields a polynomial, and thus, $I(\lambda)$ is a polynomial in λ , by contrast with the boson case.

40.4.1 Example of a Yukawa-like QFT

We now consider the vacuum amplitude or partition function of the Yukawa-like theory with Dirac fermions $\bar{\psi}(x)$, $\psi(x)$, and a scalar boson $\phi(x)$:

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)] [\mathrm{d}\bar{\psi}(x)] [\mathrm{d}\psi(x)] \exp [-\mathcal{S}(\phi, \bar{\psi}, \psi)], \quad (40.31)$$

in which the action is

$$\begin{aligned} \mathcal{S}(\phi, \bar{\psi}, \psi) = & \int \mathrm{d}^d x \left[-\bar{\psi}(x) (\not{\partial} + M + \sqrt{g} \phi(x)) \psi(x) \right. \\ & \left. + \frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) \right]. \end{aligned} \quad (40.32)$$

The parameter g is a loop expansion parameter. Since a fermion field has no classical limit, the expression (40.31) is not directly suited to the study of the vacuum decay. In fact, we expect the fermion fields to generate an effective interaction for the boson field $\phi(x)$, and this effective interaction will lead to the decay of the vacuum. This suggests that one should integrate over the ψ and $\bar{\psi}$ variables and study the instantons of the effective theory for $\phi(x)$. In addition, the zero-dimensional example has shown that the fermion integration gives some hints about the analytic structure of the theory. The integration over ψ and $\bar{\psi}$ yields ($\ln \det = \mathrm{tr} \ln$)

$$\begin{aligned} \mathcal{Z} = & \int [\mathrm{d}\phi(x)] \exp \left\{ -\frac{1}{2} \int \mathrm{d}^d x \left[(\nabla \phi(x))^2 + m^2 \phi^2(x) \right] \right. \\ & \left. + \mathrm{tr} \ln [\not{\partial} + M + \sqrt{g} \phi(x)] \right\}. \end{aligned} \quad (40.33)$$

We now face a new difficulty, arising from the integration: the generated effective action is not local in $\phi(x)$, and leads to non-local field equations. However, because we are concerned only with the determination of the large behaviour, we can simplify the effective action. The determinant generated by the fermion integration is, at least for the class of relevant $\phi(x)$ fields, an entire function of the coupling constant \sqrt{g} . Therefore, essential singularities can only be generated by the infinite range of the ϕ -integration. It is thus sufficient to evaluate the the determinant, generated by the field integration, for large fields $\phi(x)$ [413]. This situation has to be contrasted with what would have happened if $\psi(x)$ and $\bar{\psi}(x)$ would have been commuting variables. The integration then would have generated the inverse of the determinant function which has singularities for all zeros in g of the determinant. These singularities would have yielded essential singularities in the coupling constant after integration. Finally, we note that this difference, determinant versus inverse determinant, is responsible for the minus sign for each fermion loop in perturbation theory which makes cancellations possible.

40.4.2 Evaluation of the fermion determinant for large fields

As a preparatory exercise, we first solve a similar problem in which, however, the additional complication due to the spin structure is absent.

The Fredholm determinant of a Schrödinger operator for large potentials. We consider a Schrödinger operator, in d space dimensions, for smooth potentials of the form $\lambda V(x)$, with $V(x) \geq 0$, in the limit $\lambda \rightarrow +\infty$. We want to evaluate the logarithm of its normalized Fredholm determinant,

$$\Sigma(\lambda) = \ln \det \left\{ [-\nabla^2 + \lambda V(x)] [-\nabla^2 + \mu^2]^{-1} \right\}, \quad (40.34)$$

where μ is a mass parameter. On intuitive grounds, we expect the determinant to converge towards a local functional. To derive this property, we differentiate the determinant with respect to λ . Using the identity $d(\ln \det \Omega) = \text{tr } d\Omega \Omega^{-1}$, we find

$$\Sigma'(\lambda) = \int d^d x V(x) \langle x | [-\nabla^2 + \lambda V]^{-1} | x \rangle,$$

in the quantum bra–ket notation. When λ becomes large, the operator $[-\nabla^2 + \lambda V]^{-1}$ converges towards a local operator. Thus, we can replace V operator by its expectation value $V(x)$ in the state $|x\rangle$. In the Fourier representation,

$$\Sigma'(\lambda) \sim \frac{1}{(2\pi)^d} \int d^d x V(x) \frac{d^d p}{p^2 + \lambda V(x)} = \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} \lambda^{d/2-1} \int d^d x V^{d/2}(x), \quad (40.35)$$

a result formally valid for $d < 2$. The integration over λ yields

$$\Sigma(\lambda) \sim -\frac{\Gamma(-d/2)}{(4\pi)^{d/2}} \int d^d x (\lambda V(x))^{d/2} + \text{constant}. \quad (40.36)$$

For $d \geq 2$, we know that the expression (40.34), which has the form of a one-loop diagram in a scalar field theory, has to be renormalized. For $d = 2$, a mass counter-term has to be added. One then obtains the evaluation

$$\begin{aligned} \Sigma(\lambda) &\sim \lim_{d \rightarrow 2} \left\{ -\frac{1}{(4\pi)^{d/2}} \int d^d x \left[\Gamma(-d/2)(\lambda V(x))^{d/2} + \Gamma(1-d/2)\lambda V(x) \right] \right\}, \\ &\sim -\frac{1}{4\pi} \int d^2 x \lambda V(x) \ln(\lambda V(x)). \end{aligned} \quad (40.37)$$

In the same limit, for $d = 3$, one obtains

$$\Sigma(\lambda) \sim -\frac{1}{12\pi} \int d^3 x (\lambda V(x))^{3/2}. \quad (40.38)$$

For $d = 4$, a counter-term quadratic in $V(x)$ is required. Then, one finds

$$\Sigma(\lambda) \sim \frac{1}{32\pi^2} \int d^4 x (\lambda V(x))^2 \ln(\lambda V(x)). \quad (40.39)$$

The fermion determinant. A similar method can be used to evaluate the contribution of the fermion determinant

$$\Sigma(\lambda) \equiv \ln \det (\not{D} + \lambda V) (\not{D} + M)^{-1} \quad (40.40)$$

where we have set $\lambda V(x) = M + \sqrt{g}\phi(x)$, and $V(x)$ is assumed to be smooth.

Again, we differentiate with respect to λ , and obtain

$$\Sigma'(\lambda) = \text{tr}_\gamma \int d^d x V(x) \langle x | [\partial + \lambda V]^{-1} | x \rangle.$$

In the large λ limit, the operator becomes local, and, again, we replace the operator V by its expectation value in the state $|x\rangle$. In the Fourier representation, we obtain

$$\begin{aligned} \Sigma'(\lambda) &= \frac{1}{(2\pi)^d} \text{tr}_\gamma \int d^d x V(x) \int d^d p [ip + \lambda V(x)]^{-1} \\ &= \frac{1}{(2\pi)^d} \text{tr}_\gamma \int d^d x V(x) \int \frac{d^d p (-ip + \lambda V(x))}{p^2 + \lambda^2 V^2(x)}. \end{aligned}$$

Then, the trace over γ matrices can be taken, $\text{tr} \not{p} = 0$, and we set $\text{tr} \mathbf{1} = N$. The remaining part of the calculation is analogous to the case of the Schrödinger equation. Integrating, one finds

$$\Sigma'(\lambda) = \frac{N}{(4\pi)^{d/2}} \Gamma(1 - d/2) \lambda^{d-1} \int d^d x |V(x)|^d.$$

Finally, integrating over λ and returning to the initial parametrization, one obtains the large field behaviour,

$$\Sigma \sim -\frac{N}{2} \frac{\Gamma(-d/2)}{(4\pi)^{d/2}} \int d^d x |M + \sqrt{g}\phi(x)|^d \sim -\frac{N}{2} \frac{\Gamma(-d/2)}{(4\pi)^{d/2}} g^{d/2} \int d^d x |\phi(x)|^d. \quad (40.41)$$

40.4.3 The large-order behaviour

We now determine the essential singularity of the field theory at $g = 0$ from the properties of the effective local action

$$\mathcal{S}_{\text{eff.}}(\phi) = \int d^d x \left[\frac{1}{2} (\nabla \phi(x))^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{N}{2} \frac{\Gamma(-d/2)}{(4\pi)^{d/2}} g^{d/2} \phi^d(x) \right]. \quad (40.42)$$

It should be understood that for d even, the necessary counter-terms are provided to render the action finite. We have to look for instanton solutions of the corresponding field equations. Since this particular model is not interesting in itself, we will not solve the field equations explicitly, but only assume the existence of a solution. We then rescale the field ϕ ,

$$\phi(x) \mapsto \phi(x) g^{-d/2(d-2)} \quad (40.43)$$

to factorize the dependence on g in front of the classical action. The classical action, calculated for a solution, thus takes the form

$$\mathcal{S}(\phi_c) = (A/g)^{d/d-2}, \quad (40.44)$$

where A does not depend on g . Introducing this form into the Cauchy representation, we find

$$\mathcal{Z}_k \underset{k \rightarrow \infty}{\sim} \int_0^\infty \frac{e^{-(A/g)^{d/d-2}}}{g^{k-1}} dg. \quad (40.45)$$

The integration yields the large-order estimate

$$\mathcal{Z}_k \sim A^{-k} \Gamma [k(d-2)/d]. \quad (40.46)$$

We observe that, as expected, this theory is less divergent than a self-interacting boson field theory. The boson result is recovered (in a cut-off field theory) for d large, because the Pauli principle becomes decreasingly effective when the dimension increases. For $d = 2$, the expression (40.46) becomes

$$\mathcal{Z}_k \sim A^{-k} (\ln k)^k, \quad (40.47)$$

in agreement with a rigorous bound [414] that states $|\mathcal{Z}_k| < (k!)^\varepsilon$ for all $\varepsilon > 0$.

Remark. To compare the contributions of boson and fermion interactions, we have implicitly assumed a loop expansion. Then, at each order, the boson contributions always dominate the large-order behaviour. If we group the diagrams differently, this may no longer be the case. Let us again consider the theory defined by the action (40.32) in four dimensions. In four dimensions, this theory cannot be renormalized without the addition of a $\lambda\phi^4$ counter-term. Thus, renormalization requires introducing a boson self-interaction. But it is consistent with renormalization to consider λ as being of order g^2 . Then, both interaction terms $\bar{\psi}\psi\phi$ and ϕ^4 give contributions of the same order to the large-order behaviour.

40.4.4 The example of QED

Potentially quite interesting applications of the preceding analysis are gauge theories. While non-Abelian theories involve the additional problem of degenerate classical minima [415] (see also Section 40.1.3 and Chapter 42), QED [416, 417] has been more extensively studied.

The action has formally the same structure as in the Yukawa theory, but one additional complication then arises. The fermion integration yields the determinant (see Section 21.3),

$$D(e) = \det(\not{D} + m), \quad \text{with } \not{D} = \sum_{\mu} \gamma_{\mu} D_{\mu}, \text{ and } D_{\mu} = \partial_{\mu} + ieA_{\mu}(x). \quad (40.48)$$

To estimate $D(e)$ for large charge e , we can use the equation (we assume a dimension d even, see for example, Section A21.3)

$$D^2(e) = \det \left(m^2 - \sum_{\mu} D_{\mu}^2 - \frac{1}{2} e \sum_{\mu, \nu} \sigma_{\mu\nu} F_{\mu\nu}(x) \right). \quad (40.49)$$

In the large e limit, the last term, which is of order e , is negligible with respect to D_{μ}^2 which is of order e^2 (N_d is the loop factor):

$$\ln D(e) \sim \frac{1}{2} N_d \text{tr} \ln \left(m^2 - \sum_{\mu} D_{\mu}^2 \right). \quad (40.50)$$

However, the determination of the large-coupling constant behaviour is more subtle than before. A direct calculation of the determinant has not been performed. One difficulty is related to the property that, due to gauge invariance, the gauge degree of freedom of the gauge field cannot be considered as slowly varying.

In particular, the constant field approximation is not meaningful. Moreover, since perturbation theory implies gauge fixing, it is sufficient to calculate the determinant in a fixed gauge.

Therefore, it has been conjectured, on the basis of studying the determinant for special gauge fields, that the behaviour of the determinant is given for large e by

$$\ln D(e) \sim C(d) \int d^d x |e[A_T]_\mu(x)|^d, \quad \text{with } C^{-1}(d) = d(4\pi)^{(d-1)/2} \Gamma((d+1)/2),$$

where $[A_T]_\mu$ is the transverse part of A_μ :

$$[A_T]_\mu(x) = A_\mu(x) - \nabla^{-2} \partial_\mu \sum_\nu \partial_\nu A_\nu(x).$$

This result is gauge invariant, as it should, but not local, except in the gauge $\sum_\mu \partial_\mu A_\mu = 0$. It agrees for $d = 2$ with the exact result (30.59) obtained from the Abelian anomaly ($C(2) = 1/2\pi$). For $d = 4$, the case of physical interest, $C(4) = 1/12\pi^2$. The effective classical field theory then is scale invariant. Arguments related to conformal invariance can be used to construct some ansatz for the instanton solutions. Two kind of solutions have been explored in Refs. [416] and [417]. Taking the minimal action solution, one obtains an evaluation of the form

$$\mathcal{Z}_k \sim (-1)^k A^{-k} \Gamma(k/2), \quad A = 4.886, \quad (40.51)$$

the expansion parameter being $\alpha = e^2/4\pi$. It is worth mentioning that this evaluation is probably not very useful as a practical method to predict new orders in QED, for several reasons. First, the theory is not asymptotically free and thus has a potential renormalon problem, which can be understood by inserting in a Feynman diagram the one-loop corrected photon propagator. Second, the cancellation coming from the sign of fermion loops does not seem to be very effective at low orders. Therefore, an alternative calculation, which leads to a large-order behaviour at a fixed number of fermion loops, seems to be more useful. Predictions of this kind made for diagrams with one fermion loop, seem to agree better with numerical estimates [416].

A40 large-order behaviour: Additional remarks

large-order behaviour for simple integrals. We consider the integral of Chapter 37,

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp \left[-\left(\frac{1}{2}x^2 + \frac{1}{4}gx^4 \right) \right] dx. \quad (A40.1)$$

We can expand $I(g)$ in a power series,

$$I(g) = \sum_0^{\infty} I_k g^k. \quad (A40.2)$$

The coefficients I_k enumerate the number of vacuum Feynman diagrams with the proper weights in a ϕ^4 field theory. In Section 37.2, we have evaluated the imaginary part of $I(g)$ for g negative by the steepest descent method. The contributions of the non-trivial saddle points yield

$$\text{Im } I(g) \underset{g \rightarrow 0_-}{\sim} 2^{-1/2} e^{1/4g}. \quad (A40.3)$$

Therefore, for k large, I_k behaves as

$$I_k \underset{k \rightarrow \infty}{\sim} \frac{1}{\pi\sqrt{2}} (-4)^k (k-1)! . \quad (A40.4)$$

The result suggests the following interpretation of the large-order behaviour formulae obtained in Chapter 40: in the case of the anharmonic oscillator, and the ϕ^4 field theory, the number of Feynman diagrams is of the order of $4^k k!$, for k large, and a typical diagram behaves at large orders as $(4S)^{-k}$, where k is the order but also, up to an additive constant, the number of loops, and S the instanton action.

QM: Other perturbative expansions. Although we have only discussed large-order behaviour estimates for loop expansions, it is possible to generalize the analysis for perturbative expansions in other parameters. For example, we consider the action

$$\mathcal{S}(q) = \int \left[\frac{1}{2}\dot{q}(t)^2 + \frac{1}{2}q^2(t) + \lambda V(q(t)) \right] dt, \text{ with } V(q) = \sum_2^{2N} V_n q^n,$$

and we want to evaluate the large-order behaviour of the expansion in powers of λ .

The divergence of the perturbative expansion is a consequence of the infinite range of the q integration. Therefore, it is dominated by the large q behaviour of V , and thus related to the instantons associated to the action in which $V(q)$ is replaced by its term of highest degree. Then, after a rescaling of $q(t)$, we find that the classical solution $q_c(t)$ takes the form

$$q_c(t) = \lambda^{-1/(2N-2)} f(t). \quad (A40.5)$$

The term of degree n in $V(q)$ gives a contribution to the classical action proportional to $\lambda^{1-n/(2N-2)}$. For $\lambda \rightarrow 0$, one verifies that the term of highest degree gives indeed the largest contribution to the action. The saddle point in λ in the dispersion relation for large-order k is of the order

$$\lambda \sim k^{-(N-1)}.$$

Thus, the term of degree n in the potential generates a factor of the form

$$\exp \left[c_n k^{2/(n+2-2N)} \right],$$

which is relevant, at leading order, only for $n \geq 2N-2$.

41 Critical exponents and equation of state from series summation

Universal quantities near the phase transition of $O(N)$ symmetric vector models, can be determined, in the framework of the $(\phi^2)^2$ field theory, and the corresponding renormalization group (RG), in the form of perturbative series.

The $O(N)$ symmetric $(\phi^2)^2$ field theories do not describe the universal properties only of ferromagnetic systems. In Section 15.8 we have explained how the $N = 0$ limit is related to the statistical properties of polymers, in Section 15.9 how the Ising-like $N = 1$ model describes the physics of the liquid–vapour transition and, in Section 15.10, why the superfluid helium transition corresponds to $N = 2$.

Universal quantities have been calculated within two different schemes, the Wilson–Fisher $\varepsilon = 4 - d$ expansion [75], and perturbative expansion at fixed dimensions 2 and 3, as suggested by Parisi [146], using series reported in Refs. [418–420]. In both cases, in Sections 40.2 and 40.3, we have shown that the series are divergent (see Refs. [421] for mathematical details), and the expansion parameters are not small.

In fixed dimensions smaller than 4, the series are proven to be Borel summable [422]. For the ε expansion, there are reasons to believe that the renormalon contributions cancel, and that the property is equally true, but a proof is lacking. With this assumption, in both cases, although the series are divergent, they define unique functions.

Since the expansion parameters are not small, summation methods are then required to determine these functions. A specific summation method, based on a parametric Borel transformation and mapping has been successfully applied to the series, and has led to a precise evaluation of critical exponents and other universal quantities [423–444, 160].

In the first part of the chapter we define Borel summable series. We describe methods to sum Borel summable divergent series, as generated in quantum mechanics (QM) and in quantum field theory (QFT). In particular, we show how, in some methods based on the Borel transformation, the knowledge of the large-order behaviour of perturbation theory (see Sections 40.1.1, 40.2) can be efficiently incorporated.

41.1 Divergent series: Borel summability, Borel summation

Asymptotic series. We consider a function $f(z)$, analytic in a sector S ,

$$S : \quad |\operatorname{Arg} z| \leq \alpha/2, \quad |z| \leq |z_0|. \quad (41.1)$$

We assume that, in S , f has the asymptotic (Taylor series) expansion

$$f(z) = \sum_{k=0}^{\infty} f_k z^k. \quad (41.2)$$

This means that the series (41.2) diverges for all $z \neq 0$, and that, in S , it satisfies a bound of the form

$$\left| f(z) - \sum_{k=0}^N f_k z^k \right| \leq C_{N+1} |z|^{N+1}, \quad \forall N, \quad (41.3)$$

in which

$$C_N |z|^N \xrightarrow[\text{for } N \rightarrow +\infty]{} \infty, \quad \forall z \neq 0.$$

Although the series (41.2) diverges, it can nevertheless be used to estimate the function $f(z)$ for $|z|$ small. At $|z|$ fixed, we can look for a minimum in the bound (41.3) when N varies. If $|z|$ is small enough, the bound first decreases with N and then, since the series is divergent, it eventually increases. A truncation of the series at the minimum yields the best possible estimate of $f(z)$, with a finite error $\varepsilon(z)$. Assuming, for definiteness, that the coefficients C_N have the form

$$C_N = M A^{-N} (N!)^\beta, \quad (41.4)$$

one can estimate $\varepsilon(z)$ explicitly. One finds

$$\varepsilon(z) = \min_{\{N\}} C_N |z|^N \sim \exp \left[-\beta (A/|z|)^{1/\beta} \right]. \quad (41.5)$$

We see that an asymptotic series does not in general define a unique function. Indeed, if we have found one function, we can add to it any function analytic in the sector (41.1) and which is smaller than $\varepsilon(z)$ in the whole sector. The new function still satisfies the condition (41.3).

41.1.1 Borel summability. Borel summation

In the case of the specific examples (41.4), a classical theorem of complex analysis states that no analytic function can satisfy a bound of the form (41.3) in a sector of angle α larger than $\pi\beta$. Therefore, for $\alpha > \pi\beta$, although the series is divergent, it defines a unique function.

In the marginal case in which the series is asymptotic only in the open interval $|\operatorname{Arg} z| \in (-\pi\beta/2, \pi\beta/2)$, additional conditions have to be imposed to prove uniqueness.

Borel transformation. We focus from now on to the case $\beta = 1$, which is typical for perturbative expansions, and $\alpha > \pi$, but the generalization to arbitrary β is simple. The unique function can be determined by introducing its Borel transform (Watson's theorem),

$$B_f(z) = \sum_{k=0}^{\infty} \frac{f_k}{k!} z^k. \quad (41.6)$$

Formally, in the sense of power series,

$$f(z) = \int_0^{\infty} ds e^{-s} B_f(sz). \quad (41.7)$$

The bound (41.3) and the estimate (41.4) lead to the bound on the coefficients B_k ,

$$|B_k| < M A^{-k}. \quad (41.8)$$

Thus, $B_f(z)$ is analytic at least in a circle of radius A and uniquely defined by the series. Moreover, it can be proved that $B_f(z)$ is also analytic in a sector

$$|\operatorname{Arg} z| \in [0, \frac{1}{2}(\alpha - \pi)], \quad (41.9)$$

and does not increase faster than an exponential in the sector, so that integral (41.7) converges for $|z|$ small enough and inside the sector

$$|\operatorname{Arg} z| < \alpha/2.$$

In addition, it can be shown that the integral (41.7) satisfies a bound of type (41.3). Hence, this integral representation yields the unique function that has the asymptotic expansion (41.2) in the sector S . The series (41.2) is then called *Borel summable*.

41.1.2 Large-order behaviour and Borel summability

For a large class of potentials in QM, and for a number of field theories, we know that instanton contributions for small values of the loop expansion parameter g behave like

$$Cg^{-b} e^{-a/g}. \quad (41.10)$$

The corresponding contribution to the perturbative coefficients for large-order k of the loop expansion is then,

$$(C/\pi)k^{b-1}a^kk!. \quad (41.11)$$

Therefore, the coefficients B_k of the Borel transform $B(z)$ (equation (41.6)) behave as

$$B_k \sim (C/\pi)k^{b-1}a^k. \quad (41.12)$$

This asymptotic estimate implies that the singularity of $B(z)$ closest to the origin is located at the point $z = 1/a$, and that $B(z)$ has an algebraic singularity of the form

$$\frac{C}{\pi} \int_0 \frac{dg e^{-a/g}}{g^{b+1}} \sum_k \frac{1}{k!} \left(\frac{z}{g}\right)^k = \frac{C}{\pi} \int_0 \frac{dg e^{-(a-z)/g}}{g^{b+1}} = (C/\pi)\Gamma(b)(a-z)^{-b}.$$

Therefore, the integral (41.7) does not exist if the classical action $A = 1/a$ is positive. The perturbation series in such theories is not Borel summable. In the light of this remark, we can draw several conclusions.

(i) The field equations have no real instanton solutions. In particular, this is the case if one has expanded around the unique absolute minimum of the potential. If complex instanton solutions exist, the corresponding classical action is non-positive, and the perturbative expansion is presumably Borel summable. This is only a presumption, because various features of the perturbative expansion, invisible at large orders, could prevent Borel summability. For instance, the perturbative expansion could contain contributions all of the same sign, growing faster than any exponential of the order k , but much smaller than $k!$ (e.g. $\sqrt{k!}$). Then, $B(z)$ would grow too rapidly for large-argument z ($\ln B(z) \sim z^2$ in the example) and the Borel integral would not converge at infinity.

(ii) One has expanded around a relative minimum of the potential, and real instanton solutions are found, corresponding to barrier penetration: the perturbative expansion is not Borel summable.

In this case, one additional piece of information may be useful for determining the solution, if the metastable situation originates from a stable situation by analytic continuation. Then, a possible solution is to integrate in the Borel transform just above the cut, which is on the real positive axis. Therefore, from a real perturbative expansion, one derives a complex result, but this is exactly what one expects. It is easy to verify that the imaginary part is, for g small, exactly what one would have calculated directly. Actually, this is only the solution of the problem in the simplest case, when no other instanton singularities cross the contour of integration in the analytic continuation.

(iii) Real instantons connect degenerate non-continuously connected classical minima (Section 40.1.3). The theory then is not Borel summable. Integration above or below the axis yields a complex result for a real quantity. The half sum of the integral above and below is real, but even in the simple example of the quartic double well-potential, one can verify that this does not give the correct answer. In Chapter 42, in the case of several one-dimensional potentials we show that, in addition to the perturbative expansion, one has to take into account multi-instanton contributions. The corresponding problem has not been solved in field theory examples yet.

Field theory examples. Field theory examples of degenerate non-continuously connected classical minima are provided by the two-dimensional $CP(N-1)$ models (Section 39.5), and four-dimensional $SU(2)$ gauge theory (Section 39.6). In these models, real instantons connect the degenerate minima, and the corresponding classical action is positive. Therefore, the perturbative expansion is not Borel summable. Note that the instanton contribution may not necessarily dominate the large-order behaviour, because, as the example of the $\phi_{d=4}^4$ massless field theory (Section 40.3.2) illustrates, when the classical field theory is scale invariant, the perturbative expansion might be dominated by contributions unobtainable by semi-classical methods and related to ultraviolet (UV) or infrared (IR) singularities.

41.2 Borel transformation: Series summation

All summation methods rely on some additional knowledge about the analytic properties of the function that is expanded.

In the framework of the $(\phi^2)^2$ field theory, methods based on the Borel transformation have been systematically used to sum perturbative series at fixed dimension $d < 4$, and to sum Wilson–Fisher’s $\varepsilon = 4 - d$ expansions. The main motivations are:

(i) Borel summability of the perturbative expansion at the fixed dimensions 2 and 3 has been rigorously established [422]. For the ε -expansion, Borel summability is quite plausible.

(ii) The information drawn from the large-order behaviour analysis (which has been determined in all cases and compares favourably with the first available terms of the series, see Section 40.2) can easily be incorporated.

The Borel transformation reduces the problem to determining the analytic continuation of the Borel transform, defined by its Taylor series in a circle, to a neighbourhood of the real positive axis. This continuation can be performed by many methods and the optimal choice depends somewhat on the available additional information. We give here two examples that have been used.

Padé approximants. In the absence of a precise knowledge of the location of the singularities of the Borel transform in the complex plane, one can use the Padé approximation [109, 418] (the Padé–Borel method). From the series, one derives $[M, N]$ Padé approximants, which are rational functions P_M/Q_N satisfying

$$B_f(z) = \frac{P_M(z)}{Q_N(z)} + O(z^{N+M+1}), \quad (41.13)$$

where P_M and Q_N are polynomials of degrees M and N , respectively. If one knows $(K+1)$ terms of the series, one can construct all Padé approximants with $N+M \leq K$. This method is well adapted to meromorphic functions. The main shortcoming of the method is that, for a rather broad class of functions, Padé approximants are known to converge only in measure [446], and thus spurious poles may occasionally appear close to, or on the real positive axis. Even when Padé approximants converge, this property may lead to instabilities in the results when M and N vary, and make an empirical evaluation of errors for short series difficult.

Borel transformation and conformal mapping. After Borel transformation, to sum the expansion of the Borel transform efficiently, it is necessary to know, or to guess, some of its analytic properties. Since in the case of the $(\phi^2)^2$ field theory and the perturbative expansion at fixed dimension, all known instanton actions are negative, it is plausible that all singularities of the Borel transform are located on the negative real axis.

The Borel transform is then analytic in a cut-plane, the location and nature of the singularity closest to the origin being given by the large order estimates (Section 41.1.2) [425].

For what concerns the ε -expansion, the situation is more subtle, because it is related to the four-dimensional perturbation theory, which has renormalon singularities that may even prevent Borel summability. However, one may argue, and this is supported by numerical evidence, that the ε -expansion of physical quantities is free of renormalon singularities.

In Refs. [424, 430], it has been assumed that the Borel transform is also analytic in a cut-plane. However, since in both cases this maximal analyticity is only a conjecture, the reliability of results can only be estimated by checking their stability with respect to reasonable variations of the summation method. Moreover, the comparison between the two families of results, perturbation theory at fixed dimensions $d < 4$, ε -expansion provides an internal consistency check of field theory methods.

Mapping of a cut-plane onto a disk. Analytic continuation can then be achieved by using a mapping that preserves the origin and maps the domain of analyticity onto a disk [423, 425]. In the transformed variable, the new series converges in the whole domain of analyticity. Let us explain the method with an example.

We assume that the function $B_f(z)$, Borel transform of a function f , is analytic in a cut-plane, the cut running along the real negative axis from $-\infty$ to $-1/a$. The mapping

$$z \mapsto u, \quad u(z) = \frac{\sqrt{1+az} - 1}{\sqrt{1+az} + 1} \Leftrightarrow z = \frac{4}{a} \frac{u}{(1-u)^2}, \quad (41.14)$$

maps the cut-plane onto a circle of radius 1. From the original series for the Borel transform, one derives a series in powers of the new variable u :

$$B_f(z) = \sum \frac{f_k}{k!} z^k \Rightarrow B_f[z(u)] = \sum_0^\infty B_k u^k. \quad (41.15)$$

Introducing this expansion into the Borel transformation, one obtains an expansion of $f(z)$ of the form

$$f(z) = \sum_0^\infty B_k I_k(z), \quad (41.16)$$

in which the functions $I_k(z)$ have the integral representation:

$$I_k(z) = \int_0^\infty e^{-t} [u(zt)]^k dt. \quad (41.17)$$

To determine the natural domain of convergence of the new expansion, one needs the behaviour of $I_k(z)$ for $k \rightarrow \infty$. Using for $u(z)$ the explicit expression (41.14), one can evaluate it by the steepest descent method. The saddle point equation is

$$-1 + \frac{k}{t} \frac{1}{\sqrt{1+azt}} = 0, \quad (41.18)$$

which, for k large, yields

$$t \sim k^{2/3}/(az)^{1/3}. \quad (41.19)$$

It follows that $I_k(z)$ behaves for k large as

$$I_k(z) \sim \exp \left[-3k^{2/3}/(az)^{1/3} \right]. \quad (41.20)$$

Three situations can then be encountered:

- (i) The coefficients B_k either decrease or at least do not grow too rapidly,

$$|B_k| < M e^{\varepsilon k^{2/3}}, \quad \forall \varepsilon > 0.$$

Then, the expansion (41.16) converges at least in the region

$$\operatorname{Re} z^{-1/3} > 0 \Rightarrow |\operatorname{Arg} z| < 3\pi/2. \quad (41.21)$$

This, in particular, implies that the function $f(z)$ must be analytic in the cut plane and even a part of the second sheet of the cut.

- (ii) The coefficients behave like

$$\ln |B_k| \sim ck^{2/3}, \quad c > 0, \quad \text{for } k \text{ large.} \quad (41.22)$$

The domain of convergence is

$$\operatorname{Re} z^{-1/3} > \frac{1}{3}ca^{1/3}. \quad (41.23)$$

This condition implies analyticity in a finite domain containing a part of the second sheet since for $|z|$ small, the right-hand side is negligible.

(iii) The coefficients B_k grow faster than $\exp(ck^{2/3})$. This is quite possible, since the only constraint on the coefficients B_k is that the series (41.15) has a radius of convergence 1. For instance, the coefficients B_k could grow like $\exp(ck^{4/5})$. In such a situation, the new series is also divergent. Such a situation arises when the singularities on the boundary of the domain of analyticity are too strong. One must map a smaller fraction of the domain of analyticity onto a disk.

41.3 Summing the perturbative expansion of the $(\phi^2)^2$ field theory

We discuss the summation of series by several variants of the method of Borel transformation and parametric mapping, whose simplest form is described in Section 41.2 [425, 430].

41.3.1 RG functions and exponents

Critical exponents and a number of other universal quantities have been calculated within the framework defined in Section 16.1 (equations (16.3, 16.4)), that is, in the massive $(\phi^2)^2$ field theory, as perturbative series at fixed dimension. For example, in three dimensions, in the normalization

$$\tilde{g} = (N + 8)g/(48\pi), \quad (41.24)$$

such that $\beta(\tilde{g}) = -\varepsilon\tilde{g} + \tilde{g}^2 + O(\tilde{g}^3)$, for $N = 1$, the RG β -function has the expansion [419],

$$\begin{aligned} \beta(\tilde{g}) = & -\tilde{g} + \tilde{g}^2 - \frac{308}{729}\tilde{g}^3 + 0.3510695978\tilde{g}^4 - 0.3765268283\tilde{g}^5 \\ & + 0.49554751\tilde{g}^6 - 0.749689\tilde{g}^7 + O(\tilde{g}^8). \end{aligned} \quad (41.25)$$

To determine exponents or other universal quantities, one must first determine the IR stable zero g^* of the function $\beta(g)$, which is given by a few terms of a divergent expansion.

Unlike the ε -expansion, there is no small parameter in which to expand since g^* has a value of order 1. Already at this stage a summation method is required.

A further problem arises from the property that RG functions, unlike the universal quantities in the ε -expansion, depend explicitly on the renormalization scheme. On the other hand, because one-loop diagrams have, in three dimensions, a simple analytic expression, it has been possible to calculate quite early the RG functions of the N -vector model up to six- and partially seven-loop order [418–420].

The main drawback of this procedure is that the values of the critical exponents depend strongly on the value of \tilde{g}^* . Therefore, an error in the estimation of \tilde{g}^* biases all exponents. A variant, which avoids this problem, has thus been used as a check. A pseudo- ε parameter has been introduced by setting (for $d = 3$)

$$\beta(\tilde{g}, \varepsilon) = \tilde{g}(1 - \varepsilon) + \beta(\tilde{g}). \quad (41.26)$$

The two functions $\beta(\tilde{g}, \varepsilon)$ and $\beta(\tilde{g})$ coincide for $\varepsilon = 1$, and the zero of $\beta(\tilde{g}, \varepsilon)$ is calculated as a power series in ε . Critical exponents are then also calculated as series in ε , and these series are summed. However, there are indications that the specific mapping $\tilde{g} \mapsto \varepsilon$ introduces singularities, because the apparent convergence is poorer. Nevertheless, all variants give consistent results, which is satisfactory. Moreover, a comparison of the different results gives useful indications about the apparent errors.

41.3.2 Equation of state

The equation of state for $N = 1$, as well a number of universal ratios of amplitude, have also been calculated [444]. Numerical results are displayed in Section 41.8.

In the framework of the massive theory, the calculation of physical quantities in the ordered phase leads to additional technical problems because the theory is parametrized in terms of the disordered phase correlation length m^{-1} , which is singular at T_c . Moreover, the normalization of correlation functions is singular at T_c (equation (16.17)).

In the example of Ising-like systems ($N = 1$), the free-energy F has the general form

$$F(M) - F(0) = \frac{m^d}{g} \varphi(g^{1/2} m^{1-d/2} \tilde{M}, g),$$

in which g has to be set to its fixed point value g^* , and \tilde{M} is related to the magnetization M by the field renormalization (16.15):

$$\tilde{M} = M m^{-\eta/2}.$$

The derivative of F with respect to M yields the equation of state under the form

$$H = g^{-1/2} m^{1+d/2-\eta/2} \mathfrak{h}(g^{1/2} m^{1-d/2} \tilde{M}, g), \quad \mathfrak{h}(z, g) = \partial \varphi(z, g) / \partial z. \quad (41.27)$$

For example, at one-loop order, the function $\mathfrak{h}(z, g)$ is given by

$$\begin{aligned} \mathfrak{h}(z, g) &= z + \frac{z^3}{6} + \frac{gz}{2} \frac{1}{(2\pi)^d} \int \frac{d^d p}{p^2 + 1 + z^2/2} + O(2 \text{ loops}) \\ &= z + \frac{z^3}{6} + \frac{\pi N_d}{4 \sin \pi d/2} g z \left[(1 + z^2/2)^{d/2-1} - 1 - \frac{1}{4}(d-2)z^2 \right]. \end{aligned} \quad (41.28)$$

In terms of the deviation from the critical temperature $t = m^{1/\nu}(g^*)^{1/2\beta} \sim T - T_c$, equation (41.27) takes the form

$$H = H_0 t^{\beta\delta} \mathfrak{h}(Mt^{-\beta}) \quad (H_0 \text{ being a constant}). \quad (41.29)$$

The expression is adequate for the description of the disordered phase when $Mt^{-\beta}$ is small, but all terms in the loop expansion become singular for $t \rightarrow 0$.

The ordered phase. This problem does not completely prevent calculations near the coexistence curve, that is, for $t < 0$. Since, at the fixed point g^* , all functions have simple power law singularities at T_c , it is possible to proceed by analytic continuation in the complex t -plane. The scaling variable $z = Mt^{-\beta}$ picks up a phase below T_c :

$$\text{for } t = |t| e^{i\pi}, \quad z = |z| e^{-i\pi\beta}. \quad (41.30)$$

The scaling variable $H(-t)^{-\beta\delta}$ then is given by

$$H(-t)^{-\beta\delta} = H_0 e^{i\pi\beta\delta} \mathfrak{h}(z) = H_0 |\mathfrak{h}(z)|. \quad (41.31)$$

In particular, it is possible to evaluate ratios of amplitudes of singularities above and below T_c ; one can calculate the complex zero of $\mathfrak{h}(z, g)$ as a power series in g and substitute it in other quantities. The result is complex, but its modulus taken at $g = g^*$ converges towards the correct result. This can be illustrated by the example of the magnetic susceptibility. From equation (41.27), one derives

$$C^+ / C^- = e^{-i\pi\gamma} \mathfrak{h}'(z_0(g^*), g^*) / \mathfrak{h}'(0, g^*) = |\mathfrak{h}'(z_0(g^*), g^*)|, \quad (41.32)$$

in which z_0 is the complex zero of $\mathfrak{h}(z, g)$. The expression can be used to obtain a series expansion for C^+ / C^- .

However, this method does not make an extrapolation of the equation of state to t small possible. Following the lines of Section 16.9.3, it is natural to introduce the parametric representation (16.147) [163],

$$z = x_0^{-\beta} \theta (1 - \theta^2)^{-\beta},$$

and consider the function

$$h(\theta) = (1 - \theta^2)^{\beta\delta} \mathfrak{h}(z(\theta)).$$

However, in an expansion at fixed dimension, if we just replace all quantities by their perturbative expansion, the singularity of $h(\theta)$ at $\theta = 1$ (*i.e.* $t = 0$) does not cancel any more. Therefore, inspired by results coming from the ε -expansion, one also expands $h(\theta)$ in powers of θ . The method is the following: one first determines by Borel summation, as explained in Section 41.4, the first terms of the expansion of the function $\mathfrak{h}(z)$ in powers of z . As expected, the apparent precision decreases with increasing degree. One determines the corresponding coefficients of the expansion of $h(\theta)$ in powers of θ (*note* $z \sim \theta$). These coefficients are polynomials in $x_0^{-\beta}$. One then adjusts the arbitrary constant x_0 to minimize the last term, following the ODM method [447] (see Section A41.1). This strategy has been applied to the $N = 1$ series, which are known up to order g^5 . A general representation of the equation of state has been obtained. Some amplitude ratios have been inferred (see Table 41.6). It would be interesting to apply this method to $N \neq 1$ series.

Table 41.1

Series summed by the method based on Borel transformation and mapping for the zero \tilde{g}^ of the RG- β function, and the exponents γ and ν in the $\phi_{d=3}^4$ field theory, after determination of \tilde{g}^* .*

k	2	3	4	5	6	7
\tilde{g}^*	1.8774	1.5135	1.4149	1.4107	1.4103	1.4105
ν	0.6338	0.6328	0.6297	0.6302	0.6302	0.6302
γ	1.2257	1.2370	1.2386	1.2398	1.2398	1.2398

41.4 Summation method: Practical implementation

The general principles and the theoretical justification of the summation method based on Borel summation and conformal mapping are explained in Section 41.2. We add here a few details about its specific implementation for the calculation of critical exponents and other universal quantities. In Table 41.1, we display three examples of transformed series, to illustrate the convergence.

The method. Several different variants based on the Borel (really Borel–Leroy) transformation and parametric conformal mapping have been implemented and tested. Let $R(z)$ be the function whose expansion has to be summed (z here represents the coupling constant \tilde{g} or ε):

$$R(z) = \sum_{k=0}^{\infty} R_k z^k. \quad (41.33)$$

One transforms the series into

$$R(z) = \sum_{k=0}^{\infty} B_k(\rho) \int_0^{\infty} t^{\rho} e^{-t} [u(zt)]^k dt, \quad \rho \geq 0, \quad \text{with } u(z) = \frac{\sqrt{1+az}-1}{\sqrt{1+za}+1}. \quad (41.34)$$

The coefficients B_k are calculated by identifying the expansion of the right-hand side of equation (41.34) in powers of z with the expansion (41.33). The constant a has been determined by the large-order behaviour analysis (equations (38.11, 38.12)). At fixed dimension (with a normalization such that $\beta(\tilde{g}) = -\varepsilon\tilde{g} + \tilde{g}^2 + O(\tilde{g}^3)$, see equation (41.24)),

$$a(d=3) = \frac{48\pi}{(N+8)A} = 0.147774232 \times \frac{9}{(N+8)}, \quad (41.35)$$

$$a(d=2, N=1) = \frac{8\pi}{3A} = 0.238659217, \quad (41.36)$$

and for the ε -expansion, $a = 3/(N+8)$.

The free-parameter ρ is adjusted empirically to improve the convergence of the transformed series by weakening the singularities of the Borel transform near $z = -a$. Moreover, in many cases, the conformal transformation

$$R(z) \mapsto \tilde{R}(z) = R[z/(1-\tau z)], \quad (41.37)$$

has been applied to the initial function R , where τ is left as an adjustable parameter in order to move away its closest singularities, because the location of all singularities of $R(z)$ is not known.

This transformation is necessary in the case of the ε -expansion, because the critical exponents, as functions of ε , have close singularities. It has been verified that it also improves the summation of the series at fixed dimensions.

Finally, in general, a third parameter was introduced which will not be discussed here.

Needless to say, with three parameters and short-initial series, it becomes possible to find, occasionally, some transformed series whose apparent convergence is deceptively good. Therefore, it is essential to vary the parameters in some range around the optimal values, and to examine the sensitivity of the results upon their variations. Finally, it is useful to sum independently series for exponents related by scaling relations. An underestimation of the apparent errors leads to inconsistent results. It is clear from these remarks that the quoted errors remain an educated guess, even though, compared to the central values, their determination requires much more work.

The $(\phi^2)^2$ field theory at fixed dimensions. The RG β -function has been determined, in three dimensions, up to six-loop order while the series for the dimensions of the fields ϕ and ϕ^2 have been extended to seven loops. In two dimensions, the series only are known up to four loops. They have been analysed by two methods. In the first method, the series of the RG β -function has been first summed and its zero \tilde{g}^* calculated ($\tilde{g} = g(N+8)/(48\pi)$ for $d = 3$, $\tilde{g} = 3g/8\pi$ for $d = 2$). The series of the other RG functions have then been summed for $\tilde{g} = \tilde{g}^*$.

The ε -expansion. The ε -expansion makes it possible to connect the results in three and two dimensions, a notable advantage. In particular, in the cases $N = 1$ and $N = 0$, it is possible to compare the ϕ^4 results with exact results coming from lattice models and to test both universality and the reliability of the summation procedure. Moreover, it is possible to constrain the three-dimensional results by imposing the exact two-dimensional values or the behaviour near two dimensions for $N > 1$, the changes in the results being a check of consistency [430]. Finally, as we have already emphasized, the comparison between the different results is a check of the consistency of the field theory methods, combined with the summation procedures.

Table 41.2

Critical exponents from the summed perturbative $(\phi^2)^2$ field theory for $N = 0, 1$ in dimension 2.

	γ	ν	η	β	ω
$N = 0, \varepsilon^5$	1.39 ± 0.04	0.76 ± 0.03	0.21 ± 0.05	0.065 ± 0.015	1.7 ± 0.2
$N = 0, \varepsilon^6$	1.333 ± 0.025	0.741 ± 0.004	0.201 ± 0.025	0.074 ± 0.010	1.90 ± 0.25
$N = 0$, exact	1.34375	0.75	0.2083 \dots	0.0781 \dots	?
$N = 1, \varepsilon^5$	1.73 ± 0.06	0.99 ± 0.04	0.26 ± 0.05	0.120 ± 0.015	1.6 ± 0.2
$N = 1, \varepsilon^6$	1.68 ± 0.05	0.952 ± 0.014	0.237 ± 0.027	0.113 ± 0.015	1.71 ± 0.09
$N = 1, d = 2$ fixed	1.79 ± 0.04	0.96 ± 0.04	0.18 ± 0.04	0.086 ± 0.022	1.3 ± 0.2
$N = 1$, exact	1.75	1.	0.25	0.125	?

Table 41.3

Estimates of critical exponents in the $O(N)$ -symmetric $(\phi^2)_{d=3}^2$ field theory.

N	0	1	2	3
\tilde{g}^*	1.413 ± 0.006	1.411 ± 0.004	1.403 ± 0.003	1.390 ± 0.004
g^*	26.63 ± 0.11	23.64 ± 0.07	21.16 ± 0.05	19.06 ± 0.05
γ	1.1596 ± 0.0020	1.2396 ± 0.0013	1.3169 ± 0.0020	1.3895 ± 0.0050
ν	0.5882 ± 0.0011	0.6304 ± 0.0013	0.6703 ± 0.0015	0.7073 ± 0.0035
η	0.0284 ± 0.0025	0.0335 ± 0.0025	0.0354 ± 0.0025	0.0355 ± 0.0025
β	0.3024 ± 0.0008	0.3258 ± 0.0014	0.3470 ± 0.0016	0.3662 ± 0.0025
α	0.235 ± 0.003	0.109 ± 0.004	-0.011 ± 0.004	-0.122 ± 0.010
ω	0.812 ± 0.016	0.799 ± 0.011	0.789 ± 0.011	0.782 ± 0.0013
$\theta = \omega\nu$	0.478 ± 0.010	0.504 ± 0.008	0.529 ± 0.009	0.553 ± 0.012

41.5 Field theory estimates of critical exponents for the $O(N)$ model

We now display the values of critical exponents inferred from an RG analysis of the $O(N)$ -symmetric $(\phi^2)^2$ field theory, based on summed perturbative expansions. Let us point out that, even if we focus here mainly on the determination of critical exponents, the advantages of field theory methods are that universality can be proven, and that all universal quantities can, in principle, be calculated. We give the examples of the equation of state, and amplitude ratios in Section 41.8.

41.5.1 Dimension 2

In Table 41.2, we display the results for $N = 0$ and $N = 1$ coming from the summed ε -expansion for $\varepsilon = 2$, including the early results at order ε^5 [424] and the more recent results at order ε^6 [160] (note that the values of γ and β are inferred by scaling), and for $N = 1$ from the perturbative expansion at fixed dimension [425]. With the latter method, the apparent errors are larger because the series are shorter. For $N = 1$, we compare the values with the Ising model exponents and, for $N = 0$, with the values proposed in Ref. [426]. The values for $N = 0$ are consistent with a convergence towards the exact values. For $N = 1$, the values of the exponents are close to the exact values but the convergence of the summed ε expansion seems to be slower. For the less precise results inferred from the fixed dimension series, the main observation is that the exponent η is underestimated.

For $N = 1$, one obtains a fixed point coupling constant $\tilde{g}^* = 1.85 \pm 0.010$. In lattice models, it is known only from high-temperature series [427–429]: $\tilde{g}^* = 1.755(2)$.

Both for $N = 0$ and $N = 1$, the identification of the correction exponent ω remains a problem. For example, only analytic corrections to scaling have been found in the Ising model, which makes the identification of the correction exponent ω difficult. However, an analysis based on conformal invariance predicts a correction exponent $\omega = 4/m$ for ϕ^{2m-2} field theories and for $m > 3$. One may conjecture that the amplitudes of the singularities involving the correction exponent ω vanish when m approaches 3 for $d = 2$, or for $m = 3$ when d approaches 2.

Table 41.4

Critical exponents in the $(\phi^2)_{d=3}^2$ field theory derived from the ε -expansion.

N	0	1	2	3
$\gamma(\varepsilon^5)$	1.1575 ± 0.0060	1.2355 ± 0.0050	1.311 ± 0.007	1.382 ± 0.009
$\gamma(\varepsilon^6)$	1.1566 ± 0.0010	1.2356 ± 0.0014	1.313 ± 0.002	1.385 ± 0.004
$\nu(\varepsilon^5)$	0.5875 ± 0.0025	0.6290 ± 0.0025	0.6680 ± 0.0035	0.7045 ± 0.0055
$\nu(\varepsilon^6)$	0.5874 ± 0.0003	0.6292 ± 0.0005	0.6690 ± 0.0010	0.7059 ± 0.0020
$\eta(\varepsilon^5)$	0.0300 ± 0.0050	0.0360 ± 0.0050	0.0380 ± 0.0050	0.0375 ± 0.0045
$\eta(\varepsilon^6)$	0.0310 ± 0.0007	0.0362 ± 0.0006	0.0380 ± 0.0006	0.0378 ± 0.0005
$\beta(\varepsilon^5)$	0.3025 ± 0.0025	0.3257 ± 0.0025	0.3465 ± 0.0035	0.3655 ± 0.0035
$\beta(\varepsilon^6)$	0.3028 ± 0.0004	0.3260 ± 0.0004	0.3472 ± 0.0007	0.3663 ± 0.0012
$\omega(\varepsilon^5)$	0.828 ± 0.023	0.814 ± 0.018	0.802 ± 0.018	0.794 ± 0.018
$\omega(\varepsilon^6)$	0.841 ± 0.013	0.820 ± 0.007	0.804 ± 0.003	0.795 ± 0.007
$\theta = \omega\nu(\varepsilon^5)$	0.486 ± 0.016	0.512 ± 0.013	0.536 ± 0.015	0.559 ± 0.017
$\theta = \omega\nu(\varepsilon^6)$	0.494 ± 0.009	0.516 ± 0.008	0.538 ± 0.003	0.561 ± 0.012

41.5.2 Dimension 3

Table 41.3 displays the results obtained from summed perturbation series at fixed dimension 3 [430], incorporating seven-loop terms for γ and η reported in Ref. [419]. The exponent $\theta = \omega\nu$ characterizes corrections to scaling in the temperature variable,

Table 41.4 displays the results for $\varepsilon = 1$ from the ε expansions, at order ε^5 [430] and the more recent results, which take into account the ε^6 contributions (θ is inferred from the reported values of ν and ω) [160].

Discussion. Since 1980, the fixed-dimension, three-dimensional results have only marginally changed [425, 430], and within errors. No additional progress can be expected in the near future, since a seven-loop calculation of the RG β function (a formidable task) is still lacking. By contrast, the ε -expansion predictions have steadily improved [424, 430, 160] because the order ε^5 (in two steps) and more recently, the order ε^6 have become available. It should be noted that the successive values are very close, suggesting that the errors in Ref. [430] were too conservative.

Considering the reasonable agreement between exact two-dimensional results and summed ε expansion, in Ref. [430], as an alternative method, a summation procedure has been used that automatically incorporates the $d = 2, \varepsilon = 2$ values. The three-dimensional results have slightly changed, but within errors.

Comparing the two sets of results coming from the perturbation series at fixed dimension, and the ε -expansion, one notes that the general agreement is excellent, especially for the exponents ν and β , even if some tension exists for η . A comparison with the most precise results provided by other methods for $N = 0$ and $N = 1$ favours the ε expansion.

41.6 Other three-dimensional theoretical estimates

The N -vector model, with nearest-neighbour interactions, has been studied on several lattices. Critical exponents are derived from an analysis of high-temperature series by different ratio methods, Padé or differential approximants (Sections 41.2 and A41.2) [431].

Table 41.5

Estimates of critical exponents of the $(\phi^2)_{d=3}^2$ field theory, by other theoretical methods.

N	0	1	2	3
γ	1.159653(1)	1.237075	1.3178(2)	1.3963(22)
ν	0.587597	0.629971	0.6717(1)	0.7116(10)
η	0.031043(3)	0.036298(2)	0.0381(2)	0.0378(3)
β	0.302919	0.326419	0.3486(1)	0.3692(5)
ω	0.900(15)	0.830(2)	0.811(10)	0.791(22)
$\theta = \omega\nu$	0.528(8)	0.523(2)	0.545(6)	0.563(13)

A number of results are obtained from computer simulations using Monte-Carlo methods [432, 433]. More recently, new results have been provided by the *conformal bootstrap method* [91, 434, 435].

Note that, when the first field theory results were published [425], they were in strong disagreement with some of high-temperature series analyses. With time, high-temperature and Monte-Carlo predictions have converged towards the field theory results, and now all available results are remarkably consistent, as Table 41.5 indicates.

However, in the $N = 0$ Monte-Carlo results [433], one notes one discrepancy for the correction exponent ω .

$N = 1$ is now dominated by the conformal bootstrap results [435] and the fixed-dimension result for ω seems now to be somewhat low.

Estimates for $N = 2$ and $N = 3$, have been obtained by using a combination of Monte-Carlo simulations and analysis of high-temperature series [432].

41.7 Critical exponents from experiments

We have discussed the N -vector model in the ferromagnetic language, even though most of our experimental knowledge comes from physical systems which are not magnetic but belong to the universality class of the N -vector model. For example, $N = 0$ describes the statistical properties of long polymer chains, that is, long not intersecting chains or self-avoiding walks (see Section 15.8). The case $N = 1$ (Ising-like systems) describes liquid-vapour transitions in classical fluids, critical binary fluids and uniaxial antiferromagnets. The helium superfluid transition corresponds to $N = 2$. Finally, for $N = 3$, the experimental information comes from ferromagnetic systems.

An early review where experiments, high-temperature series and field-theory results are compared, using the results then available, can be found in the proceedings of the Cargèse summer school 1980 [436].

Critical exponents and polymers. In the case of polymers, only the exponent ν is easily accessible. An old, not reproduced since, result is $\nu = 0.586 \pm 0.004$ [437], in excellent agreement with theory.

Ising-like systems $N = 1$. For Ising-like systems, the results are somewhat scattered and difficult to review. Earlier reviews for fluids containing, in addition, estimates of universal ratios are Refs. [438]. Ref. [439] reports the values $\alpha = 0.0.113 \pm 0.005$ and for the correction exponent θ (sometimes denoted Δ), the estimate $\theta = 0.50 \pm 0.3$.

In the more recent review of Sengers and Shanks [440], one finds $\eta = 0.032 \pm 0.013$, $\nu = 0.629 \pm 0.003$.

However, the authors also quote the more precise value $\eta = 0.030 \pm 0.0015$ [441] and $\nu = 0.632 \pm 0.002$, $\eta = 0.041 \pm 0.005$ [442].

The conclusion is that the exponents derived from the analysis of fluid experiments are globally totally consistent with theoretical values.

Helium superfluid transition, N = 2. The helium fluid makes extremely precise measurements very close to T_c possible, and this explains the precision of the determination of critical exponents. However, the order parameter is not directly accessible and, therefore, only ν and α have been determined. Reported values are [143]

$$\nu = 0.6705 \pm 0.0006, \quad \nu = 0.6708 \pm 0.0004, \quad \text{and} \quad \alpha = -0.01285 \pm 0.00038.$$

The agreement with RG values is quite remarkable, but the precision of ν is now a challenge to field theory.

Ferromagnetic systems, N = 3. The precision of experimental results is limited. One finds in the literature the estimates

$$\gamma = 1.40 \pm 0.03, \quad \nu \in [0.700 - 0.725], \quad \beta = 0.35 \pm 0.03, \quad \alpha \in [-0.09 - -0.012], \quad \theta = 0.54 \pm 0.10.$$

Table 41.6
Amplitude ratios: Models and binary critical fluids ($N = 1$).

	ε -expansion	Fixed dim. $d = 3$	Lattice models	Experiment
A^+/A^-	0.527 ± 0.037	0.537 ± 0.019	$\begin{cases} 0.523 \pm 0.009 \\ 0.560 \pm 0.010 \end{cases}$	0.56 ± 0.02
C^+/C^-	4.73 ± 0.16	4.79 ± 0.10	$\begin{cases} 4.75 \pm 0.03 \\ 4.95 \pm 0.15 \end{cases}$	4.3 ± 0.3
f_1^+/f_1^-	1.91	2.04 ± 0.04	1.96 ± 0.01	1.9 ± 0.2
R_ξ^+	0.28	0.270 ± 0.001	0.266 ± 0.001	$0.25 - 0.32$
R_c	0.0569 ± 0.0035	0.0574 ± 0.0020	0.0581 ± 0.0010	0.050 ± 0.015
$R_\xi^+ R_c^{-1/3}$	0.73	0.700 ± 0.014	0.650	$0.60 - 0.80$
R_χ	1.648 ± 0.036	1.669 ± 0.018	1.75	1.75 ± 0.30
Q_2	1.13		1.21 ± 0.04	1.1 ± 0.3
Q_3	0.96		0.896 ± 0.005	

Table 41.7
Correction amplitude ratios for $N = 1$.

	ε – expansion	Fixed dim. $d = 3$	HT series	Experiment
a_ξ^+/a_χ^+	0.56 ± 0.15	0.65 ± 0.05	0.70 ± 0.03	
a_C^+/a_ξ^+	2.03	1.45 ± 0.11		
a_C^+/a_χ^+	1.02	0.94 ± 0.10	1.96	0.87 ± 0.13
a_C^+/a_C^-	2.54	1.0 ± 0.1		$0.7 - 1.35$
a_χ^+/a_χ^-	0.3	0.315 ± 0.013		
a_C^+/a_M		1.10 ± 0.25		1.85 ± 0.10
a_M/a_χ^+		0.90 ± 0.21		$0.08 - 1.4$

41.8 Amplitude ratios

A classical review on amplitude ratios can be found in Ref. [443].

Amplitude ratios evaluated by field theory and RG methods (see Section 16.9.4 for definitions) are less precise than exponents, because the series are shorter. Note that a determination of the equation of state based on field theory methods (Section 41.3.2) is also available Ref. [444].

Table 41.6 contains a comparison of amplitude ratios as obtained, for $N = 1$, from field theory, lattice calculations for Ising-like models, and experiments on binary mixtures [445].

Some results are available for uniaxial magnetic systems and liquid–vapour transitions. For the latter systems, a few reported values are

$$C^+/C^- = 5. \pm 0.2 , \quad R_c = 0.047 \pm 0.010 , \quad R_\chi = 1.69 \pm 0.14 .$$

For A^+/A^- , results range from 0.48 to 0.53. The set of results, with indeed large errors, shows a satisfactory agreement with the field theory based RG predictions.

Finally, we give a few results concerning ratios of amplitudes of corrections to the leading scaling behaviour. If, in addition to the correlation length and the susceptibility amplitudes a_ξ and a_χ , one considers also the specific heat amplitude a_C , and the coexistence curve magnetization amplitude a_M , one can form three independent ratios. The results for $N = 1$ are displayed in Table 41.7.

Table 41.8
Amplitude ratios for $N = 2$ and $N = 3$.

	N	Field theory	HT series	Experiment
A^+/A^-	2	1.056 ± 0.004	1.08	$1.054 \pm .001$
R_ξ^+	2	0.36	0.36	
R_c	2	0.123 ± 0.003		
A^+/A^-	3	1.52 ± 0.02	1.52	1.40–1.52
R_ξ^+	3	0.42	0.42	0.45
R_c	3	0.189 ± 0.009	0.165	

A few amplitude ratios have been calculated and measured for helium ($N = 2$) and ferromagnets ($N = 3$). We give in Table 41.8 the examples of A^+/A^- , R_ξ^+ and R_c .

A review of all available data (critical exponents, amplitude ratios, and so on), clearly shows that the RG predictions are remarkably consistent with the whole experimental information available. The diversity of experimental systems is a spectacular confirmation of the concept of universality, and of the RG ideas combined with field methods.

A41 Some other summation methods

For illustration purposes, we give two examples of summation methods that do not necessarily involve a Borel transformation.

A41.1 Order-dependent mapping method (ODM)

The ODM method requires, to be applicable, some knowledge of the analyticity properties of the function itself [447]. As we have discussed, functions may be singular at an expansion point and nevertheless have a (necessarily divergent) series expansion in a sector. In the examples we have met, the function has a cut extending to the origin, and its discontinuity decreases exponentially, close to the origin. The idea of the ODM method is then to pretend that the function is analytic, in addition to its true domain of analyticity, in a small disk centred at the origin, of adjustable radius ρ , and to map this extended domain onto a circle centred at the origin, keeping the origin fixed. If the function would really be analytic in such a domain, the expansion in the transformed variable would converge in the whole domain of analyticity, and the continuation problem would be solved. Since the original series is really only asymptotic, the series in the transformed variable is also asymptotic. However, as a result of the transformation, the coefficients of the new series now depend on an adjustable parameter ρ .

For instance, let us assume that $f(z)$ is analytic in a cut-plane. We then use the mapping,

$$z = 4\rho u / (1 - u)^2. \quad (\text{A41.1})$$

The transformed series has the form

$$f(z(u)) = \sum_0^{\infty} P_k(\rho) u^k, \quad (\text{A41.2})$$

in which the coefficients $P_k(\rho)$ are polynomials of degree k in the parameter ρ . In more general situations, one can often use a mapping of the form

$$z = \rho h(u), \quad h(u) = O(u). \quad (\text{A41.3})$$

The k th order approximation is obtained by truncating the series at order k and choosing ρ as one of the zeros of the polynomial $P_k(\rho)$. The zero cannot actually be chosen arbitrarily but qualitatively speaking, must be the zero of largest modulus for which the derivative $P'_k(\rho)$ is small. The idea behind the method is the following: with the original series, the best approximation is obtained by truncating the series at z fixed, at an order dependent on z such that the modulus of the last term taken into account is minimal. By introducing an additional parameter, one modifies the situation: one first chooses the order of truncation, and then tries to adjust the parameter ρ in such a way that, at z again fixed, the last term taken into account is minimal.

The k th order approximant has the form

$$\{f(z)\}_k = \sum_{l=0}^k P_l(\rho_k) [u(z)]^l, \quad P_k(\rho_k) = 0. \quad (\text{A41.4})$$

Under some conditions, it can be shown that if the terms f_k of the original series grow, for k large, like $(k!)^\beta$, then the sequence ρ_k decreases like $1/k^\beta$. Proof of the convergence of the method in some cases can be found in Ref. [448].

The method has been successfully applied to test problems like the quartic anharmonic oscillator with a mapping

$$z = \rho u / (1 - u)^{3/2},$$

the complex PT symmetric potential $x^2 + igx^3$ [449], and to one physics relevant example, the hydrogen atom in a strong magnetic field [450].

A41.2 Linear differential approximants

Padé approximants [109] provide the simplest example of a general class of approximants, which are obtained as solutions to equations (algebraic or differential) with polynomial coefficients [451]. These polynomials are chosen to be the polynomials of the lowest degree for which the solution of the equation has the same power series expansion, up to a given order, as the function one wants to approximate. To be more concrete, we describe linear differential approximants.

Let $f(z)$ be a function for which we know a power series expansion. We can construct approximants $\bar{f}_k(z)$ to this function by looking for solutions of the differential equation

$$\sum_{n=0}^N P_n(z) \left(\frac{\partial}{\partial z} \right)^n \bar{f}_k(z) = R(z), \quad (A41.5)$$

in which the polynomials $P_n(z)$ and $R(z)$ form a set of polynomials of the lowest possible total degree such that

$$f(z) - \bar{f}_k(z) = O(z^{k+1}). \quad (A41.6)$$

In the generic situation, the degrees $[P_n]$ and $[R]$ of the polynomials P_n and R satisfy

$$\sum_{n=0}^N [P_n] + [R] = k + 1. \quad (A41.7)$$

The advantage of these kinds of approximants is that they are extremely flexible. It is possible to use much available information about the function by imposing additional constraints on the polynomials P_n and R .

Furthermore, while Padé approximants generate only approximants with poles, the more general approximants can have a large class of new singularities. The drawback of this flexibility is that this leads to approximations that are much more unstable. It is necessary to select among the large number of approximants one can construct, those for which one has reasons to believe that they are best adapted to the function one wants to approximate.

Due to the generality of the problem, a systematic study of this class of approximants is lacking. Finally, the method can be generalized to power series in several variables. One then derives partial differential equations with polynomial coefficients in all variables.

42 Multi-instantons in quantum mechanics (QM)

In general, a linear combination of instanton solutions is not a solution of the imaginary-time equations of motion, because the equations are not linear. Moreover, in QM, all solutions of the classical equations can depend only on one time collective coordinate (in this respect, in field theory, the situation is different, see Sections 39.5, [399]). However, a linear combination of largely separated instantons (a *multi-instanton* configuration) renders the action almost stationary, because each instanton solution differs from a constant solution by only exponentially small corrections at large distances (in field theory this is only true if the theory is massive). In the context of QM, we examine the possible contributions of such multi-instanton configurations. However, the generalization to quantum field theory (QFT) is not trivial, and is, to a large extent, still to be worked out, even if some progress has been reported [452].

In several simple situations, multi-instantons are expected to play a role. In the case in which instantons are found, when calculating the contribution to $\text{tr } e^{-\beta H}$ at finite β , we have always kept only the solution that describes the classical trajectory once. We have argued that the other solutions, in which the trajectory is described n -times, have in the large β limit an action n -times larger, and, therefore, give subleading contributions to the path integral. In the large β limit, these configurations have the properties one expects from n -instantons. However, there is a subtlety: naively, one expects these configurations to give contributions of order β , because a classical solution depends only on one time parameter. On the other hand, the ground state energy E_0 has an expansion of the form

$$E_0 = E_0^{(0)} + E_0^{(1)} + \dots,$$

in which $E_0^{(0)}$ and $E_0^{(1)}$ are the perturbative and one-instanton contributions, respectively, and the dots represent possible multi-instanton contributions. The ground state energy has been derived from a semi-classical calculation, at β large, of the partition function, which then has the form

$$\text{tr } e^{-\beta H} \sim e^{-\beta E_0} \sim e^{-\beta E_0^{(0)}} \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \left(E_0^{(1)} \right)^n. \quad (42.1)$$

Thus, the existence of a one-instanton contribution to the energy implies the existence of n -instanton contributions to the partition function proportional to β^n instead of β .

Another example is provided by large-order behaviour estimates of perturbation theory for potentials with degenerate minima (Section 40.1.3). When one starts from a situation in which the minima are almost degenerate, one obtains, in the degenerate limit, a contribution that has the interpretation of the contribution of two, infinitely separated, instantons, multiplied by an infinite multiplicative coefficient. The divergence has the following interpretation: in the degenerate limit, the fluctuations that correspond to changing the distance between largely separated instanton and anti-instanton, induce a vanishingly small variation of the action. It follows that, to correctly calculate the limit, one has to introduce a second collective coordinate which describes these fluctuations, although there is no corresponding symmetry of the action.

It can then also be understood where, in the first example, the factor β^n comes from. Although a given classical trajectory can only generate a factor β , these new configurations depend on n independent collective coordinates over which one has to integrate.

To summarize, multi-instanton contributions do exist. However, multi-instantons are not solutions of the classical equation of motion. They correspond to configurations of largely separated instantons connected in a way that has to be examined. They become solutions of the equation of motion only asymptotically, in the limit of infinite separation, and depend on n times more collective coordinates than the one-instanton configuration.

In Sections 42.1 and 42.2, we first return to two examples that we have already discussed in Chapter 39: the quartic *double-well* potential, and the *periodic cosine* potential. We then discuss general potentials with degenerate minima. We also calculate the large order behaviour in the case of the $O(\nu)$ -symmetric anharmonic oscillator.

The determination, at leading order, of the many-instanton contributions has led to conjecture the exact form of the semi-classical expansion for potentials with degenerate minima, generalizing the exact Bohr–Sommerfeld quantization condition. Although a rigorous analysis of the problem relies on Wentzel–Kramers–Brillouin (WKB) exact methods and resurgence theory, the original multi-instanton viewpoint remains quite intuitive, and could still be useful in QFT, which cannot be described in terms of differential equations.

The appendix contains some technical remarks about the calculation of multi-instanton contributions, a simple example of a non Borel-summable expansion, and a discussion of the generalized Bohr–Sommerfeld quantization condition within the framework of Schrödinger’s equation and WKB approximation.

42.1 The quartic double-well potential

We first consider the Hamiltonian of the double-well potential (see Section 39.1 for details),

$$H = -\frac{1}{2} (\mathrm{d}/\mathrm{d}q)^2 + V(q\sqrt{g})/g, \quad \text{with} \quad V(q) = \frac{1}{2}q^2(1-q)^2, \quad g > 0. \quad (42.2)$$

We derive properties of the spectrum by considering both the partition function $\mathrm{tr} e^{-\beta H}$ and a twisted partition function $P \mathrm{tr} e^{-\beta H}$, where P is the reflection operator corresponding to the symmetry $q \mapsto g^{-1/2} - q$ of the Hamiltonian.

The partition function is given by the path integral (equation (39.8))

$$\mathcal{Z}(\beta) = \int [\mathrm{d}q(t)] \exp [-\mathcal{S}(q)], \quad \text{with} \quad \mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\frac{1}{2}\dot{q}^2(t) + V(q(t)\sqrt{g})/g \right] \mathrm{d}t, \quad (42.3)$$

and the paths satisfy periodic boundary conditions: $q(-\beta/2) = q(\beta/2)$. The path integral representation of the twisted partition function $\mathrm{tr} P e^{-\beta H}$ differs by the boundary conditions, which are $q(-\beta/2) + q(\beta/2) = g^{-1/2}$.

In the infinite β limit, the instanton solutions are (equation (39.11)),

$$q_c(t) = f(\mp(t - t_0))/\sqrt{g}, \quad \text{with} \quad f(t) = 1/(1 + e^{-t}) = 1 - f(-t), \quad (42.4)$$

where the constant t_0 characterizes the instanton position, and the corresponding classical action is (equation (39.12))

$$\mathcal{S}(q_c) = \frac{1}{g} \int \mathrm{d}t \left[\frac{1}{2}\dot{f}^2(t) + V(f(t)) \right] = \frac{1}{6g}. \quad (42.5)$$

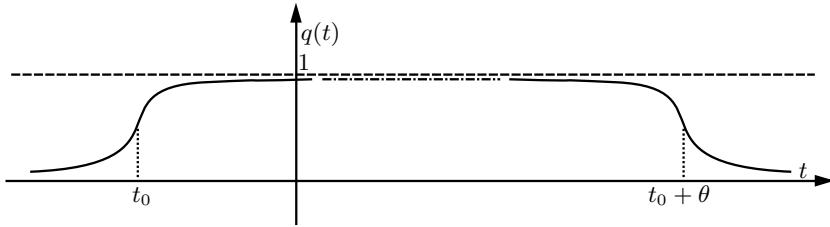


Fig. 42.1 A two-instanton configuration

42.1.1 The two-instanton configuration

We define a two-instanton (really an instanton–anti-instanton) configuration, as a sufficiently differentiable path, which depends on an additional time parameter, characterizing the separation between instantons. For large separations, it must decompose into the superposition of two instantons (Fig. 42.1), and minimize the variation of the action.

For this purpose, we could introduce a constraint in the path integral fixing the separation between instantons and solve the equation of motion with a Lagrange multiplier for the constraint (see Sections 37.4.1 and A42.2.1). Instead, we use a method which, at least at leading order is simpler, and shows that the result is universal [406, 407, 408].

We consider a path $q_c(t)$, sum of instantons separated by a distance $\theta > 0$, up to an additive constant adjusted in such a way as to satisfy the boundary conditions. It is convenient to introduce the notation

$$\underline{u}(t) = f(t + \theta/2), \quad v(t) = 1 - \underline{u}(-t) = \underline{u}(t - t_0), \quad (42.6)$$

where $f(t)$ is the function (42.4). We then consider the path $q_c(t)$, with

$$q_c(t)\sqrt{g} = \underline{u}(t) + \underline{u}(-t) - 1 = \underline{u}(t) - v(t) \quad (42.7)$$

(Again time translation $t \mapsto t + t_0$ generates a set of degenerate configurations). The path has the following properties: it is continuous and differentiable and, when θ is large, it differs, near each instanton, from the instanton solution only by exponentially small terms of order $e^{-\theta}$. Although the calculation of the corresponding action is simple, we describe it stepwise to show the generality of the ansatz. The action of the path (42.7) is

$$\begin{aligned} \mathcal{S}(q_c) &= \int dt \left[\frac{1}{2} \dot{q}_c^2(t) + V(q_c(t)\sqrt{g})/g \right] = \frac{2}{6g} + \frac{1}{g} \Sigma(\underline{u}, v), \text{ with} \\ \Sigma(\underline{u}, v) &= \int dt \left[-\dot{\underline{u}}(t)\dot{v}(t) + V(\underline{u}(t) - v(t)) - V(\underline{u}(t)) - V(v(t)) \right], \end{aligned} \quad (42.8)$$

where equation (42.5) for \underline{u} and v has been used. Since $q_c(t)$ is even, the integration can be restricted to the region $t < 0$, where $v(t)$ is small. After an integration by parts of the term $\dot{\underline{u}}(t)\dot{v}(t)$, we find

$$\Sigma(\underline{u}, v) = -2v(0)\dot{\underline{u}}(0) + 2 \int_{-\infty}^0 dt \left[v(t)\ddot{\underline{u}}(t) + V(\underline{u}(t) - v(t)) - V(\underline{u}(t)) - V(v(t)) \right].$$

We now expand in powers of v , and take into account the equation of motion for \underline{u} . The leading terms is of order $e^{-\theta}$, and we thus stop at order v^2 . We obtain

$$\Sigma(\underline{u}, v) = -2v(0)\dot{\underline{u}}(0) + 2 \int_{-\infty}^0 dt \left[\frac{1}{2} v^2(t) V''(\underline{u}(t)) - V(v(t)) \right]. \quad (42.9)$$

The function v decreases exponentially away from the origin. Therefore, the main contributions to the integral come from the neighbourhood of $t = 0$, where $V''(\underline{u}) \sim 1$. Moreover, $V(v)$ can be replaced at leading order by $\frac{1}{2}v^2$, and the two terms cancel.

The only term left is the integrated contribution,

$$v(0)\dot{u}(0) \sim e^{-\theta},$$

and thus

$$\mathcal{S}(q_c) = g^{-1} \left[\frac{1}{3} - 2e^{-\theta} + O(e^{-2\theta}) \right]. \quad (42.10)$$

It will become clearer later why we need the classical action only up to order $e^{-\theta}$ (a contribution that we call later *instanton interaction*). It is also useful to keep β large, but finite, in the calculation. Symmetry between θ and $\beta - \theta$ then implies

$$\mathcal{S}(q_c) = g^{-1} \left[\frac{1}{3} - 2e^{-\theta} - 2e^{-(\beta-\theta)} \right]. \quad (42.11)$$

As a verification, we calculate the extremum of $\mathcal{S}(q)$ at β fixed and obtain

$$\theta_c = \beta/2, \Rightarrow \mathcal{S}(q_c) = g^{-1} \left[\frac{1}{3} - 4e^{-\beta/2} + O(e^{-\beta}) \right].$$

In Chapter 39, for the same Hamiltonian, we have found for β large (equation (39.14))

$$\mathcal{S}(q_c) = g^{-1} \left[\frac{1}{6} - 2e^{-\beta} + O(e^{-2\beta}) \right]. \quad (42.12)$$

Both results are consistent. Indeed, to compare them, we have to replace β by $\beta/2$ in equation (42.12) and multiply the action by a factor 2, since the action corresponds to a trajectory described twice in the total time β .

The variation of the action. We now show that if we infinitesimally (for θ large) modify the configuration to further decrease the variation of the action, the change $r(t)$ of the path will be of order $e^{-\theta}$, and the variation of the action of order $e^{-2\theta}$, at least. Setting

$$q(t) = q_c(t) + r(t), \quad (42.13)$$

and expanding the action up to second order in $r(t)$, we find,

$$\begin{aligned} \mathcal{S}(q_c + r) &= \mathcal{S}(q_c) + \int \left[\dot{q}_c(t)\dot{r}(t) + \frac{1}{\sqrt{g}} V'(q_c(t)\sqrt{g})r(t) \right] dt \\ &\quad + \frac{1}{2} \int dt [r^2(t) + V''(q_c\sqrt{g})r^2(t)] + O([r(t)]^3). \end{aligned} \quad (42.14)$$

In the term linear in $r(t)$, we integrate by parts $\dot{r}(t)$ in order to use the property that $q_c(t)$ approximately satisfies the equation of motion. In the term proportional to $r^2(t)$, we replace V'' by 1, since we expect $r(t)$ to be large only far from the instantons. We then verify that the term linear in r is of order $e^{-\theta}$, while the quadratic term is of order 1. A shift of r , to eliminate the linear term in r , would then generate a contribution of order $e^{-2\theta}$, which is negligible at the order we calculate.

42.1.2 The n -instanton configuration and action

An n -instanton configuration corresponds to a succession of n instantons, separated by times θ_i , such that

$$\sum_{i=1}^n \theta_i = \beta. \quad (42.15)$$

At leading order, we only need to take into account ‘interactions’ between nearest neighbours. This is an essential simplifying feature of one-dimensional QM. The classical action $\mathcal{S}_c(\theta_i)$ can then be directly inferred from expression (42.11):

$$\mathcal{S}_c(\theta_i) = \frac{1}{g} \left[\frac{n}{6} - 2 \sum_{i=1}^n e^{-\theta_i} + O(e^{-(\theta_i+\theta_j)}) \right]. \quad (42.16)$$

Other interactions are negligible, because they are of higher order in $e^{-\theta}$.

Note that, for n even, the n -instanton configuration contributes to $\text{tr } e^{-\beta H}$, while for n odd it contributes to $\text{tr } P e^{-\beta H}$ (P is the reflection operator). Then, calculating

$$\mathcal{Z}_\epsilon = \frac{1}{2} \text{tr} [(1 + \epsilon P) e^{-\beta H}], \quad (42.17)$$

we obtain for $\epsilon = +1$ and $\epsilon = -1$ contributions to the even and odd eigenstate energies, respectively.

Remark. Since one keeps in the action all terms of order $e^{-\beta}$, one expects to find the contributions not only to the two lowest eigenvalues, but also to all eigenvalues that remain finite when g goes to zero (see the remark after equation (37.95) in Section 37.7).

42.1.3 The n -instanton contribution

We have calculated the n -instanton action. We now evaluate, at leading order, the contribution to the path integral of the neighbourhood of the classical path. Although the path is not a solution of the equation of motion, we have defined it in such a way that the linear terms can be neglected in the Gaussian integration. The second derivative of the action at the classical path,

$$M(t', t) = [-d_t^2 + V''(\sqrt{g}q_c(t))] \delta(t - t') \quad (42.18)$$

has the form of a quantum Hamiltonian, with a potential that consists of n largely separated wells, almost identical to the well arising in the one-instanton problem. Therefore, at leading order, the corresponding spectrum is the spectrum arising in the one-instanton problem n -times degenerate. Corrections are exponentially small in the separation (for details see Section A42.1). Moreover, by introducing n collective time variables, we suppress n times the eigenvalue 0, and generate the Jacobian of the one-instanton case to the power n . Therefore, the n -instanton contribution to \mathcal{Z}_ϵ (equation (42.17)) reduces to

$$\mathcal{Z}_\epsilon^{(n)} = e^{-\beta/2} \frac{\beta}{n} \left(\frac{e^{-1/6g}}{\sqrt{\pi g}} \right)^n \int_{\theta_i \geq 0} \delta \left(\sum \theta_i - \beta \right) \prod_i d\theta_i \exp \left(\frac{2}{g} \sum_{i=1}^n e^{-\theta_i} \right). \quad (42.19)$$

All factors have already been explained, except the factor β which comes from the integration over a global time translation, and the factor $1/n$, which arises because the configuration is invariant under a cyclic permutation of the θ_i . The factor $e^{-\beta/2}$ is the usual normalization factor.

We define the quantity κ , the ‘fugacity’ of the instanton gas,

$$\kappa = \frac{\epsilon}{\sqrt{\pi g}} e^{-1/6g}, \quad (42.20)$$

which is half the one-instanton contribution at leading order. It is then convenient to introduce the sum $\Sigma(\beta, g)$ of the leading order n -instanton contributions:

$$\Sigma(\beta, g) = e^{-\beta/2} + \sum_{n=1}^{\infty} \mathcal{Z}_\epsilon^{(n)}(\beta, g). \quad (42.21)$$

If one neglects the instanton interaction (the dilute gas approximation), one can integrate over the θ_i 's, and calculate the sum. One finds

$$\Sigma(\beta, g) = e^{-\beta/2} \left(1 + \beta \sum_{n=1}^{\infty} \frac{\kappa^n}{n} \frac{\beta^{n-1}}{(n-1)!} \right) = e^{-\beta(1/2-\kappa)}. \quad (42.22)$$

We recognize the perturbative and one-instanton contributions, at leading order, to $E_\epsilon(g)$, the ground state and the first excited state energies:

$$E_\epsilon(g) = \frac{1}{2} + O(g) - \frac{\epsilon}{\sqrt{\pi g}} e^{-1/6g} (1 + O(g)). \quad (42.23)$$

Instanton interactions. To go beyond the one-instanton approximation, we have to take into account the interaction between instantons. We then face a problem: examining expression (42.19), we discover that the interaction between instantons is *attractive*.

Therefore, for g small, the dominant contributions to the integral come from paths in which the instantons are close. For such paths, the concept of instanton is no longer meaningful, since such paths cannot be distinguished from the fluctuations around the constant, or the one-instanton solution.

Actually, this phenomenon is consistent with the large-order behaviour analysis, which has indicated that the perturbative expansion in the case of potentials with degenerate minima is not Borel summable. An ambiguity is expected at the two-instanton order. If the perturbative expansion is ambiguous at the two-instanton order, it seems impossible to calculate a correction of the same order or smaller. To proceed any further, first we must define the sum of perturbation theory. Remarkably enough, for $g < 0$, the perturbative expansion of the double-well potential is, up to a normalization and the change of g in $-g$, identical to the perturbative expansion of the $O(2)$ anharmonic oscillator, which is Borel summable [447, 453]. Therefore, we define the sum of the perturbation series as the analytic continuation of this Borel sum from g negative to $g = |g| \pm i0$. This corresponds in the Borel transformation to integrate above or below the real positive axis. We then note that, for g negative, the interaction between instantons becomes *repulsive*, and the expression (42.19) is defined. Therefore, we calculate, for g small and negative, both the sum of the perturbation series and the instanton contributions, and perform an analytic continuation to g positive of all quantities consistently. In the same way, the perturbative expansion around each multi-instanton configuration is also not Borel summable, and must be summed by the same procedure.

42.1.4 The calculation

We start from the n -instanton contribution (42.19),

$$\mathcal{Z}_\epsilon^{(n)} \sim \frac{\beta}{n} e^{-\beta/2} \kappa^n \int_{\theta_i \geq 0} \delta \left(\sum \theta_i - \beta \right) \prod_{i=1}^n d\theta_i \exp \left(\frac{2}{g} \sum_{i=1}^n e^{-\theta_i} \right). \quad (42.24)$$

To factorize the integral over the variables θ_i , we replace the δ -function by the integral representation,

$$\delta \left(\sum_{i=1}^n \theta_i - \beta \right) = \frac{1}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} ds \exp \left[-s \left(\beta - \sum_{i=1}^n \theta_i \right) \right], \text{ with } \eta > 0. \quad (42.25)$$

In terms of the function

$$I(s) = \int_0^{+\infty} e^{s\theta - \mu e^{-\theta}} d\theta, \quad (42.26)$$

the integral (42.24) can be rewritten as

$$\mathcal{Z}_\epsilon^{(n)} \sim \frac{\beta e^{-\beta/2}}{2i\pi} \frac{\kappa^n}{n} \int_{-i\infty-\eta}^{i\infty-\eta} ds e^{-\beta s} [I(s)]^n, \text{ with } \mu = -2/g. \quad (42.27)$$

By giving to s a small negative real part, we have ensured the convergence of the integral (42.26). To evaluate the integral (42.26), we set

$$\mu e^{-\theta} = t \quad (42.28)$$

and the integral becomes

$$I(s) = \int_0^\mu \frac{dt}{t} \left(\frac{\mu}{t}\right)^s e^{-t} = \int_0^{+\infty} \frac{dt}{t} \left(\frac{\mu}{t}\right)^s e^{-t} + O(e^{-\mu}/\mu), \quad (42.29)$$

for μ positive and large, that is, $g \rightarrow 0_-$. Up to an exponentially small correction, we thus obtain

$$I(s) \sim \mu^s \Gamma(-s). \quad (42.30)$$

The generating function $\Sigma(\beta, g)$ (equation (42.21)) of the leading order multi-instanton contributions then becomes

$$\begin{aligned} \Sigma(\beta, g) &= -\frac{\beta e^{-\beta/2}}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} ds e^{-\beta s} \sum_n \frac{\kappa^n}{n} \mu^{ns} [\Gamma(-s)]^n \\ &= -\frac{\beta e^{-\beta/2}}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} ds e^{-\beta s} \ln [1 - \kappa \mu^s \Gamma(-s)]. \end{aligned} \quad (42.31)$$

We set

$$E = s + 1/2, \quad \phi(E) = 1 - \kappa \mu^{E-1/2} \Gamma(1/2 - E). \quad (42.32)$$

We then integrate $\beta e^{-\beta s}$ by parts, and obtain

$$\Sigma(\beta, g) = -\frac{1}{2i\pi} \int_{-i\infty}^{+i\infty} dE e^{-\beta E} \frac{\phi'(E)}{\phi(E)}. \quad (42.33)$$

The asymptotic behaviour of the Γ -function (given by the Stirling formula) ensures the convergence of the integral and, moreover, the contour can be deformed to enclose the poles of the integrand in the half-plane $\text{Re}(E) > 0$. Integrating, we obtain the sum of residues

$$\Sigma(\beta, g) = \sum_{N \geq 0} e^{-\beta E_N}, \quad (42.34)$$

where the energies E_N are solutions of the spectral equation [408],

$$\phi(E) = 1 - \kappa \mu^{E-1/2} \Gamma(1/2 - E) = 0. \quad (42.35)$$

Since κ is small, a zero E of this equation is close to a pole of $\Gamma(1/2 - E)$:

$$E_N = N + \frac{1}{2} + O(\kappa), \quad N \geq 0. \quad (42.36)$$

We then expand the solutions of equation (42.35) as a power series of κ :

$$E_N(g) = \sum E_N^{(n)}(g) \kappa^n. \quad (42.37)$$

We obtain, at once, the many-instanton contributions to all energy levels $E_N(g)$ of the double-well potential, at leading order. It is convenient to write equation (42.35) as

$$\frac{e^{-1/6g}}{\sqrt{2\pi}} \left(-\frac{2}{g}\right)^E \Gamma(\tfrac{1}{2} - E) = -\epsilon i \Leftrightarrow \frac{\cos \pi E}{\pi} = \epsilon i \frac{e^{-1/6g}}{\sqrt{2\pi}} \left(-\frac{2}{g}\right)^E \frac{1}{\Gamma(\tfrac{1}{2} + E)}. \quad (42.38)$$

For example, the one-instanton contribution is

$$E_N^{(1)}(g) = -\frac{\epsilon}{N!} \left(\frac{2}{g}\right)^{N+1/2} \frac{e^{-1/6g}}{\sqrt{2\pi}} (1 + O(g)). \quad (42.39)$$

The two-instanton contribution is then,

$$E_N^{(2)}(g) = \frac{1}{(N!)^2} \left(\frac{2}{g}\right)^{2N+1} \frac{e^{-1/3g}}{2\pi} [\ln(-2/g) - \psi(N+1) + O(g \ln g)], \quad (42.40)$$

where ψ is the logarithmic derivative of the Γ -function.

The appearance of a factor $\ln g$ can be simply understood by noting that the interaction terms are only relevant for $g^{-1} e^{-\theta}$ of order 1, that is, θ of order $-\ln g$.

More generally, it can be verified that the n -instanton contribution has, at leading order, the form

$$E_N^{(n)}(g) = -\left(\frac{2}{g}\right)^{n(N+1/2)} \left(\frac{e^{-1/6g}}{\sqrt{2\pi}}\right)^n \left[P_n^N(\ln(-g/2)) + O(g(\ln g)^{n-1})\right], \quad (42.41)$$

in which $P_n^N(\sigma)$ is a polynomial of degree $(n-1)$. For example, for $N=0$, one finds

$$P_2(\sigma) = \sigma + \gamma, \quad P_3(\sigma) = \frac{3}{2}(\sigma + \gamma)^2 + \frac{\pi^2}{12}, \quad (42.42)$$

in which γ is Euler's constant $\gamma = -\psi(1) = 0.577215\dots$

Analytic continuation. When we perform our analytic continuation from g negative to g positive, two things happen: the Borel sums become complex, with an imaginary part exponentially smaller by about a factor $e^{-1/3g}$ than the real part. Simultaneously, the function $\ln(-2/g)$ also becomes complex and yields an imaginary part $\pm i\pi$. Since the sum of all the contributions is real, the imaginary parts should cancel (a sign of a resurgent structure). This argument leads to an evaluation of the imaginary part of the Borel sum of the perturbation series, or of the expansion around one instanton, for example.

From the imaginary part of P_2 , we infer

$$\text{Im } E^{(0)}(g) \sim \frac{1}{\pi g} e^{-1/3g} \text{Im}[P_2(\ln(-g/2))], \quad (42.43)$$

and therefore,

$$\text{Im } E^{(0)}(g) \sim -\frac{1}{g} e^{-1/3g}. \quad (42.44)$$

Using a dispersion relation to calculate the coefficients $E_k^{(0)}$ of the expansion of

$$E^{(0)}(g) = \sum_k E_k^{(0)} g^k, \quad (42.45)$$

we obtain the large-order behaviour of the perturbative expansion,

$$E_k^{(0)} \underset{k \rightarrow \infty}{\sim} \frac{1}{\pi} \int_0^\infty \text{Im}[E^{(0)}(g)] \frac{dg}{g^{k+1}}, \Rightarrow E_k^{(0)} \sim -\frac{1}{\pi} 3^{k+1} k!. \quad (42.46)$$

From the imaginary part of P_3 , we derive the large-order behaviour of the expansion of

$$E^{(1)}(g) = -\frac{1}{\sqrt{\pi g}} e^{-1/6g} \left(1 + \sum_{k=1}^{\infty} E_k^{(1)} g^k \right). \quad (42.47)$$

We express that the imaginary parts of $E^{(1)}(g)$ and $E^{(3)}(g)$ cancel at leading order:

$$\text{Im } E^{(1)}(g) \sim - \left(\frac{e^{-1/6g}}{\sqrt{\pi g}} \right)^3 \text{Im} [P_3(\ln(-g/2))]. \quad (42.48)$$

We derive the coefficients $E_k^{(1)}$ from the dispersion integral,

$$E_k^{(1)} = \frac{1}{\pi} \int_0^\infty \left\{ \text{Im} [E^{(1)}(g)] \sqrt{\pi g} e^{1/6g} \right\} \frac{dg}{g^{k+1}}. \quad (42.49)$$

Then, using equations (42.42) and (42.48), we find

$$E_k^{(1)} \sim -\frac{1}{\pi} \int_0^\infty 3 [\ln(2/g) + \gamma] e^{-1/3g} \frac{dg}{g^{k+2}}. \quad (42.50)$$

Finally, at leading order for k large, we can replace g by its saddle point value $1/3k$ in $\ln g$, and obtain

$$E_k^{(1)} = -\frac{3^{k+2}}{\pi k!} \left[\ln 6k + \gamma + O\left(\frac{\ln k}{k}\right) \right]. \quad (42.51)$$

Both results (42.46) and (42.51) have been shown to agree with the numerical behaviour of the corresponding series by calculating about 100 terms of the series.

The behaviour of the real part of P_2 for g has also been verified numerically [395].

42.2 The periodic cosine potential

The structure of the low-lying levels of the periodic cosine potential are discussed in Section 39.2, and we borrow the results here.

To avoid the proliferation of big integer factors, it is convenient to use a specific normalization of the coupling constant. We write the Hamiltonian as

$$H = -\frac{1}{2} \left(\frac{d}{dq} \right)^2 + \frac{1}{16g} (1 - \cos 4q\sqrt{g}). \quad (42.52)$$

For g small, the spectrum has a band structure. The unitary operator T , which translates wave functions by one period $\pi/2\sqrt{g}$ of the potential,

$$T\psi(q) \equiv \psi(q + \pi/2\sqrt{g}),$$

commutes with the Hamiltonian, and can thus be diagonalized simultaneously. An eigenstate in a band, that we denote by $|N, \varphi, g\rangle$, is characterized by the eigenvalue $e^{i\varphi}$ of T , and the corresponding eigenfunction $\psi_N(\varphi, g, q)$ of H :

$$T|N, \varphi, g\rangle = e^{i\varphi} |N, \varphi, g\rangle, \quad H|N, \varphi, g\rangle = \mathcal{E}_N(\varphi, g)|N, \varphi, g\rangle. \quad (42.53)$$

For g small, the eigenvalue $\mathcal{E}_N(\varphi, g)$ is a periodic function of the angle φ , and $\mathcal{E}_N(\varphi, g) = N + 1/2 + O(g)$.

The partition function in the φ -sector. We define the partition function in the sector corresponding to an angle φ as the sum

$$\mathcal{Z}(\beta, \varphi, g) = \sum_{N=0} e^{-\beta \mathcal{E}_N(\varphi, g)}. \quad (42.54)$$

In particular, for β large,

$$\mathcal{Z}(\beta, \varphi, g) \underset{\beta \rightarrow \infty}{\sim} e^{-\beta \mathcal{E}_0(\varphi, g)}. \quad (42.55)$$

We also define the quantity (equation (A39.8))

$$\begin{aligned} \mathcal{Z}_l(\beta, g) &\equiv \text{tr} (T^l e^{-\beta H}) \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\varphi \sum_N e^{-\beta \mathcal{E}_N(\varphi, g)} e^{il\varphi} = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \mathcal{Z}(\beta, \varphi, g) e^{il\varphi}. \end{aligned} \quad (42.56)$$

The inverse of last relation is,

$$\mathcal{Z}(\beta, \varphi, g) = \sum_{l=-\infty}^{+\infty} e^{-il\varphi} \mathcal{Z}_l(\beta, g). \quad (42.57)$$

The path integral representation of $\mathcal{Z}_l(\beta, g)$ is

$$\mathcal{Z}_l(\beta, g) = \int_{q(\beta/2)=q(-\beta/2)+l\pi/2\sqrt{g}} [dq(t)] \exp [-\mathcal{S}(q)], \quad (42.58)$$

with

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} dt \left[\frac{1}{2} \dot{q}^2(t) + \frac{1}{16g} (1 - \cos 4q(t)\sqrt{g}) \right]. \quad (42.59)$$

The factor $e^{-il\varphi}$ can be incorporated in the path integral, since

$$-il\varphi = -\frac{2\sqrt{g}}{\pi} \varphi [q(\beta/2) - q(-\beta/2)] = -\frac{2i\sqrt{g}}{\pi} \int_{-\beta/2}^{+\beta/2} dt \dot{q}(t).$$

This corresponds to adding to $\mathcal{S}(q)$ the integral of a local density,

$$\mathcal{S}(q) \mapsto \mathcal{S}(q) + \frac{2i\sqrt{g}}{\pi} \int_{-\beta/2}^{+\beta/2} dt \dot{q}(t). \quad (42.60)$$

The sum over l then is obtained by summing over all periodic trajectories. In the infinite β limit, the configurations for which $q(\beta/2) - q(-\beta/2)$ is not a multiple of the period are suppressed, since their classical action is necessarily infinite. The sum (42.57) can thus be expressed without restriction on the path $q(t)$ as

$$e^{-\beta \mathcal{E}_0(\varphi, g)} \underset{\beta \rightarrow \infty}{\sim} \int [dq(t)] \exp \left[-\mathcal{S}(q) - i \frac{2\sqrt{g}}{\pi} \varphi \int_{-\beta/2}^{+\beta/2} dt \dot{q}(t) \right]. \quad (42.61)$$

These expressions have natural generalizations in the case of the θ -vacuum of the $CP(N-1)$ model, and non-Abelian gauge theories (see Sections 39.5 and 39.6).

Remark. In an expansion of an eigenvalue $\mathcal{E}_N(\varphi, g)$ in a Fourier series,

$$\mathcal{E}_N(\varphi, g) = \sum_{l=-\infty}^{+\infty} E_N(l, g) e^{il\varphi}, \quad E_N(l, g) = E_N(-l, g), \quad (42.62)$$

for g small, $E_{N,l}(g)$ is dominated by l -instanton contributions. In particular, for the ground state energy $\mathcal{E}_0(\varphi, g)$ in the φ sector, , the $l = 1$ term behaves like

$$E_0(l=1, g) \sim \frac{1}{\sqrt{\pi g}} e^{-1/2g}. \quad (42.63)$$

Multi-instanton configurations. This example, and the double-well potential example differ in one important aspect. In the double-well case, each configuration is a succession of instantons and anti-instantons. By contrast, here, at each step, the instanton can go to the next minimum of the potential, or the preceding one. Therefore, we assign a sign $\epsilon = +1$ to an instanton and a sign $\epsilon = -1$ to an anti-instanton. A simple calculation, similar to the calculation of Section 42.1 (for details, see Section A42.2), yields the interaction between two consecutive instantons of types ϵ_1 and ϵ_2 separated by a distance θ_{12} :

$$\frac{2\epsilon_1\epsilon_2}{g} e^{-\theta_{12}}. \quad (42.64)$$

The interaction between instantons of the same kind is repulsive, while it is attractive for different kinds. We denote the one-instanton contribution at leading order by

$$\kappa = \frac{1}{\sqrt{\pi g}} e^{-1/2g}. \quad (42.65)$$

With this notation, the n -instanton contribution reads

$$\mathcal{Z}^{(n)}(\beta, \varphi, g) = \beta e^{-\beta/2} \frac{\kappa^n}{n} \int_{\theta_1 \geq 0} \delta\left(\sum_{i=1}^n \theta_i - \beta\right) J_n(\theta_i), \quad (42.66)$$

with

$$J_n(\theta_i) = \sum_{\epsilon_i=\pm 1} \exp\left(\sum_{i=1}^n -\frac{2}{g} \epsilon_i \epsilon_{i+1} e^{-\theta_i} - i \epsilon_i \varphi\right). \quad (42.67)$$

The additional term $-i \epsilon_i \varphi$ originates from the expression (42.57). We have identified ϵ_{n+1} and ϵ_1 .

In contrast with the example of the double-well potential, the interaction between instantons contains both attractive and repulsive terms. Thus, we have to begin with g complex to perform the analytic continuation of both the Borel sums and the instanton contributions.

Following the same steps as in the case of the double-well potential, we obtain

$$\begin{aligned} \mathcal{Z}^{(n)}(\beta, \varphi, g) &= \frac{\beta}{2i\pi} e^{-\beta/2} \frac{\kappa^n}{n} \oint ds e^{-\beta s} \Gamma^n(-s) \\ &\times \sum_{\{\epsilon_i=\pm 1\}} \exp\left[\sum_{i=1}^n -i \epsilon_i \varphi - s \ln(\epsilon_i \epsilon_{i+1} g/2)\right]. \end{aligned} \quad (42.68)$$

We introduce the notation

$$\sigma = \ln(g/2), \quad (42.69)$$

and choose to make the analytic continuation from above so that

$$\ln\left(\frac{1}{2}g\epsilon_i\epsilon_{i+1}\right) = \sigma - \frac{1}{2}i\pi(1 - \epsilon_i\epsilon_{i+1}). \quad (42.70)$$

The expression (42.68) can then be written as

$$\begin{aligned} \mathcal{Z}^{(n)}(\beta, \varphi, g) &\sim \frac{\beta}{2i\pi} e^{-\beta/2} \frac{\kappa^n}{n} \oint ds e^{-\beta s} \left([\Gamma(-s) e^{-\sigma s}]^n\right. \\ &\times \left. \sum_{\{\epsilon_i=\pm 1\}} \exp\left[\sum_{i=1}^n -i \epsilon_i \varphi + \frac{1}{2}i\pi s(1 - \epsilon_i \epsilon_{i+1})\right]\right). \end{aligned} \quad (42.71)$$

The summation over the set $\{\epsilon_i\}$ corresponds to the calculation of the partition function of a one-dimensional Ising model, whose transfer matrix is

$$\mathbf{M} = \begin{bmatrix} e^{-i\varphi} & e^{i\pi s} \\ e^{i\pi s} & e^{i\varphi} \end{bmatrix}. \quad (42.72)$$

The sum then is $\text{tr } \mathbf{M}^n$. The eigenvalues m_{\pm} of \mathbf{M} are

$$m_{\pm} = \cos \varphi \pm (e^{2i\pi s} - \sin^2 \varphi)^{1/2}. \quad (42.73)$$

The expression (42.71) can then be written as

$$\mathcal{Z}^{(n)}(\beta, \varphi, g) \sim \frac{\beta e^{-\beta/2}}{2i\pi} \frac{\kappa^n}{n} \oint ds e^{-\beta s} [\Gamma(-s) e^{-\sigma s}]^n (m_+^n + m_-^n). \quad (42.74)$$

The sum $\Sigma(\beta, g)$ of all leading order multi-instanton contributions can be calculated and yields,

$$\begin{aligned} \Sigma(\beta, g) &= e^{-\beta/2} - \frac{\beta e^{-\beta/2}}{2i\pi} \oint ds e^{-\beta s} \ln \{ [1 - \kappa \Gamma(-s) e^{-\sigma s} m_+(s)] \\ &\quad \times [1 - \kappa \Gamma(-s) e^{-\sigma s} m_-(s)] \}. \end{aligned} \quad (42.75)$$

The argument of the logarithm can also be written as

$$\begin{aligned} &[1 - \kappa \Gamma(-s) e^{-\sigma s} m_+(s)] [1 - \kappa \Gamma(-s) e^{-\sigma s} m_-(s)] \\ &= 1 - \kappa \Gamma(-s) e^{-\sigma s} \left[2 \cos \varphi + \kappa \frac{2i\pi e^{(i\pi-\sigma)s}}{\Gamma(1+s)} \right]. \end{aligned} \quad (42.76)$$

An integration by parts of $\beta e^{-\beta s}$ in the integral (42.75) yields the final result:

$$\Sigma(\beta, g) = \sum_{N=0}^{\infty} e^{-\beta E_N(g)}, \quad (42.77)$$

in which $E_N(g) = \frac{1}{2} + s_N(\kappa, \sigma)$ is a solution expandable in powers of κ of the equation [408],

$$\left(\frac{2}{g}\right)^{-E} \frac{e^{1/2g}}{\Gamma(\frac{1}{2} - E)} + \left(-\frac{2}{g}\right)^E \frac{e^{-1/2g}}{\Gamma(\frac{1}{2} + E)} = \frac{2 \cos \varphi}{\sqrt{2\pi}}. \quad (42.78)$$

Note the symmetry in the change $g, E \mapsto -g, -E$. However, this symmetry is slightly misleading, because the equation is actually quadratic in $\Gamma(\frac{1}{2} - E)$ and only one root, corresponding to m_+ , is relevant for $g > 0$.

42.3 General potentials with degenerate minima

We now consider a general analytic potential having two degenerate minima located at the origin and another point x_0 :

$$\begin{aligned} V(x) &= \frac{1}{2}x^2 + O(x^3), \\ V(x) &= \frac{1}{2}\omega^2(x - x_0)^2 + O((x - x_0)^3). \end{aligned} \quad (42.79)$$

For definiteness, we assume $\omega > 1$.

In such a situation, the classical equation of motion have instanton solutions connecting the two minima of the potential. However, there is no ground state degeneracy beyond the classical level. Therefore, the one-instanton solution does not contribute anymore to the path integral. Only periodic classical paths are relevant; the leading contribution now comes from the two-instanton configuration.

To determine the potential between instantons and the normalization of the path integral, it is convenient to first calculate the contribution at β finite of a trajectory described n times, and then take the large β limit of the expression. Using the expressions derived in Section 37.5, we infer

$$\{\text{tr } e^{-\beta H}\}_{(n)} = i(-1)^n \frac{\beta}{n\sqrt{\pi g}} \sqrt{\frac{\omega C}{n(1+\omega)}} e^{-\omega\beta/[2n(1+\omega)]} e^{-nA(\beta)/g}, \quad (42.80)$$

with the definitions

$$C = x_0^2 \omega^{2/(1+\omega)} \exp \left\{ \frac{2\omega}{1+\omega} \left[\int_0^{x_0} dx \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} - \frac{1}{\omega(x_0-x)} \right) \right] \right\}, \quad (42.81)$$

and

$$A(\beta) = 2 \int_0^{x_0} \sqrt{2V(x)} dx - 2C \frac{(1+\omega)}{\omega} e^{-(\beta/n)\omega/(1+\omega)} + \dots \quad (42.82)$$

Note that n does not have the same meaning here as in Section 42.1. Since ω is different from 1, the one-instanton configuration does not contribute and n instead counts the number of instanton and anti-instanton pairs in the terminology of Section 42.1. Therefore, n here actually corresponds to $2n$ in the $\omega = 1$ limit.

42.3.1 The n -instanton action

We denote by θ_i the successive amounts of time the classical trajectory spends near x_0 , and φ_i the amounts it spends near the origin. The n -instanton action takes the form

$$A(\theta_i, \varphi_j) = na - 2 \sum_{i=1}^n (C_1 e^{-\omega\theta_i} + C_2 e^{-\varphi_i}) \quad (42.83)$$

with $\sum_{i=1}^n (\theta_i + \varphi_i) = \beta$ and

$$a = 2 \int_0^{x_0} \sqrt{2V(x)} dx. \quad (42.84)$$

By comparing the value of the action at the saddle point

$$\theta_i = \frac{\beta}{n(1+\omega)}, \quad \varphi_i = \frac{\omega\beta}{n(1+\omega)} \quad (42.85)$$

with the expression (42.82), we see that we can choose

$$C_1 = C/\omega, \quad C_2 = C \quad (42.86)$$

by adjusting the definitions of θ and φ .

42.3.2 The n -instanton contribution

The n -instanton contribution then has the form

$$\begin{aligned} \{\text{tr } e^{-\beta H}\}_{(n)} &= \beta e^{-\beta/2} \frac{e^{-na/g}}{(\pi g)^n} N_n \\ &\times \int_{\theta_i, \varphi_i \geq 0} \delta\left(\sum_i (\theta_i + \varphi_i) - \beta\right) \exp\left[\sum_{i=1}^n \frac{1}{2}(1-\omega)\theta_i - \frac{1}{g} A(\theta, \varphi)\right]. \end{aligned} \quad (42.87)$$

The additional term $\sum_i \frac{1}{2}(1-\omega)\theta_i$ in the integrand comes from the determinant generated by the Gaussian integration around the classical path. The normalization can be obtained by evaluating the integral over the variables θ_i and φ_i using the steepest descent method, and comparing the result with expression (42.80). The result is

$$N_n = \frac{(C\sqrt{\omega})^n}{n}. \quad (42.88)$$

The factor $1/n$ comes from the symmetry of the action under cyclic permutations of the θ_i and φ_i .

We set

$$\kappa = \frac{e^{-a/g}}{\pi g} C\sqrt{\omega}, \quad \mu = -\frac{2C}{g}. \quad (42.89)$$

As in Section 42.1, we introduce an integral representation for the δ -function:

$$\delta\left(\sum_{i=1}^n (\theta_i + \varphi_i) - \beta\right) = \frac{1}{2i\pi} \int_{-i\infty-\eta}^{+i\infty+\eta} ds \exp\left[-s\beta + s \sum_{i=1}^n (\theta_i + \varphi_i)\right], \quad \text{for } n \geq 1. \quad (42.90)$$

The expression (42.87) can be rewritten as

$$\{\text{tr } e^{-\beta H}\}_{(n)} = \beta e^{-\beta/2} \frac{\kappa^n}{n} \frac{1}{2i\pi} \int_{-i\infty-\eta}^{+i\infty-\eta} ds e^{-s\beta} [I(s)J(s)]^n, \quad (42.91)$$

where we have defined

$$I(s) = \int_0^{+\infty} e^{\theta s - \mu e^{-\theta}} d\theta, \quad (42.92)$$

$$J(s) = \int_0^{+\infty} \exp\left\{\left[\frac{1}{2}(1-\omega) + s\right]\theta - \frac{\mu}{\omega} e^{-\omega\theta}\right\} d\theta. \quad (42.93)$$

The integrals can be calculated explicitly in the small g limit:

$$I(s) = \Gamma(-s)\mu^s, \quad (42.94)$$

$$J(s) = \frac{1}{\sqrt{\mu\omega}} \Gamma\left(\frac{1}{2} - (s + \frac{1}{2})/\omega\right) \left(\frac{\mu}{\omega}\right)^{(s+1/2)/\omega}. \quad (42.95)$$

We denote by $\Sigma(\beta, g)$, the generating functional of the many-instanton contributions.

Summing over n and integrating by parts, we obtain $\Sigma(\beta, g)$ as a sum of residues:

$$\Sigma(\beta, g) = \sum_{\alpha} e^{-\beta E_{\alpha}}, \quad (42.96)$$

in which the values $E_{\alpha} = \frac{1}{2} + s_{\alpha}$ are the solutions expandable for g small of the equation

$$\left(\frac{\mu}{\omega}\right)^{E/\omega} \Gamma\left(\frac{1}{2} - E/\omega\right) \mu^E \Gamma\left(\frac{1}{2} - E\right) \frac{e^{-a/g}}{2\pi} = -1. \quad (42.97)$$

We now note that we find two series of energy levels corresponding to the poles of the two Γ -functions:

$$E_N = N + \frac{1}{2} + O(\kappa), \quad (42.98)$$

$$E_N = (N + \frac{1}{2})\omega + O(\kappa). \quad (42.99)$$

The same expression contains the instanton contributions to the two different sets of eigenvalues.

One can verify that the many-instanton contributions are singular for $\omega = 1$. But if one directly sets $\omega = 1$ in equation (42.97), one obtains

$$\mu^{2E} \Gamma^2(\tfrac{1}{2} - E) \frac{e^{-a/g}}{2\pi} = -1,$$

equation that can be rewritten as

$$\mu^E \Gamma(\tfrac{1}{2} - E) \frac{e^{-a/(2g)}}{\sqrt{2\pi}} = \pm i. \quad (42.100)$$

This is exactly the set of two equations obtained in Section 42.1.

42.3.3 Large-order estimates of perturbation theory

The expression (42.97) can be used to determine the large-order behaviour of perturbation theory by calculating the imaginary part of the leading instanton contribution, and using a dispersion integral as we have done in Section 40.1.1. For the energy $E_N(g) = N + \frac{1}{2} + O(g)$, one finds

$$\text{Im } E_N(g) \sim K_N g^{-(N+1/2)(1+1/\omega)} e^{-a/g}, \quad (42.101)$$

with

$$K_N = \frac{(-1)^{N+1}}{2\pi N!} \omega^{-(N+1/2)/\omega} (2C)^{(N+1/2)(1+1/\omega)} \sin[\pi(N + \frac{1}{2})(1 + 1/\omega)] \\ \times \Gamma[\tfrac{1}{2} - (N + \tfrac{1}{2})/\omega]. \quad (42.102)$$

From the imaginary part of $E_N(g)$, one infers at large-order k :

$$E_{Nk} \sim K_N \frac{\Gamma(k + (N + 1/2)(1 + 1/\omega))}{a^{k+(N+1/2)(1+1/\omega)}} (1 + O(k^{-1})). \quad (42.103)$$

Note that, in contrast with the instanton contribution to the real part, the expression is uniform in the limit $\omega = 1$ and yields the result (42.46).

42.4 The $O(\nu)$ -symmetric anharmonic oscillator

As a last example, we consider the analytic continuation to negative coupling of the energy levels of the quartic anharmonic oscillator with $O(\nu)$ symmetry, corresponding to the Hamiltonian

$$H = -\frac{1}{2}\nabla^2 + \frac{1}{2}\mathbf{q}^2 + g(\mathbf{q}^2)^2, \quad \mathbf{q} \in \mathbb{R}^\nu. \quad (42.104)$$

We explicitly discuss the $\nu = 2$ example, but the extension to other values of ν is then simple.

The $O(2)$ -symmetric anharmonic oscillator. The instanton solution can be written as

$$\mathbf{q}(t) = \mathbf{u}f(t), \quad (42.105)$$

in which \mathbf{u} is a fixed unit vector. The one-instanton contribution to the ground state energy is

$$E^{(1)}(g) = \frac{4i}{g} e^{1/3g} (1 + O(g)), \quad \text{for } g \rightarrow 0-. \quad (42.106)$$

From a calculation of the instanton interaction, one infers the n -instanton action:

$$A(\theta_i) = -\frac{1}{3}n - 4 \sum_i e^{-\theta_i} \cos \varphi_i, \quad (42.107)$$

in which θ_i is the distance between two successive instantons and φ_i the angle between them:

$$\cos \varphi_i = \mathbf{u}_i \cdot \mathbf{u}_{i+1}. \quad (42.108)$$

It is useful to consider the quantity (Sections 3.4 and 14.3),

$$\text{tr} [R(\alpha) e^{-\beta H}] = \int [d\mathbf{q}(t)] \exp(-S(q)), \quad \text{with} \quad \hat{\mathbf{q}}(-\beta/2) \cdot \hat{\mathbf{q}}(\beta/2) = \cos \alpha. \quad (42.109)$$

The matrix $R(\alpha)$ is a rotation matrix, which rotates vectors by an angle α . It leads to the boundary condition that $\mathbf{q}(t)$ at initial and final times should differ by an angle α .

The right-hand side of equation (42.109) can be rewritten as

$$\text{tr} [R(\alpha) e^{-\beta H}] = \sum_{l,N} e^{-il\alpha - \beta E_{l,N}}. \quad (42.110)$$

In this expression, l is the angular momentum. The boundary condition imposed on the path integral (42.109) implies a constraint on the many-instanton configuration:

$$\sum_{i=1}^n \varphi_i = \alpha, \quad (42.111)$$

constraint which can be implemented through the identity

$$\delta \left(\sum_{i=1}^n \varphi_i - \alpha \right) = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} \exp \left[il \left(\sum_{i=1}^n \varphi_i \right) - i\alpha l \right]. \quad (42.112)$$

The n -instanton contribution to expression (42.110) then takes the form

$$\{\text{tr} [R(\alpha) e^{-\beta H}] \}_{(n)} \sim \frac{\kappa^n}{2i\pi n} \beta e^{-\beta} \int ds e^{-s\beta} \sum_{l=-\infty}^{+\infty} e^{-il\alpha} [I_l(s)]^n, \quad (42.113)$$

where we have set

$$\kappa = \frac{4i}{g} e^{1/3g}, \quad \mu = -\frac{4}{g}, \quad (42.114)$$

$$I_l(s) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_{-\infty}^{+\infty} d\theta \exp(s\theta + il\varphi - \mu e^{-\theta} \cos \varphi). \quad (42.115)$$

We denote by $\Sigma_l(\beta, g)$ the generating function of n -instanton contributions at fixed angular momentum l :

$$\Sigma_l(\beta, g) = \sum_n \frac{\kappa^n}{2i\pi n} \beta e^{-\beta} \int ds e^{-s\beta} [I_l(s)]^n. \quad (42.116)$$

To evaluate $I_l(s)$, we first integrate over θ :

$$I_l(s) = \mu^s \Gamma(-s) \int_0^{2\pi} \frac{d\varphi}{2\pi} e^{il\varphi} (\cos \varphi)^s. \quad (42.117)$$

Finally, performing the last integral and using various relations among Γ functions, we obtain

$$I_l(s) = \mu^s e^{\frac{1}{2}i\pi(s+1)} \frac{\Gamma(\frac{1}{2}(l-s)) 2^{-s-1}}{\Gamma(1 + \frac{1}{2}(l+s))}. \quad (42.118)$$

It is easy to verify that

$$I_l(s) = I_{-l}(s). \quad (42.119)$$

From equation (42.116), we now derive the result

$$\text{tr } e^{-\beta H_l} = \sum_{N=0}^{\infty} e^{-\beta E_{N,l}}, \quad (42.120)$$

with $E_{N,l} = s_{N,l} + 1$ being the solution of the equation

$$e^{1/(3g)} \left(-\frac{2}{g}\right)^E e^{i\pi(E+l)/2} \frac{\Gamma(\frac{1}{2}(l+1-E))}{\Gamma(\frac{1}{2}(l+1+E))} = 1, \quad (42.121)$$

which satisfies

$$E_{N,l} = l + 2N + 1 + O(g), \quad N \geq 0. \quad (42.122)$$

Note that checks about these expressions are provided by the surprising perturbative relation between the $O(2)$ anharmonic oscillator with negative coupling and the quartic double-well potential [408, 453, 454].

The $O(\nu)$ -symmetric Hamiltonian. One can extend this result to the general $O(\nu)$ case since, at fixed angular momentum l , the Hamiltonian depends only on the combination $l + \nu/2$. Hence, making in equation (42.121) the corresponding substitution, one obtains

$$i e^{1/(3g)} \left(-\frac{2}{g}\right)^E e^{i\pi(E+l+\nu/2)/2} \frac{\Gamma(\frac{1}{2}(l+\nu/2-E))}{\Gamma(\frac{1}{2}(l+\nu/2+E))} = 1. \quad (42.123)$$

At leading order in κ , one recovers the imaginary part of the energy levels for g small and negative:

$$\text{Im } E_{N,l} \underset{g \rightarrow 0_-}{=} -\frac{1}{N!} \frac{1}{\Gamma(\frac{1}{2}\nu + l + N)} \left(\frac{2}{g}\right)^{(\nu/2)+l+2N} e^{1/3g} (1 + O(g)). \quad (42.124)$$

Using the Cauchy formula, one can derive from this expression large-order estimates of perturbation theory. At next order in κ , one obtains the two-instanton contribution which is related by the same dispersion relation to the large-order behaviour of the perturbative expansion around one instanton.

42.5 Generalized Bohr–Sommerfeld quantization formula

So far, we have considered instanton contributions only at leading order. However, the form of the result is extremely suggestive and has led to a conjecture [408] about the general form of the semi-classical expansion for potentials with degenerate minima, proved to a large extent since [395, 455], using exact WKB methods [456] and resurgence theory [457].

To be specific, we explain the conjecture for the double-well potential, although it can be easily generalized to the other problems discussed in this chapter (see Section A42.4).

We introduce the function

$$D(E, g) = E + \sum_{k=1}^{\infty} g^k D_{k+1}(E), \quad (42.125)$$

in terms of which the perturbation expansion for an energy level $E_N^{(0)}$ can be obtained by inverting

$$N + \frac{1}{2} = D(E^{(0)}, g). \quad (42.126)$$

This is the form of the usual exact Bohr–Sommerfeld quantization condition.

One verifies that, as a power series, $D(E, g) = -D(-E, -g)$.

In the case of the double-well potential, to take into account instanton contributions, we generalize the Bohr–Sommerfeld quantization condition as

$$\frac{1}{\sqrt{2\pi}} \Gamma(\frac{1}{2} - D) (-2/g)^{D(E,g)} e^{-A(g,E)/2} = \pm i, \quad (42.127)$$

with

$$A(g, E) = \frac{1}{3g} + \sum_{k=1}^{\infty} g^k A_{k+1}(E), \quad (42.128)$$

where again, in the sense of power series, $A(g, E) = -A(-g, -E)$. The functions $A(E, g)$ and $D(E, g)$ can be calculated by expanding the corresponding WKB series, which are expansions at gE fixed, for E small (for details see Section A42.4). The coefficients $D_k(E)$ and $A_k(E)$ are polynomials of degree k in E .

If we solve equation (42.127) in the one-instanton approximation, and substitute into equation (42.125), we find ($\epsilon = \pm 1$)

$$E = E^{(0)}(g) - \epsilon \left(\frac{2}{g}\right)^N \frac{1}{N!} \frac{e^{-A(g,E^{(0)})/2}}{\sqrt{\pi g}} \frac{\partial D}{\partial E} \left(E^{(0)}\right)^{-1}. \quad (42.129)$$

Thus the knowledge of the two functions D and A is equivalent to the knowledge and the perturbative and one-instanton expansions for all levels and to all orders.

If we now systematically expand equation (42.127), we find for the energy level $E_N(g) = N + 1/2 + O(g)$ the following expansion,

$$E_N(g) = \sum_0^{\infty} E_l^{(0)} g^l + \sum_{n=1}^{\infty} \frac{1}{g^{N_n}} \left(\frac{1}{\sqrt{\pi g}} e^{-1/6g} \right)^n \sum_{k=0}^{n-1} (\ln(-2/g))^k \sum_{l=0}^{\infty} e_{nkl} g^l. \quad (42.130)$$

All the series in powers of g appearing in this expansion are determined by the perturbative expansion of A and D . This structure has found an explanation in the framework of the theory of *resurgent* functions.

The function $A(E, g)$ has initially been determined at this order by a combination of analytic and numerical calculations.

However, later, it has been proved for the double well and cosine potentials [455], using differential equation techniques, the remarkable relation

$$\frac{\partial E}{\partial D} = -6Dg - 3g^2 \frac{\partial A}{\partial g},$$

which reduces the determination of both functions to the much simpler determination of the perturbative spectral function $D(E, g)$.

Moreover, we conjecture that all these series have to be summed for g negative first, and the value of each instanton contribution for g positive is then obtained by analytic continuation. The property that the infinite number of perturbation series around all instantons are related may, at least in QM, simplify the problem of the summation of the many-instanton contributions.

In Section A42.4, we give the generalized Bohr–Sommerfeld quantization formulae for other potentials with degenerate minima.

A42 Additional remarks

A42.1 Multi-instantons: The determinant

We can express the operator M defined by equation (42.18) as

$$M = - \left(\frac{d}{dt} \right)^2 + 1 + \sum_{i=1}^n v(t - t_i), \quad (A42.1)$$

in which $v(t)$ is a potential localized around $t = 0$, and t_i are the positions of the instantons:

$$v(t) = O\left(e^{-|t|}\right), \quad \text{for } |t| \rightarrow \infty. \quad (A42.2)$$

We want to evaluate

$$\det(M M_0^{-1}) = \det \left\{ 1 + \left[-(d/dt)^2 + 1 \right]^{-1} \sum_{i=1}^n v(t - t_i) \right\}. \quad (A42.3)$$

Using the identity $\ln \det = \text{tr} \ln$, we expand the right-hand side in powers of $v(t)$ and find

$$\begin{aligned} \ln \det MM_0^{-1} &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \int \left[\Delta(u_1 - u_2) \sum_{i_1=1}^n v(u_2 - t_{i_1}) \Delta(u_2 - u_3) \cdots \Delta(u_k - u_1) \right. \\ &\quad \times \left. \sum_{i_k=1}^n v(u_1 - t_{i_k}) \right] \prod_{j=1}^k du_j, \end{aligned} \quad (A42.4)$$

with the definition

$$\Delta(t) = \left\langle 0 \left| \left[-(d/dt)^2 + 1 \right]^{-1} \right| t \right\rangle \sim \frac{1}{2} e^{-|t|}, \quad \text{for } 1 \ll t \ll \beta. \quad (A42.5)$$

It is clear from the behaviour of $v(t)$ and $\Delta(t)$ that, when the instantons are largely separated, only the terms in which one retains the same instanton contribution from each potential survive. Therefore,

$$\begin{aligned} \ln \det MM_0^{-1} &= n \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \int \Delta(u_1 - u_2) v(u_2) \cdots \Delta(u_k - u_1) v(u_1) \prod_{j=1}^k du_j, \\ &\quad \text{for } |t_i - t_j| \gg 1. \end{aligned} \quad (A42.6)$$

We recognize n times the logarithm of the one-instanton determinant.

A42.2 The instanton interaction

Like in Section 42.3, we assume that the potential has two degenerate minima located at the points $x = 0$ and $x = x_0$, with

$$\begin{aligned} V(x) &= \frac{1}{2}x^2 + O(x^3) \\ V(x) &= \frac{1}{2}\omega^2(x - x_0)^2 + O((x - x_0)^3). \end{aligned} \quad (A42.7)$$

The one-instanton solution $q_c(t)$, which goes from 0 to $q_0 = x_0/\sqrt{g}$, can be written as

$$q_c(t) = f(t)/\sqrt{g}. \quad (A42.8)$$

We choose the function $f(t)$ in such a way that it satisfies

$$\begin{aligned} x_0 - f(t) &\sim \sqrt{C} e^{-\omega t}/\omega, \quad \text{for } t \rightarrow +\infty, \\ f(t) &\sim \sqrt{C} e^t, \quad \text{for } t \rightarrow -\infty. \end{aligned} \quad (A42.9)$$

By solving the equation of motion, it is easy to calculate the constant

$$C = x_0^2 \omega^{2/(1+\omega)} \exp \left\{ \frac{2\omega}{1+\omega} \left[\int_0^{x_0} dx \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} - \frac{1}{\omega(x_0 - x)} \right) \right] \right\}. \quad (A42.10)$$

We recognize the constant (42.81).

We now construct instanton–anti-instanton pair configurations $q(t)$, which correspond to trajectories starting from, and returning to, $q = q_0$ or $q = 0$. Since we want also to deal with the case of two successive instantons, we assume, but only in this case, that $V(x)$ is an even function and has, therefore, a third minimum at $x = -x_0$.

Following the discussion of Section 42.2, we take as a two-instanton configuration:

$$q_1(t) = \frac{1}{\sqrt{g}}(f_+(t) + \epsilon f_-(t)), \quad \epsilon = \pm 1, \quad (A42.11)$$

with

$$f_+(t) = f(t - \theta/2), \quad f_-(t) = f(-t - \theta/2), \quad (A42.12)$$

where θ parametrizes the instanton separation. The case $\epsilon = 1$ corresponds to an instanton–anti-instanton pair starting from $q = q_0$ at time $-\infty$, approaching $q = 0$ at intermediate times, and returning to q_0 . The case $\epsilon = -1$ corresponds to a sequence of two instantons going from $-q_0$ to q_0 . Moreover, for the classical trajectory that, instead, goes from the origin to q_0 and back, we choose

$$q_2(t) = [f(t + \theta/2) + f(\theta/2 - t) - x_0]/\sqrt{g}. \quad (A42.13)$$

To calculate the classical action corresponding to $q_1(t)$, we separate the action into two parts, corresponding at leading order to the two instanton contributions:

$$\mathcal{S}(q_1) = \mathcal{S}_+(q_1) + \mathcal{S}_-(q_1), \quad (A42.14)$$

with

$$\begin{aligned} \mathcal{S}_+(q_1) &= \int_0^{+\infty} \left[\frac{1}{2} \dot{q}_1^2(t) + V(\sqrt{g}q_1(t))/g \right] dt, \\ \mathcal{S}_-(q_1) &= \int_{-\infty}^0 \left[\frac{1}{2} \dot{q}_1^2(t) + V(\sqrt{g}q_1(t))/g \right] dt. \end{aligned} \quad (A42.15)$$

The value $t = 0$ of the separation point is somewhat arbitrary and can be replaced by any value which remains finite when θ goes to infinity. We then use the properties that, for θ large, $f_+(t)$ is small for $t < 0$, and $f_-(t)$ is small for $t > 0$, to expand both terms. For example, for \mathcal{S}_+ , we find

$$\begin{aligned} \mathcal{S}_+(q_1) = \frac{1}{g} \int_0^{+\infty} dt & \left\{ \left[\frac{1}{2} \dot{f}_+^2(t) + V(f_+(t)) \right] + \epsilon \left[\dot{f}_-(t) \dot{f}_+(t) + V'(f_+(t)) f_-(t) \right] \right. \\ & \left. + \frac{1}{2} \left[\dot{f}_-^2(t) + V''(f_+(t)) f_-^2(t) \right] \right\}. \end{aligned} \quad (A42.16)$$

Since $f_-(t)$ decreases exponentially, only values of t small compared to $\theta/2$ contribute to the last term of equation (A42.16), which is proportional to V'' . For such values of t , we note that

$$\frac{1}{2} V''(f_+(t)) f_-^2(t) \sim V(f_-(t)). \quad (A42.17)$$

For the terms linear in $f_-(t)$, we integrate by parts the kinetic term, and use the equation of motion

$$\ddot{f}(t) = V'[f(t)]. \quad (A42.18)$$

Only the integrated term survives and yields

$$\int_0^{+\infty} dt \left[\dot{f}_-(t) \dot{f}_+(t) + V'(f_+(t)) f_-(t) \right] = -\dot{f}(-\theta/2) f(-\theta/2). \quad (A42.19)$$

The contribution \mathcal{S}_- can be evaluated by exactly the same method. We note that the sum of the two contributions reconstructs twice the classical action a . We then find

$$\mathcal{S}(q_1) = \frac{1}{g} \left[2a - 2\epsilon f(-\theta/2) \dot{f}(-\theta/2) + \dots \right], \quad (A42.20)$$

with

$$a = \int_0^{x_0} \sqrt{2V(x)} dx. \quad (A42.21)$$

Finally, replacing, for θ large, f by its asymptotic form (A42.9), we obtain the classical action

$$\mathcal{S}(q_1) = g^{-1} [2a - 2\epsilon C e^{-\theta} + O(e^{-2\theta})], \quad (A42.22)$$

and, therefore, the instanton interaction.

The calculation of the classical action corresponding to $q_2(t)$ follows the same steps, and one finds

$$\mathcal{S}(q_2) = \frac{1}{g} \left\{ 2a - 2[f(\theta/2) - x_0] \dot{f}(\theta/2) + \dots \right\}, \quad (A42.23)$$

expression which, for θ large, is equivalent to

$$\mathcal{S}(q_2) = \frac{1}{g} [2a - 2(C/\omega) e^{-\omega\theta}]. \quad (A42.24)$$

Finally, in the case of a finite time interval β with periodic boundary conditions, we can combine both results to find the action of a periodic trajectory passing close to $q = 0$ and $q = q_0$:

$$\mathcal{S}(q) = g^{-1} [2a - 2C (e^{-\beta+\theta} + e^{-\omega\theta}/\omega)], \quad (A42.25)$$

in agreement with equations (42.83, 42.86).

A42.2.1 Multi-instantons from constraints

Although multi-instanton configurations do not correspond to solutions of the equation of motion, it is nevertheless possible to modify the classical action by introducing constraints, and integrating over all possible constraints, generalizing the method of Section 37.4.1. The main problem with such a method is to find a set of constraints that are both theoretically acceptable and convenient for practical calculations.

One can, for instance, fix the positions of the instantons by introducing in the path integral (in the example of the double-well potential):

$$1 = \prod_{i=1}^n \left[\int dt \dot{q}_{\epsilon_i}^2(t - t_i) \right] \delta \left[\int dt \dot{q}_{\epsilon_i}(t - t_i)(q(t) - q_{\epsilon_i}(t - t_i)) \right] dt_i, \quad (A42.26)$$

where t_i are the instanton positions, and ϵ_i a successions of \pm indicating instantons and anti-instantons. One then uses an integral representation of the δ -functions, so that the path integral becomes

$$\begin{aligned} & \left(\frac{\|\dot{q}_+\|^2}{2i\pi} \right)^n \int \prod_{i=1}^n dt_i d\kappa_i \int [dq(t)] \prod_{i=1}^n \exp[-\mathcal{S}(q, \kappa_i)], \quad \text{with} \\ & \mathcal{S}(q, \kappa_i) = \mathcal{S}(q) + \sum_{i=1}^n \kappa_i \int dt \dot{q}_{\epsilon_i}(t - t_i)(q(t) - q_{\epsilon_i}(t - t_i)). \end{aligned} \quad (A42.27)$$

The arguments of Section 37.4.1 can then be generalized to recover the results of Section 42.1.

A42.3 A simple example of non-Borel summability

We illustrate here the problem of non-Borel summability with the example of a simple integral, which shares some of the features of the problem in QM which we have studied in the chapter. We consider the function

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dq \exp \left[-\frac{1}{g} V(q\sqrt{g}) \right], \quad (A42.28)$$

where $V(x)$ is an entire function with an absolute minimum at $x = 0$ with $V(0) = 0$. For g small, $I(g)$ can be calculated by steepest descent, expanding V around $q = 0$. This yields an expansion of the form,

$$I(g) = \sum_{k \geq 0} I_k g^k. \quad (A42.29)$$

One can express the integral (A42.28) as a generalized Borel or Laplace transform:

$$I(g) = \frac{1}{\sqrt{2\pi}} \int dq \int dt \delta[V(q\sqrt{g}) - t] e^{-t/g}. \quad (A42.30)$$

Interverting the q and t integration, we can integrate over q and find,

$$I(g) = \frac{1}{\sqrt{2\pi g}} \int_0^\infty dt e^{-t/g} \sum_i \frac{1}{|V'[x_i(t)]|}, \quad (A42.31)$$

in which $\{x_i(t)\}$ are the solutions of the equation

$$V[x_i(t)] = t. \quad (A42.32)$$

When the function $V(x)$ is monotonous both for x positive and negative, equation (A42.32) has two solutions for all values of t , and equation (A42.31) is directly the Borel representation of the function $I(g)$, which has a Borel summable power series expansion.

By contrast, we now assume that $V(x)$ has a second local minimum which gives a negligible contribution to $I(g)$ for g small. A simple example is

$$V(x) = \frac{1}{2}x^2 - \frac{1}{3a}x^3(1+a) + \frac{1}{4a}x^4, \quad \frac{1}{2} < a < 1, \quad (\text{A42.33})$$

which has a minimum at $x = 1$. Between its two minima, the potential $V(x)$ has a maximum, located at $x = a$, whose contribution dominates the large-order behaviour of the expansion in powers of g :

$$I_k \underset{k \rightarrow \infty}{\propto} \Gamma(k) [V(a)]^{-k}, \quad V(a) > 0, \quad (\text{A42.34})$$

(in the example (A42.33) $V(a) = a^2(1-a/2)/6$), and the series is not Borel summable.

The *naive* Borel transform of $I(g)$ is obtained by retaining in equation (A42.31) only the roots of equation (A42.32) that exist for t small. The singularities of the Borel transform then correspond to the zeros of $V''(x)$.

For the potential (A42.33), the expression (A42.31) has the form (θ is the step function)

$$I(g) = \frac{1}{\sqrt{2\pi g}} \int_0^{+\infty} dt e^{-t/g} \left[\frac{1}{|V'(x_1(t))|} + \frac{\theta(V(a)-t)}{|V'(x_2(t))|} + \frac{\theta(V(a)-t)\theta(t-V(1))}{|V'(x_3(t))|} + \frac{\theta(t-V(1))}{|V'(x_4(t))|} \right], \quad (\text{A42.35})$$

with the definitions (see Fig. 42.2) $x_1(t) \leq 0 \leq x_2(t) \leq a \leq x_3(t) \leq 1 \leq x_4(t)$.

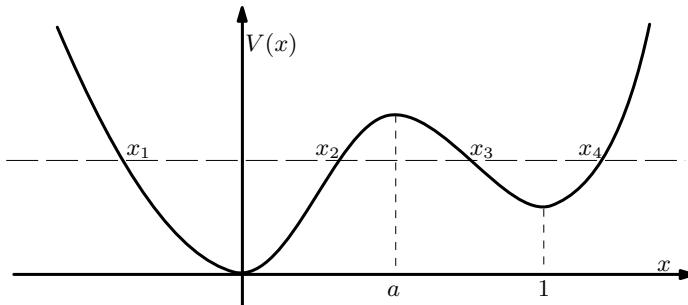


Fig. 42.2 The four roots of equation (A42.32)

The idea of the analytic continuation is to integrate each contribution up to $t = +\infty$ following a contour which passes below or above the cut along the positive real axis. This means that we consider $x_2(t)$ to be solution of the equation

$$V[x_2(t)] = t \pm i\epsilon. \quad (\text{A42.36})$$

The sign is arbitrary. Let us, for instance, choose the positive sign. We then have to subtract this additional contribution. We proceed in the same way for $x_3(t)$ for $t > V(a)$.

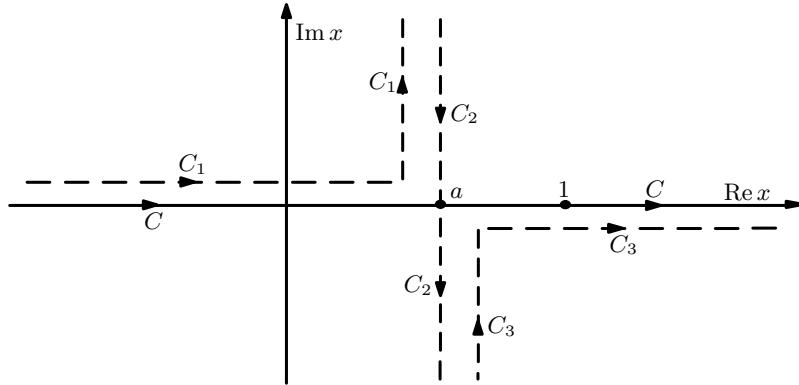


Fig. 42.3 The different contours in the x plane

Since $x_2(t)$ and $x_3(t)$ meet at $t = V(a)$, the analytic continuation corresponds to take for $x_3(t)$ the other solution of

$$V[x_3(t)] = t \mp i\epsilon. \quad (\text{A42.37})$$

Therefore, we have to subtract from the total expression the contributions of two roots of the equation. But it is easy to verify that this is just the contribution of the saddle point located at $x = a$, which corresponds to a maximum of the potential.

Therefore, we have succeeded in writing expression (A42.35) as the sum of three saddle point contributions (see Fig. 42.3). There is some arbitrariness in the decomposition, which here corresponds to the choice $\epsilon = \pm 1$.

In the complex x plane, we have replaced the initial contour C on the real positive axis, by a sum of three contours C_1 , C_2 , and C_3 , corresponding to the three saddle points located at $0, a, 1$.

A42.4 Multi-instantons and WKB approximation

We now give a more general form of the conjecture presented in Section 42.5, and indicate how a few properties of the functions $A(E, g)$ and $D(E, g)$ can be obtained from the ‘exact’ WKB expansion [456].

A42.4.1 The conjecture

The double-well potential. For the double-well potential, we have conjectured the spectral equation (42.127),

$$\frac{1}{\sqrt{2\pi}} \Gamma\left(\frac{1}{2} - D\right) (-2/g)^{D(E,g)} e^{-A(g,E)/2} = \epsilon i, \quad (\text{A42.38})$$

where the functions $D(E, g)$ and $A(E, g)$ have the expansions (42.125) and (42.128), respectively. The conjecture can be generalized to potentials with two asymmetric wells:

$$\begin{aligned} & \Gamma\left(\frac{1}{2} - D_1(E, g)\right) \left(-\frac{2C_1}{g}\right)^{D_1(E,g)} \Gamma\left(\frac{1}{2} - D_2(E, g)\right) \left(-\frac{2C_2}{g}\right)^{D_2(E,g)} \frac{e^{-A(g,E)}}{2\pi} \\ &= -1, \end{aligned} \quad (\text{A42.39})$$

where $D_1(E, g)$ and $D_2(E, g)$ are determined by the perturbative expansions around each of the two minima of the potential, and C_1, C_2 are numerical constants.

The cosine potential. For the cosine potential $\frac{1}{16}(1 - \cos 4q)$, the conjecture reads

$$\left(\frac{2}{g}\right)^{-D(E,g)} \frac{e^{A(E,g)/2}}{\Gamma(\frac{1}{2} - D(E,g))} + \left(\frac{-2}{g}\right)^{D(E,g)} \frac{e^{-A(g,E)/2}}{\Gamma(\frac{1}{2} + D(E,g))} = \frac{2 \cos \varphi}{\sqrt{2\pi}}. \quad (A42.40)$$

The $O(\nu)$ -symmetric anharmonic oscillator. For the $O(\nu)$ anharmonic oscillator, the conjecture

$$i e^{-A(E,g)} \left(-\frac{2}{g}\right)^D e^{i\pi(D+l+\nu/2)/2} \frac{\Gamma(\frac{1}{2}(l + \nu/2 - D))}{\Gamma(\frac{1}{2}(l + \nu/2 + D))} = 1, \quad (A42.41)$$

corresponds to the expansion of the energy levels for $g < 0$ and, in particular, yields the instanton contributions to the large-order behaviour.

A42.4.2 The WKB expansion

We consider the Schrödinger equation,

$$-\frac{1}{2}\psi''(q) + (V(q\sqrt{g})/g)\psi(q) = E\psi(q). \quad (A42.42)$$

We first assume that $q = 0$ is the absolute minimum of $V(q)$ and, moreover, $V(q) \sim \frac{1}{2}q^2$, for q small. In terms of $x = \sqrt{g}q$, the equation can be rewritten as

$$-g^2\psi''(x) + 2V(x)\psi(x) = 2gE\psi(x). \quad (A42.43)$$

The WKB expansion is an expansion for $g \rightarrow 0$ at Eg fixed, in contrast with the perturbative expansion where E is fixed. It can be constructed by going over to the corresponding Riccati equation, setting

$$S(x) = -g\psi'(x)/\psi(x), \quad (A42.44)$$

where S satisfies

$$gS'(x) - S^2(x) + S_0^2(x) = 0, \quad S_0^2(x) = 2V(x) - 2gE. \quad (A42.45)$$

One then expands systematically in powers of g , at Eg fixed, starting from $S(x) = S_0(x)$. It is convenient to decompose $S(x)$ into an odd and even part, setting,

$$S(x, g, E) = S_+(x, g, E) + S_-(x, g, E), \quad S_{\pm}(x, -g, -E) = \pm S_{\pm}(x, g, E). \quad (A42.46)$$

It follows that

$$gS'_-(x) - S_+^2(x) - S_-^2(x) + S_0^2(x) = 0, \quad (A42.47)$$

$$gS'_+(x) - 2S_+(x)S_-(x) = 0. \quad (A42.48)$$

It is then possible to express the wave function in terms of S_+ only:

$$\psi(x) = (S_+(x))^{-1/2} \exp\left[-\frac{1}{g} \int_{x_0}^x dx' S_+(x')\right]. \quad (A42.49)$$

The spectrum can then be determined by the condition

$$\frac{1}{2i\pi} \oint_C dz \frac{\psi'(z)}{\psi(z)} = N, \quad (A42.50)$$

where N is the number of nodes of the eigenfunction, and C a contour which encloses them.

In the semi-classical limit, C encloses the cut of $S_0(x)$, which joins the two turning points solutions of $S_0(x) = 0$ (x_1, x_2 in Fig. 42.2). In terms of S_+ , equation (A42.50) becomes

$$-\frac{1}{2i\pi g} \oint_C dz S_+(z) = N + \frac{1}{2}. \quad (\text{A42.51})$$

If we replace S_+ by its WKB expansion, and expand each term in a power series of Eg , we obtain the function $D(E, g)$ (the perturbative expansion):

$$-\frac{1}{2i\pi g} \oint_C dz S_+(z) = D(E, g) = -D(-E, -g). \quad (\text{A42.52})$$

Potentials with degenerate minima. In the case of potentials with degenerate minima, two functions D_1 and D_2 (in the notation of equation (A42.39)) appear, corresponding to the expansions around each minimum. An additional contour integral arises corresponding to barrier penetration effects. The expansion for Eg small of its WKB expansion yields the function $A(g, E)$:

$$\begin{aligned} \frac{1}{g} \oint_{C'} dz S_+(z) &= A(E, g) + \ln(2\pi) - \sum_{i=1}^2 \ln \Gamma\left(\frac{1}{2} - D_i(E, g)\right) \\ &\quad + D_i(E, g) \ln(-2g/C_i), \end{aligned} \quad (\text{A42.53})$$

where C' encloses $[x_2, x_3]$ in Fig. 42.2. In the WKB expansion, the functions $\Gamma(\frac{1}{2} - D_i)$ have to be replaced by their asymptotic expansion for D_i large. Still a calculation of $A(E, g)$ at a finite order in g requires the WKB expansion and the asymptotic expansion of the Γ function only at a finite order.

$O(\nu)$ -symmetric potentials. These expressions can be generalized to the case of $O(\nu)$ -symmetric potentials. The perturbative expansion can be obtained by inverting a relation of the form

$$\mu + 2N + 1 = D(E, g, \mu), \quad \mu = l + \nu/2 - 1, \quad (\text{A42.54})$$

where the function $D(E, g, \mu)$ is given by a contour integral surrounding all zeros of the wave function on the real axis (including the negative real axis) of the even part (in the sense of equation (A42.46)) of $-g(\psi'_l/\psi_l + (\nu - 1)/2|q|)$. The following properties can then be verified

$$D(E, g, \mu) = -D(-E, -g, \mu), \quad D(E, g, \mu) = D(E, g, -\mu), \quad (\text{A42.55})$$

and the coefficient of order g^k in the expansion of D is a polynomial of degree $[(k+1)/2]$ in μ . In the WKB expansion, the functions D and A again correspond to different contour integrals around turning points.

Bibliography

- [1] G.C. Wick, *Phys. Rev.* 80 (1950) 268.
- [2] F.A. Berezin, *The Method of Second Quantization* (Akad. Nauk, Moscow 1965) translated from Russian by N. Mugibayashi and A. Jeffrey (Academic Press, New York 1966), reprinted (Elsevier 2012).
- [3] F. Bloch, *Zeit. f. Phys.* 74 (1932) 295.
- [4] P.A.M. Dirac, *Physik. Z. Sowjetunion* 3 (1933) 64;
R.P. Feynman, *Rev. Mod. Phys.* 367 (1948) 20; *Phys. Rev.* 80 (1950) 440;
R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York 1965).
- [5] E.S. Abers, B.W. Lee, *Phys. Rep.* 9C (1973) 1.
- [6] J. Zinn-Justin, *Intégrale de chemin en mécanique quantique: introduction* (EDP Sciences and CNRS Editions, Les Ulis 2003); English version, *Path integrals in Quantum Mechanics* (Oxford Univ. Press 2005); Russian translation (Fizmatlit, Moscow 2006).
- [7] L.S. Schulman, *Techniques and application of path integration* (Dover Pub., Mineola 2005).
- [8] S.G. Brush, *Rev. Mod. Phys.* 33 (1961) 79;
J. Ginibre, *Some applications of functional integration in statistical mechanics*, Les Houches physics school *Statistical Mechanics and Quantum Field Theory* 1970, p.327, C. De Witt, R. Stora eds. (Gordon and Breach, New York 1971).
- [9] J. Schwinger, *Phys. Rev.* 82 (1951) 914; *PNAS* 37 (1951) 452.
- [10] G. Källen, *Quantum Electrodynamics* (Springer, Berlin 1972).
- [11] R.P. Feynman, *Phys. Rev.* 84 (1951) 108 (Appendix B);
C. Garrod, *Rev. Mod. Phys.* 38 (1966) 483;
F.A. Berezin, *Theor. Math. Phys.* 6 (1971) 194.
- [12] *Methods in Field Theory*, Les Houches physics school 1975, R. Balian, J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).
- [13] The generalization to constrained systems is given in
L.D. Faddeev, *Theor. Math. Phys.* 1 (1969) 3.
- [14] An early discussion of the quantization problem for velocity-dependent potentials, in the operator formalism, can be found in
T.D. Lee, C.N. Yang, *Phys. Rev.* 128 (1962) 885, Appendix C.
- [15] V. Bargmann, *Comm. Pure and Appl. Math.* 14 (1961) 187.
- [16] J.R. Klauder, *Ann. Phys.* 11 (1960) 123, where integrals over complex phase space variables are introduced.
- [17] S.S. Schweber, *J. Math. Phys.* 3 (1962) 831, where a general formulation is found and the relation with the holomorphic formalism of Ref. [15] is shown explicitly.
- [18] See Chapter 7 in Ref. [6].

- [19] V.N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics* (Reidel Publishing, Dordrecht 1983).
- [20] J. Solyom, *Advances in Physics* 28 (1979) 201.
- [21] L.D. Faddeev's contribution to Ref. [12], p.1-77.
- [22] H.D.I. Abarbanel, *Cargèse Lectures in Physics* vol. 5, D. Bessis ed. (Gordon and Breach, New York 1972).
- [23] E. Brézin, C. Itzykson, J. Zinn-Justin, *Phys. Rev.* D1 (1970) 2349.
- [24] F.J. Dyson, *Phys. Rev.* 75 (1949) 486.
- [25] N.N. Bogoliubov, D.V. Shirkov, *Introduction to Theory of Quantized Fields* (Interscience, New York, London, 1959).
- [26] R.P. Feynman, *Acta Phys. Polon.* 24 (1963) 697; *Magic without Magic: J.A. Wheeler, J. Klauder* ed. (Freeman, San Francisco 1972).
- [27] B. Zumino in *Theory and Phenomenology in Particle Physics*, Erice 1968, A. Zichichi ed. (Academic Press, New York 1968).
- [28] L.S. Brown, *Quantum Field Theory*, p.129 (Cambridge Univ. Press 1992).
- [29] A.N. Vasiliev, *Functional Methods in Quantum Field Theory and Statistical Mechanics*, (St Petersburg Univ. Press 1976), translated by P.A. Millard (Gordon and Breach, Amsterdam 1998).
- [30] J. Schwinger, *PNAS* 37 (1951) 455.
- [31] The Legendre transform of the generating functional of connected correlation functions has been discussed in the context of statistical physics in
 C. De Dominicis, *J. Math. Phys.* 4 (1963) 255;
 C. De Dominicis, P.C. Martin, *J. Math. Phys.* 5 (1964) 14, 31.
 It has been introduced in quantum field theory in
 G. Jona-Lasinio, *Nuovo Cimento* 34 (1964) 1790.
- [32] The loop expansion is discussed in the context of quantum gravity in
 B.S. de Witt, *Phys. Rev.* 162 (1967) 1195, 1239.
- [33] Y. Nambu, *Phys. Lett.* 26B (1968) 626.
- [34] S. Tomonaga, *Prog. Theor. Phys.* 1 (1946) 27.
- [35] J. Schwinger, *Phys. Rev.* 74 (1948) 1439; *Phys. Rev.* 75 (1949) 651; *Phys. Rev.* 76 (1949) 790.
- [36] F.J. Dyson, *Phys. Rev.* 75 (1949) 486, 1736.
- [37] R.P. Feynman, *Phys. Rev.* 76 (1949) 749, 769.
- [38] F. Englert, C. De Dominicis, *Nuovo Cimento* 53A (1968) 1007, 1021.
- [39] A.A. Slavnov, *Theor. Math. Phys.* 22 (1975) 123.
- [40] J. Zinn-Justin in *Trends in Elementary Particle Physics*, Bonn summer institute 1974, H. Rollnik, K. Dietz eds. *Lectures Notes in Physics* 37 (1975) 2-39 (Springer, Berlin).
- [41] B.W. Lee, J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3121.
- [42] R. Jackiw, *Phys. Rev.* D9 (1974) 1686.
- [43] S. Coleman, E. Weinberg, *Phys. Rev.* D7 (1973) 1888.

- [44] S. Coleman, *Laws of Hadronic Matter* (discussion of lecture 3), Erice 1973, A. Zichichi ed. (Academic Press, New York 1975).
- [45] K. Symanzik, *Comm. Math. Phys.* 16 (1970) 48; see also Ref. [44].
- [46] The use of the solution of the classical field equation as argument of the generating functional or background field method (see Section A7.2) was suggested in B.S. DeWitt in *Relativity, Groups and Topology*, Les Houches physics school 1963, C. DeWitt, B. DeWitt eds. (Gordon and Breach, New York 1964); *Phys. Rev.* 162 (1967) 1195;
G. 't Hooft, *Nucl. Phys.* B62 (1973) 444.
- [47] The decoupling of non-critical or massive modes is discussed in
K. Symanzik, *Acta Phys. Austr. Suppl.* XI (1973) 199;
T. Applequist, J. Carazzone, *Phys. Rev.* D11 (1975) 2856.
- [48] W. Pauli, F. Villars, *Rev. Mod. Phys.* 21 (1949) 434.
- [49] C.N. Yang, T.D. Lee, *Phys. Rev.* 87 (1952) 404;
T.D. Lee, C.N. Yang, *Phys. Rev.* 87 (1952) 410;
M.E. Fisher, *Phys. Rev. Lett.* 40 (1978) 1610.
- [50] *Phase Transitions and Critical Phenomena*, vol. 6, C. Domb, M.S. Green eds. (Academic Press, New York 1976).
- [51] E. Brézin, J.C. Le Guillou, J. Zinn-Justin in Ref. [50], p.125-247.
- [52] The consistency of the lattice regularization is proven (except for chiral fermions) in
T. Reisz, *Comm. Math. Phys.* 117 (1988) 79, 639.
- [53] J. Schwinger, *Phys. Rev.* 82 (1951) 664.
- [54] N.N. Bogoliubov, O.S. Parasiuk, *Acta Math.* 97 (1957) 227.
- [55] K.G. Wilson, *Phys. Rev.* 179 (1969) 1499.
- [56] E.C.G. Stueckelberg, A. Peterman, *Helv. Phys. Acta* 26 (1953) 499;
M. Gell-Mann, F.E. Low, *Phys. Rev.* 95 (1954) 1300.
- [57] C.G. Callan's contribution to Ref. [12], p.41-77.
- [58] Combining the proof of renormalizability and of RG equations is proposed in
A.S. Blaer, K. Young, *Nucl. Phys.* B83 (1974) 493.
- [59] J. Polchinski, *Nucl. Phys.* B231 (1984) 269.
- [60] C.G. Callan, *Phys. Rev.* D2 (1970) 1541;
K. Symanzik, *Comm. Math. Phys.* 18 (1970) 227.
- [61] F.J. Wegner, A. Houghton, *Phys. Rev.* A8 (1973) 401;
F.J. Wegner, *J. Physics* C7 (1974) 2098.
- [62] K.G. Wilson, J.B. Kogut, *Phys. Rep.* 12 (1974) 75.
- [63] F.J. Wegner's contribution to Ref. [50], p.7-124.
- [64] J. Zinn-Justin, *Transitions de phase et groupe de renormalisation*, EDP Sciences et CNRS Editions (Les Ulis 2005), English version *Phase transitions and Renormalization Group* (Oxford Univ. Press 2007).
- [65] The asymptotic behaviour of renormalized integrals was studied in
S. Weinberg, *Phys. Rev.* 118 (1960) 838;
K. Hepp, *Comm. Math. Phys.* 2 (1966) 301.

- [66] E.R. Speer, *J. Math. Phys.* 9 (1967) 1404;
 K. Hepp, *Comm. Math. Phys.* 2 (1966) 301;
 H. Epstein, V. Glaser, *Annales de l'IHP*, section A19 (1973) 211.
- [67] W. Zimmermann, *Lectures on Elementary Particles and Quantum Field Theory, Brandeis 1970*, S. Deser, M. Grisaru, H. Pendleton eds. (MIT Press, Cambridge 1970), where a systematic discussion of renormalization theory in different terminology and with different techniques, as well as historical references can be found.
- [68] T. Appelquist, *Ann. Phys. (NY)* 54 (1969) 27;
 M.C. Bergère, J.-B. Zuber, *Comm. Math. Phys.* 35 (1974) 113;
 M. Bergère, Y.M.P. Lam, *Comm. Math. Phys.* 39 (1974) 1.
- [69] K. Symanzik, *Comm. Math. Phys.* 7 (1973) 34.
- [70] S. Weinberg, *Phys. Rev.* D8 (1973) 3497.
- [71] G. 't Hooft, *Nucl. Phys.* B61 (1973) 455.
- [72] J. Zinn-Justin, *Cargèse Lectures 1973*, Saclay preprint T73/049 unpublished, where the bare RG equations are derived, material later incorporated into the review [51].
- [73] For an early, rigorous discussion of the issue, see
 J. Fröhlich, *Nucl. Phys.* B200 (1982) 281.
- [74] A pedagogical presentation has been given in
 S. Coleman, *Dilatations, Erice Lectures 1971* reprinted in *Aspects of Symmetry* (Cambridge Univ. Press 1985).
- [75] K.G. Wilson, M.E. Fisher, *Phys. Rev. Lett.* 28 (1972) 240.
- [76] J. Ashmore, *Lett. Nuovo Cimento* 4 (1972) 289;
 G. 't Hooft, M. Veltman, *Nucl. Phys.* B44 (1972) 189;
 C.G. Bollini, J.J. Giambiagi, *Phys. Lett.* 40B (1972) 566; *Nuovo Cimento* 12B (1972) 20;
 G. 't Hooft, M. Veltman, in *Particle Interactions at Very High Energies*, D. Speiser, F. Halzen, J. Weyers eds. (Plenum Press, New York 1973) vol. 4B;
 G. Leibbrandt, *Rev. Mod. Phys.* 47 (1975) 849.
- [77] J.C. Collins, *Nucl. Phys.* B80 (1974) 341.
- [78] The properties of renormalization-group functions in the minimal subtraction scheme are discussed in Refs. [71] and [77].
- [79] For a pedagogical presentation see D.J. Gross's contribution to Ref. [12], p.141-250.
- [80] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev.* D8 (1973) 434, 2418.
- [81] A.A. Vladimirov, *Teor. Mat. Fiz.* 43 (1980) 210, (*Theor. Math. Phys.* 43 (1980) 417);
 A.A. Vladimirov, D.I. Kazakov, O.V. Tarasov, *Zh. Eksp. Teor. Fiz.* 77 (1979) 1035, (*Sov. Phys. JETP* 50 (1979) 521).
- [82] W. Zimmermann, *Ann. Phys. (NY)* 77 (1973) 536.
- [83] E. Brézin, C. De Dominicis, J. Zinn-Justin, *J. Lett. Nuovo Cimento* 9 (1974) 483.
- [84] The short distance expansion was postulated in Ref. [55] and
 L.P. Kadanoff, *Phys. Rev. Lett.* 23 (1969) 1430;
 K.G. Wilson, *Phys. Rev.* D3 (1971) 1818.
- [85] K.G. Wilson, W. Zimmermann, *Comm. Math. Phys.* 24 (1972) 87.

- [86] K. Symanzik, *Comm. Math. Phys.* 18 (1970) 227.
- [87] W. Zimmermann, *Ann. Phys. (NY)* 77 (1973) 570.
- [88] RG equations for the leading coefficients of the SDE are established in
K. Symanzik, *Comm. Math. Phys.* 23 (1971) 49;
C.G. Callan, *Phys. Rev. D* 5 (1972) 3202.
- [89] The extension of Wilson's work to the light-cone expansion is found in
R.A. Brandt, G. Preparata, *Nucl. Phys.* B27 (1972) 541.
- [90] K. Symanzik, *Nucl. Phys.* B226 (1983) 187, 205.
- [91] S. El-Showk *et al*, *Phys. Rev. D* 86 (2012) 025022.
- [92] Y. Choquet-Bruhat, C. DeWitt-Morette, *Analysis, Manifolds and Physics: Part II* (North-Holland, Amsterdam 1989).
Concerning spin in four dimensions, see also
P. Moussa and R. Stora in *Methods in Subnuclear Physics*, M. Nikolic ed. (Gordon and Breach New York 1966).
P. Moussa in *Particle Physics*, Les Houches 1971, C. De Witt, C. Itzykson eds. (Gordon and Breach, New York 1973).
- [93] L.L. Foldy, S.A. Wouthuysen, *Phys. Rev.* 78 (1950) 29.
- [94] For a discussion of the *PCT* theorem and the spin-statistics connection, starting from first principles, see
R.F. Streater, A.S. Wightman, *PCT, Spin & Statistics and All That* (Benjamin, New York 1964).
- [95] H.B. Nielsen, M. Ninomiya, *Nucl. Phys.* B185 (1981) 20.
- [96] K.G. Wilson in *New Phenomena in Subnuclear Physics*, Erice 1975, A. Zichichi ed. (Plenum, New York 1977).
- [97] T. Banks, L. Susskind, J. Kogut, *Phys. Rev. D* 13 (1976) 1043.
- [98] P.H. Ginsparg, K.G. Wilson, *Phys. Rev. D* 25 (1982) 2649.
- [99] H. Neuberger, *Phys. Lett.* B417 (1998) 141; *Phys. Lett.* B427 (1998) 353;
P. Hasenfratz, V. Laliena, F. Niedermayer, *Phys. Lett.* B427 (1998) 125.
- [100] M. Lüscher, *Phys. Lett.* B428 (1998) 342.
- [101] A.N. Vasiliev (in French) in *Cargèse Lectures in Physics* vol. 5, Cargèse Summer Institute 1970, D. Bessis ed. (Gordon and Breach, New York 1971), later incorporated into Ref. [29].
- [102] For early derivations of the Nambu-Goldstone theorem, see
Y. Nambu, *Phys. Rev. Lett.* 4 (1960) 380;
J. Goldstone, *Nuovo Cimento* 19 (1961) 155;
Y. Nambu, G. Jona-Lasinio, *Phys. Rev.* 122 (1961) 345; *Phys. Rev.* 124 (1961) 246;
J. Goldstone, A. Salam, S. Weinberg, *Phys. Rev.* 127 (1962) 965;
S. Bludman, A. Klein, *Phys. Rev.* 131 (1963) 2364;
G. Jona-Lasinio, *Nuovo Cimento* 34 (1964) 1790.
- [103] The linear σ -model has been proposed in
M. Gell-Mann, M. Lévy, *Nuovo Cimento* 16 (1960) 705.

- [104] B.W. Lee, *Nucl. Phys.* B9 (1969) 649;
 J.-L. Gervais, B.W. Lee, *Nucl. Phys.* B12 (1969) 627;
 J.L. Basdevant, B.W. Lee, *Phys. Rev.* D2 (1970) 1680.
- [105] K. Symanzik, *Comm. Math. Phys.* 16 (1970) 48.
- [106] B.W. Lee, *Chiral Dynamics* (Gordon and Breach, New York 1972).
- [107] S. Coleman in *Laws of Hadronic Matter*, Erice 1973, A. Zichichi ed. (Academic Press, New York 1975).
- [108] D. Bessis, M. Pusterla, *Nuovo Cimento* 54A (1968) 243.
- [109] G.A. Baker and P. Graves-Morris, *Encyclopedia of Mathematics and its Applications* vols. 13 and 14, Gian-Carlo Rota ed. (Addison-Wesley, New York 1981).
- [110] J. Zinn-Justin, *Phys. Rep.* 1C (1970) 55.
- [111] J. Gasser, H. Leutwyler, *Annals of Physics* 158 (1984) 142.
- [112] M.L. Goldberger, S.B. Treiman, *Phys. Rev.* 110 (1958) 1178;
 S.L. Adler, *Phys. Rev.* 139 (1965) B1638; *Phys. Rev.* 140 (1965) B736;
 W.I. Weisberger, *Phys. Rev.* 143 (1966) 1302;
 S. Weinberg, *Phys. Rev. Lett.* 17 (1966) 616;
 $SU(3) \times SU(3)$ symmetry breaking is discussed in
 M. Gell-Mann, R.J. Oakes, B. Renner, *Phys. Rev.* 175 (1968) 2195.
- [113] C.G. Callan, S. Coleman, R. Jackiw, *Ann. Phys. (NY)* 59 (1970) 42.
- [114] Conformal invariance, and its spectacular consequences in two dimensions, are thoroughly discussed in *Fields, Strings and Critical Phenomena*, Les Houches physics school 1988 E. Brézin, J. Zinn-Justin eds. (Elsevier, Amsterdam 1989). See, in particular, the lectures
 P. Ginsparg, *Applied Conformal Field Theory* p.1-168;
 J. Cardy, *Conformal Invariance and Statistical Mechanics* p.169-246.
- [115] It has been pointed out that at a zero of the RG β -function both dilatation and conformal invariance are restored in
 B. Schroer, *Lett. Nuovo Cimento* 2 (1971) 867.
- [116] P. Weiss, *J. Phys. Radium, Paris* 6 (1907) 667;
 M.E. Fisher, *Rep. Prog. Phys.* 30 (1967) 615;
 L.P. Kadanoff *et al*, *Rev. Mod. Phys.* 39 (1967) 395.
- [117] L.D. Landau, *Phys. Z. Sowjetunion* 11 (1937) 26, reprinted in *Collected Papers of L.D. Landau*, D. ter Haar ed. (Pergamon, New York 1965).
- [118] The impossibility of spontaneous breaking of continuous symmetries in one and two dimensions has been proven by
 N.D. Mermin, H. Wagner, *Phys. Rev. Lett.* 17 (1966) 1133;
 S. Coleman, *Comm. Math. Phys.* 31 (1973) 259.
- [119] V.L. Ginzburg, *Soviet Physics-Solid State* 2 (1961) 1824.
- [120] K.G. Wilson, *Phys. Rev.* B4 (1971) 3184.
- [121] For a proof of spontaneous symmetry breaking of discrete symmetries, see
 R. Peierls, *Phys. Rev.* 54 (1938) 918.

- [122] B. Widom, *J. Chem. Phys.* 43 (1965) 3892, 3898;
C. Domb, D.L. Hunter, *Proc. Phys. Soc. (London)* 86 (1965) 1147.
- [123] E.H. Graf, D.M. Lee, J.D. Reppe, *Phys. Rev. Lett.* 19 (1967) 417;
M. Blume, *Phys. Rev.* 141 (1966) 517;
H.W. Capel, *Physica* 32 (1966) 966.
- [124] G. Horwitz, H.B. Callen, *Phys. Rev.* 124 (1961) 1757;
F. Englert, *Phys. Rev.* 129 (1963) 567.
- [125] Section A14.2 directly follows from
J. Zinn-Justin, in *Perspectives in Particles and Fields*, Cargèse 1983, vol. B126 p.15-43, M. Lévy *et al.* eds. (Springer, Boston 1985).
- [126] S.F. Edwards, P.W. Anderson, *J. Phys.* F5 (1975) 965.
- [127] M. Mézard, G. Parisi, M. Virasoro, *Spin Glass and Beyond* (World Scientific, Singapore 1987).
- [128] The idea of RG transformations was earlier proposed in a simplified form in
L.P. Kadanoff, *Physics* 2 (1966) 263.
- [129] K.G. Wilson, *Phys. Rev.* B4 (1971) 3174.
- [130] C. Wetterich, *Phys. Lett.* B301 (1993) 90;
J. Berges, N. Tetradis, C. Wetterich, *Phys. Rep.* 363 (2002) 223.
- [131] F.J. Wegner, *Phys. Rev.* B5 (1972) 4529; *Phys. Rev.* B6 (1972) 1891.
- [132] P.K. Mitter, *Phys. Rev.* D7 (1973) 2927.
- [133] C. Di Castro, *Lett. Nuovo Cimento.* 5 (1972) 69;
B. Schroer, *Phys. Rev.* B8 (1973) 4200;
C. Di Castro, G. Jona-Lasinio, L. Peliti, *Ann. Phys. (NY)* 87 (1974) 327;
F. Jegerlehner, B. Schroer, *Acta Phys. Austr. Suppl.* XI (1973) 389 (Springer, Berlin).
- [134] G. Parisi, *Statistical Field Theory* (Addison-Wesley, New York 1988).
- [135] S.K. Ma, *Modern Theory of Critical Phenomena* (Taylor & Francis, New York, 2000).
- [136] P.G. de Gennes, *Phys. Lett.* 38A (1972) 339.
- [137] J. des Cloizeaux, G. Jannink, *Les Polymères en Solution* (Editions de Physique, Paris 1987), English version *Polymers in Solution: Modelling and Structure*, (Oxford Univ. Press 1990).
- [138] V.J. Emery, *Phys. Rev.* B11 (1975) 239;
B. Duplantier, *C.R. Acad. Sci. Paris* 290B (1980) 199.
- [139] J. des Cloizeaux, *J. Physique (Paris)* 36 (1975) 281;
L. Schäfer, T.A. Witten, *J. Chem. Phys.* 66 (1977) 2121; *J. Physique (Paris)* 41 (1980) 459.
- [140] T.D. Lee, C.N. Yang, *Phys. Rev.* 87 (1952) 410.
- [141] The derivation of Section 15.9.1 has been inspired by the review of the application of field integrals to classical and quantum Bose fluids,
F.W. Wiegel, *Phys. Rep.* 16C (1975) 57.
- [142] A description of helium's superfluid transition in terms of the two-component ϕ^4 field theory, in the framework of Landau's theory, can be found in

- V.L. Ginzburg, L.P. Pitaevskii, *Zh. Eksperim. i Teor. Fiz.* 34 (1958) 1240 (*Soviet Phys. JETP* 34 (1958) 858).
- [143] L.S. Goldner, G. Ahlers, *Phys. Rev.* B45 (1992) 13129;
J.A. Lipa, D.R. Swanson, J. Nissen, T.C.P. Chui, U.E. Israelson, *Phys. Rev. Lett.* 76 (1996) 944.
- [144] E.K. Riedel, F.J. Wegner, *Phys. Rev. Lett.* 29 (1972) 349.
- [145] G. Baym, J.-P. Blaizot, J. Zinn-Justin, *Euro. Phys. Lett.* 49 (2000) 150.
- [146] G. Parisi, *Cargèse Lectures 1973*, published in *J. Stat. Phys.* 23 (1980) 49.
- [147] E. Brézin, D.J. Wallace, K.G. Wilson, *Phys. Rev. Lett.* 29 (1972) 591; *Phys. Rev. B7* (1973) 232;
G.M. Avdeeva, A.A. Migdal, *JETP Lett.* 16 (1972) 178.
- [148] R.B. Griffith, *Phys. Rev.* 158 (1967) 176.
- [149] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev. B10* (1974) 892.
- [150] D.J. Wallace, *J. Phys. C: Solid State Phys.* 6 (1973) 1390;
A. Aharony, *Phys. Rev. B8* (1973) 4270;
I.J. Ketley, D.J. Wallace, *J. Phys. A: Math. Nucl. Gen.* 6 (1973) 1667.
- [151] L. Michel, *Phys. Rev. B29* (1984) 2777.
- [152] E. Vicari, J. Zinn-Justin, *New J. Phys.* 8 (2006) 321.
- [153] E. Riedel, F. Wegner, *Z. Phys.* 225 (1969) 195;
M.E. Fisher, P. Pfeuty, *Phys. Rev. B6* (1972) 1889;
K.G. Wilson, *Phys. Rev. Lett.* 28 (1972) 548;
Y. Yamazaki, M. Suzuki, *Prog. Theor. Phys.* 50 (1973) 1097;
Y. Yamazaki, *Phys. Lett.* 49A (1974) 215.
- [154] D.J. Wallace, R.K.P. Zia, *Ann. Phys.* 92 (1975) 42.
- [155] E. Brézin, D. Amit, J. Zinn-Justin, *Phys. Rev. Lett.* 32 (1974) 151;
E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev. Lett.* 32 (1974) 473.
- [156] M.E. Fisher, J.S. Langer, *Phys. Rev. Lett.* 20 (1968) 665;
M.E. Fisher, A. Aharony, *Phys. Rev. B10* (1974) 2818.
- [157] K.G. Wilson, *Phys. Rev. Lett.* 28 (1972) 548.
- [158] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, B.G. Nickel, *Phys. Lett.* 44A (1973) 227.
- [159] S.G. Gorishny, S.A. Larin, F.V. Tkachov, *Phys. Lett.* 101A (1984) 120;
H. Kleinert, J. Neu, V. Schulte-Frohlinde, K.G. Chetyrkin, S.A. Larin, *Phys. Lett.* B272 (1991) 39, Erratum *Phys. Lett.* B319 (1993) 545.
- [160] M.K. Kompaniets, E. Panzer, *Phys. Rev.* D96 (2017) 036016.
- [161] O Schnetz, *Phys. Rev.* D97 (2018) 085018.
- [162] D.J. Wallace, R.P.K. Zia, *J. Phys. C* 7 (1974) 3480.
- [163] B.D. Josephson, *J. Phys. C* 2 (1969) 1113;
P. Schofield, J.D. Litster, J.T. Ho, *Phys. Rev. Lett.* 23 (1969) 1098.
- [164] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Lett.* 47A (1974) 285;
H.B. Tarko, M.E. Fisher, *Phys. Rev. Lett.* 31 (1973); *Phys. Rev. B11* (1975) 1217.
- [165] M.E. Fisher, H.B. Tarko, *Phys. Rev. B11* (1975) 1131.

- [166] A. Aharony, P.C. Hohenberg, *Phys. Rev.* B13 (1976) 3081; *Physica* 86-88B (1977) 611.
- [167] Y. Okabe, K. Ideura, *Prog. Theor. Phys.* 66 (1981) 1959;
J.F. Nicoll, P.C. Albright, *Phys. Rev.* B31 (1985) 4576.
- [168] F. Wegner, *Phys. Rev.* B5 (1972) 4529; *Phys. Rev.* B6 (1972) 1891, and the contribution to Ref. [50].
- [169] F.J. Wegner, E.K. Riedel, *Phys. Rev.* B7 (1973) 248.
- [170] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, Section VIII p.193 of Ref. [51].
- [171] S.K. Kehrein, F. Wegner, *Nucl. Phys.* B424 (1994) 521.
- [172] M.E. Fisher, S.K. Ma, B.G. Nickel, *Phys. Rev. Lett.* 29 (1972) 917.
- [173] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *J. Phys. A: Math. Gen.* 9 (1976) L119.
- [174] The logarithmic corrections in four dimensions, for the ϕ^4 theory, and for the uniaxial system with dipolar forces have been obtained, with the use of the parquet approximation, by
A.I. Larkin, D.E. Khmel'nitskii, *Zh. Eksp. Teor. Fiz.* 56 (1969) 647 (*JETP* 29, 1123).
- [175] A. Aharony, *Phys. Rev.* B8 (1973) 3363;
E. Brézin, J. Zinn-Justin, *Phys. Rev.* B13 (1976) 251.
- [176] G. Ahlers, A. Kornblit, H.J. Guggenheim, *Phys. Rev. Lett.* 34 (1975) 1227.
- [177] R. Abe, *Prog. Theor. Phys.* 49 (1973) 113;
R. Abe, S. Hikami, *Prog. Theor. Phys.* 49 (1973) 442;
S.K. Ma, *Phys. Rev. Lett.* 29 (1972) 1311.
- [178] K.G. Wilson, *Phys. Rev.* D7 (1973) 2911.
- [179] S.K. Ma, *Phys. Rev.* A10 (1974) 1818.
- [180] S.K. Ma, [50] p.250-292 and Ref. [51].
- [181] M. Moshe, J. Zinn-Justin, *Phys. Rep.* 385 (2003) 69.
- [182] P. Di Francesco, P. Ginsparg, J. Zinn-Justin, *Phys. Rep.* 254 (1995) 1;
P. Ginsparg, G. Moore, *Lectures on 2D Gravity and 2D Strings*,
Arxiv: hep-th/9304011 (1993).
- [183] I. Yaref'eva, E.R. Nissimov, S.J. Pacheva, *Comm. Math. Phys.* 71 (1980) 213;
A.N. Vasil'ev, M.Yu. Nalimov, *Teor. Mat. Fiz.* 55 (1983) 163.
- [184] E. Brézin, D.J. Wallace, *Phys. Rev.* B7 (1973) 1967.
- [185] M.E. Fisher, A. Aharony, *Phys. Rev. Lett.* 31 (1973) 1238;
A. Aharony, *Phys. Rev.* B10 (1974) 2834;
R. Abe, S. Hikami, *Prog. Theor. Phys.* 51 (1974) 1041.
- [186] A.N. Vasil'ev, Yu.M. Pis'mak, Yu.R. Honkonen, *Theor. Math. Phys.* 47 (1981) 465;
Theor. Math. Phys. 50 (1982) 127.
- [187] I. Kondor, T. Temesvari, *J. Physique Lett. (Paris)* 39 (1978) L99; Erratum *J. Physique Lett. (Paris)* 39 (1978) L415.
- [188] Y. Okabe, M. Oku, *Prog. Theor. Phys.* 59 (1978) 1825; *Prog. Theor. Phys.* 60 (1978) 1277, 1287; *Prog. Theor. Phys.* 61 (1979) 443.
- [189] A. Pelissetto, E. Vicari, *Nucl. Phys.* B519 (1998) 626.

- [190] S.E. Derkachov, A.N. Manashov *Nucl. Phys.* B522 (1998) 301; *Phys. Rev. Lett.* 79 (1997) 1423.
- [191] S. Weinberg, *Phys. Rev.* 166 (1968) 1568;
J. Schwinger, *Phys. Rev.* 167 (1968) 1432.
- [192] I.S. Gerstein, R. Jackiw, B.W. Lee, S. Weinberg, *Phys. Rev.* D3 (1971) 2486;
J. Honerkamp, K. Meetz, *Phys. Rev.* D3 (1971) 1996;
J. Honerkamp, *Nucl. Phys.* B36 (1972) 130.
- [193] L.D. Faddeev, A.A. Slavnov, *Theor. Math. Phys.* 8 (1971) 843.
- [194] F. David, *Comm. Math. Phys.* 81 (1981) 149.
- [195] A.A. Slavnov, *Nucl. Phys.* B31 (1971) 301.
- [196] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev.* D14 (1976) 2615.
- [197] F.J. Wegner, *Z. Phys.* B78 (1990) 33;
G.E. Castilla, S. Chakravarty, *Phys. Rev. Lett.* 71 (1993) 384.
- [198] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev.* B14 (1976) 4976.
- [199] D. Bessis, J. Zinn-Justin, *Phys. Rev.* D5 (1972) 1313.
- [200] A.M. Polyakov, *Phys. Lett.* 59B (1975) 79.
- [201] E. Brézin, J. Zinn-Justin, *Phys. Rev. Lett.* 36 (1976) 691; *Phys. Rev.* B14 (1976) 3110.
- [202] W.A. Bardeen, B.W. Lee, R.E. Shrock, *Phys. Rev.* D14 (1976) 985.
- [203] S. Hikami, *Nucl. Phys.* B215[FS7] (1983) 555;
F. Wegner, *Nucl. Phys.* B316 (1989) 663.
- [204] J. Zinn-Justin, *Nucl. Phys.* B367 (1991) 105.
- [205] D.J. Gross, A. Neveu, *Phys. Rev.* D10 (1974) 3235.
- [206] L.N. Mihaila, N. Zerf, B. Ihrig, I.F. Herbut, M.M. Scherer, *Phys. Rev.* B96 (2017) 165133.
- [207] J.A. Gracey, *Int. J. Mod. Phys.* A9 (1994) 567, 727; *Nucl. Phys.* B 644 (2002) 433.
- [208] J. A. Gracey, *Phys. Rev.* D96 (2017) 065015.
- [209] W. Wetzel, *Phys. Lett.* 153B (1985) 297;
N.A. Kivel, A.S. Stepanenko, A.N. Vasilev, *Nucl. Phys.* B424 [FS] (1994) 619.
- [210] J.A. Gracey, T. Luthe, Y. Schröder, *Phys. Rev.* D94 (2016) 125028.
- [211] A. Hasenfratz, P. Hasenfratz, K. Jansen, J. Kuti, Y. Shen, *Nucl. Phys.* B365 (1991) 79.
- [212] R. Dashen, B. Hasslacher, A. Neveu, *Phys. Rev.* D12 (1975) 244;
A.B. Zamolodchikov, Al.B. Zamolodchikov, *Phys. Lett.* 72B (1978) 481.
- [213] J. Schwinger ed., *Selected Papers in Quantum Electrodynamics* (Dover, New York 1958);
J. Bjorken and S. Drell, *Relativistic Quantum Mechanics, and Relativistic Quantum Fields* (McGraw-Hill, New York 1964, 1965);
A.I. Akhiezer and V.B. Berestetskii, *Quantum Electrodynamics* (Interscience, New York 1965);
J.D. Jackson, *Classical Electrodynamics*, ISBN 0-471-30932-X (Wiley-VCH 1998).

- [214] M. Peskin, D.V. Schroeder, *An introduction to quantum field theory*, (CRC Press, Taylor and Francis 2018).
- [215] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York 1980).
- [216] T. Aoyama, T. Kinoshita, M. Nio, *Phys. Rev.* D97 (2018) 036001.
- [217] G.C. Rossi, M. Testa, *Nucl. Phys.* B163 (1980) 129.
- [218] R.L. Arnowitt, S.I. Fickler, *Phys. Rev.* 127 (1962) 1821.
- [219] L.D. Faddeev, V.N. Popov, *Phys. Lett.* 25B (1967) 29.
- [220] D.R. Yennie, S.C. Frautschi, H. Suura, *Ann. Phys. (NY)* 13 (1961) 379; T. Kinoshita, *J. Math. Phys.* 3 (1962) 650; T.D. Lee, M. Nauenberg, *Phys. Rev.* 133B (1964) 1549.
- [221] K.G. Wilson, *Phys. Rev.* D10 (1974) 2445.
- [222] J.C. Ward, *Phys. Rev.* 78 (1950) 182; Y. Takahashi, *Nuovo Cimento* 6 (1957) 371.
- [223] V.L. Ginzburg, L.D. Landau, *Zh. Eksp. Teor. Fiz.* 20 (1950) 1064; English translation in: L. D. Landau, *Collected papers* (Oxford: Pergamon Press, 1965) p. 546.
- [224] P.W. Anderson, *Phys. Rev.* 130 (1963) 439; P.W. Higgs, *Phys. Rev. Lett.* 13 (1964) 508; *Phys. Rev.* 145 (1966) 1156; R. Brout, F. Englert, *Phys. Rev. Lett.* 13 (1964) 321; G.S. Guralnik, C.R. Hagen, T.W.B. Kibble, *Phys. Rev. Lett.* 13 (1964) 585.
- [225] B.W. Lee, *Phys. Rev.* D5 (1972) 823.
- [226] A general class of gauges is constructed in G. 't Hooft, *Nucl. Phys.* B35 (1971) 167.
- [227] G. Parisi, Y.S. Wu, *Acad. Sci. Sin.* 24 (1981) 483.
- [228] D. Zwanziger, *Nucl. Phys.* B192 (1981) 259, *Nucl. Phys.* B209 (1982) 336; J. Zinn-Justin, D. Zwanziger, *Nucl. Phys.* B295 [FS21] (1987) 297; K. Okano, *Nucl. Phys.* B289 (1987) 109.
- [229] P.H. Damgaard, H. Hüffel, *Phys. Rep.* 152 (1987) 227.
- [230] H.G.B. Casimir, *Proc. Kon. Ned. Akad. Wetenschap.* B51 (1948) 793; M. Bordag, G.L. Klimchitskaya, U. Mohideen, V.M. Mostepanenko, *Advances in the Casimir effect*, Int. Ser. Monogr. Phys. 145 (Oxford Univ. Press 2009); G.L. Klimchitskaya, U. Mohideen, V.M. Mostepanenko, *Rev. Mod. Phys.* 81 (2009) 1827.
- [231] C.N. Yang, R.L. Mills, *Phys. Rev.* 96 (1954) 191; R. Utiyama, *Phys. Rev.* 101 (1956) 1597; M. Gell-Mann, S. Glashow, *Ann. Phys. (NY)* 15 (1961) 437.
- [232] B.S. DeWitt, *Phys. Rev.* 162 (1967) 1239.
- [233] S. Mandelstam, *Phys. Rev.* 175 (1968) 1580; E.S. Fradkin, I.V. Tyutin, *Phys. Rev.* D2 (1970) 2841.
- [234] K. Fujikawa, B.W. Lee, A.I. Sanda, *Phys. Rev.* D6 (1972) 2923.
- [235] B.W. Lee, J. Zinn-Justin, *Phys. Rev.* D7 (1973) 1049.
- [236] C. Becchi, A. Rouet, R. Stora, *Comm. Math. Phys.* 42 (1975) 127; *Ann. Phys.* 98 (1976) 28.

- [237] I.V. Tyutin, Lebedev preprint FIAN 39 (1975), unpublished.
- [238] G. 't Hooft, *Nucl. Phys.* B33 (1971) 173.
- [239] A.A. Slavnov, *Theor. Mat. Phys.* 13 (1972) 1064.
- [240] G. 't Hooft, M. Veltman, *Nucl. Phys.* B44 (1972) 189.
- [241] F.J. Wegner, *J. Math. Phys.* 12 (1971) 2259.
- [242] P.W. Higgs, *Phys. Rev. Lett.* 13 (1964) 508;
F. Englert, R. Brout, *Phys. Rev. Lett.* 13 (1964) 321.
- [243] B.W. Lee, J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3137, Erratum *Phys. Rev.* D8 (1973) 4654.
- [244] See Ref. [243], Appendix C.
- [245] *Gauge Theory of Weak and Electromagnetic Interactions*, C.H. Lai ed. (World Scientific, Singapore 1981).
- [246] J.C. Taylor, *Gauge Theories of Weak Interactions* (Cambridge Univ. Press 1976);
L.D. Faddeev and A.A. Slavnov, *Gauge Fields: Introduction to Quantum Field Theory* (Benjamin, Reading MA 1980);
L.D. Faddeev, *Gauge fields: and introduction to quantum theory* (CRC Press, Taylor and Francis 1991, 2018).
- [247] V.N. Gribov, *Nucl. Phys.* B139 (1978) 1.
- [248] T.W.B. Kibble, *Phys. Rev.* 155 (1967) 1554.
- [249] B.W. Lee, J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3155.
- [250] S.L. Glashow, *Nucl. Phys.* 22 (1961) 579;
S. Weinberg *Phys. Rev. Lett.* 19 (1967) 1264;
A. Salam, *Elementary Particle Theory*, N. Svartholm ed. (Almqvist and Wiksell, Stockholm 1968).
- [251] N. Cabibbo, *Phys. Rev. Lett.* 10 (1963) 531;
S.L. Glashow, J. Iliopoulos, L. Maiani, *Phys. Rev.* D2 (1970) 1285;
M. Kobayashi, T. Maskawa, *Prog. Theor. Phys.* 49 (1973) 652.
- [252] D.J. Gross, F. Wilczek, *Phys. Rev. Lett.* 30 (1973) 1343;
H.D. Politzer, *Phys. Rev. Lett.* 30 (1973) 1346.
- [253] M. Le Bellac, *Des Phénomènes Critiques aux Champs de Jauge* (Editions du CNRS, 1988), English translation: *Quantum and Statistical Field Theory* (Oxford Univ. Press 1992).
- [254] T. van Ritbergen, J.A.M. Vermaelen, S.A. Larin, *Phys. Lett.* B400 (1997) 379.
- [255] R. Jackiw, C. Rebbi, *Phys. Rev. Lett.* 37 (1976) 172;
C.G. Callan, R.F. Dashen, D.J. Gross, *Phys. Lett.* B63 (1976) 334.
- [256] R.D. Peccei, H.R. Quinn, *Phys. Rev. Lett.* 38 (1977) 1440; *Phys. Rev.* D16 (1977) 1791;
S. Weinberg, *Phys. Rev. Lett.* 40 (1978) 223;
F. Wilczek, *Phys. Rev. Lett.* 40 (1978) 279.
- [257] J.S. Bell, R. Jackiw, *Nuovo Cimento* A60 (1969) 47;
S.L. Adler, *Phys. Rev.* 177 (1969) 2426.

- [258] W.A. Bardeen, *Phys. Rev.* 182 (1969) 1517;
 D.J. Gross, R. Jackiw, *Phys. Rev.* D6 (1972) 477;
 H. Georgi, S.L. Glashow, *Phys. Rev.* D6 (1972) 429;
 C. Bouchiat, J. Iliopoulos, Ph. Meyer, *Phys. Lett.* 38B (1972) 519;
 E. Witten, *Phys. Lett.* 117B (1982) 324.
- [259] G. 't Hooft, *Phys. Rep.* 142 (1986) 357.
- [260] See Ref. [214], and also the volume
 R. Jackiw, S.B. Treiman, E. Witten, B. Zumino, *Current Algebra and Anomalies* (World Scientific, Singapore 1985), and references therein.
- [261] K. Fujikawa, *Phys. Rev.* D21 (1980) 2848; *Phys. Rev.* D22 (1980) 1499(E); *Phys. Rev.* D23 (1981) 2262.
- [262] L. Alvarez-Gaumé, P. Ginsparg, *Nucl. Phys.* B243 (1984) 449, and references therein.
- [263] J. Wess, B. Zumino, *Phys. Lett.* B37 (1971) 95.
- [264] S. Coleman, D.J. Gross, *Phys. Rev. Lett.* 31 (1973) 851;
 S. Coleman, contribution to *Properties of the Fundamental Interactions*, Erice 1973,
 A. Zichichi ed., reprinted in *Aspects of Symmetries* (Cambridge Univ. Press 1985).
- [265] D.J.E. Callaway, *Phys. Rep.* 167 (1988) 241.
- [266] M. Lüscher, P. Weisz, *Nucl. Phys.* B318 (1989) 705.
- [267] H. Georgi, H.D. Politzer, *Phys. Rev.* D9 (1974) 416;
 D.J. Gross, F. Wilczek, *Phys. Rev.* D9 (1974) 920.
- [268] J.B. Kogut, *Rev. Mod. Phys.* 51 (1979) 659;
 J.B. Kogut in *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches school 1982, J.-B. Zuber, R. Stora eds. (North-Holland, Amsterdam 1984).
- [269] M. Lüscher, *Nucl. Phys.* B219 (1983) 233;
 M. Lüscher, G. Münster, *Nucl. Phys.* B232 (1984) 445;
 J. Koller, P. van Baal, *Nucl. Phys.* 302 (1988) 1;
 P. van Baal, *Nucl. Phys.* B307 (1988) 274.
- [270] B. Svetitsky, *Phys. Rep.* 132 (1986) 1;
 E. Marinari, *Phys. Rep.* 184 (1989) 131.
- [271] For Monte-Carlo methods in statistical physics, see, for example,
Monte Carlo Methods in Statistical Physics, K. Binder ed. (Springer-Verlag, New York 1979).
 Early articles on Monte-Carlo simulations in lattice gauge theories are reprinted in *Lattice Gauge Theories and Monte Carlo Simulations*, C. Rebbi ed. (World Scientific, Singapore 1983).
 See also the lectures
 M. Lüscher in *Critical Phenomena, Random Systems, Gauge Theories*, Les Houches school 1984, K. Osterwalder, R. Stora eds. (Elsevier, Amsterdam 1986);
 M. Lüscher in *Fields, Strings and Critical Phenomena*, Les Houches school 1988, p.451-528, E. Brézin, J. Zinn-Justin eds. (Elsevier, Amsterdam 1989).
- [272] S. Elitzur, *Phys. Rev.* D12 (1975) 3978.
- [273] A.M. Polyakov, *Phys. Lett.* 59B (1975) 82.

- [274] M. Lüscher, *Nucl. Phys.* B180 (1981) 317.
- [275] R. Balian, J.M. Drouffe, C. Itzykson, *Phys. Rev.* D11 (1975) 2098;
J.M. Drouffe, J.-B. Zuber, *Phys. Rep.* 102 (1983) 1.
- [276] M. Creutz, *Phys. Rev. Lett.* 45 (1980) 313;
M. Creutz, L. Jacobs, C. Rebbi, *Phys. Rep.* 95 (1983) 201;
I. Montvay, *Rev. Mod. Phys.* 59 (1987) 263.
- [277] Y. Aoki *et al.*, *Nature* 443 (2006) 05120.
- [278] M. Lüscher, *Nucl. Phys.* B549 (1999) 295.
- [279] J. Zinn-Justin, *Chiral Anomalies and Topology*, in *Topology and Geometry in Physics*, Bick E., Steffen F.D. eds. *Lecture Notes in Physics* 659 (2005) p.167-236 (Springer, Berlin)
- [280] S. Dürr, *et al.*, *Science* 322 (2008) 1224.
- [281] M. Lüscher, *Nucl. Phys.* B354 (1991) 531.
- [282] D.J. Gross, E. Witten, *Phys. Rev.* D21 (1980) 446.
- [283] A.A. Slavnov, *Theor. Math. Phys.* 10 (1972) 99;
J.C. Taylor, *Nucl. Phys.* B33 (1971) 436.
- [284] A.A. Slavnov, *Slavnov–Taylor identities*, (2008) Scholarpedia, 3(10): 7119.
- [285] J. Zinn-Justin, *Nucl. Phys.* B275 (1986) 135.
- [286] J. Zinn-Justin, *Mod. Phys. Lett.* A19 (1999) 1227.
- [287] Jean Zinn-Justin, *Zinn-Justin equation*, (2009) Scholarpedia, 4(1): 7120.
- [288] D.A. Ross, J.C. Taylor, *Nucl. Phys.* B51 (1973) 125;
B.W. Lee, *Phys. Lett.* 46B (1973) 214; *Phys. Rev.* D9 (1974) 933;
H. Kluberg-Stern, J.B. Zuber, *Phys. Rev.* D 12 (1975) 482;
B.W. Lee in Ref. [12], p.79-139.
- [289] J. Zinn-Justin, *Nucl. Phys.* B246 (1984) 246.
- [290] H. Kluberg-Stern, J.B. Zuber, *Phys. Rev.* D 12 (1975) 467, 3159.
- [291] J. Zinn-Justin, *Renormalization problems in gauge theories*, school of theoretical physics Karpacz 1975, *Acta Universitatis Wratislaviensis* 368 (1976) 433.
- [292] G. Barnich, M. Henneaux, *Phys. Rev. Lett.* 72 (1994) 1588.
- [293] G. Barnich, F. Brandt, M. Henneaux, *Comm. Math. Phys.* 174 (1995) 57, 93.
- [294] T. Kugo, I. Ojima, *Phys. Lett.* B73 (1978) 459; *Nucl. Phys.* B144 (1978) 234.
- [295] J. Wess, B. Zumino, *Nucl. Phys.* B70 (1974) 39.
- [296] Numerous other textbooks have been devoted to supersymmetry see, for example,
S. Ferrara, *Supersymmetry and Supergravity*, *Lecture Notes in Physics* 208 (1984)
(Springer, Berlin);
B.S. DeWitt, *Supermanifolds*, (Cambridge Univ. Press 1984).
- [297] Many articles are reprinted in
Supersymmetry, S. Ferrara ed. (North-Holland and World Scientific 1987).
- [298] A. Salam, J. Strathdee, *Phys. Lett.* 51B (1974) 353;
S. Ferrara, B. Zumino, J. Wess, *Phys. Lett.* 51B (1974) 239.
- [299] E. Witten, *Phys. Rev.* D16 (1977) 2991.

- [300] Marcus T. Grisaru, A.E.M. van de Ven, D. Zanon, *Nucl. Phys.* B277 (1986) 409; *Phys. Lett.* B173 (1986) 423.
- [301] A.Yu. Morozov, A.M. Perelomov, M.A. Shifman, *Nucl. Phys.* B248 (1984) 279.
- [302] J.A. Gracey, *Nucl. Phys.* B348 (1991) 737; *Nucl. Phys.* B352 (1991) 183; *Phys. Lett.* B262 (1991) 49.
- [303] D.Z. Freedman, P.K. Townsend, *Nucl. Phys.* B177 (1981) 282;
L. Alvarez-Gaumé, D.Z. Freedman, *Comm. Math. Phys.* 80 (1981) 443.
- [304] J. Wess, J. Bagger, *Supersymmetry and Supergravity* (Princeton Univ. Press 1983).
- [305] E. Cartan, *Leçons sur la géométrie des espaces de Riemann* (1928), reprinted by Gauthiers-Villars (Paris 1950);
A. Lichnerowicz, *Géométrie des groupes de transformations* (Dunod, Paris 1958);
E. Nelson, *Tensor Analysis* (Princeton Univ. Press 1967);
S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces* (Academic Press, New York 1978);
Y. Choquet-Bruhat, C. DeWitt-Morette, M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam 1977).
- [306] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York 1972).
- [307] B.S. DeWitt, *Phys. Rev.* 160 (1967) 113; *Phys. Rev.* 162 (1967) 1195.
- [308] B.S. DeWitt, S.W. Hawking, in *Relativity, Groups and Topology*, Les Houches school 1983, B.S. DeWitt, R. Stora eds. (Elsevier Science, Amsterdam 1984).
- [309] Selected articles are gathered in
Euclidean Quantum Gravity, G.W.Gibbons, S.W. Hawking eds. (World Scientific, Singapore 1993).
- [310] J. Ambjorn, J. Jurkiewicz, R. Loll, *Phys. Rev. Lett.* 93 (2004) 131301; *Phys. Rev. Lett.* 95 (2005) 171301.
- [311] S. Bernuzzi, A. Nagar, T. Dietrich, T. Damour, *Phys. Rev. Lett.* 114 (2015) 161103.
- [312] S. Coleman, J. Wess, B. Zumino, *Phys. Rev.* 177 (1969) 2239;
S. Gasiorowicz, D. Geffin, *Rev. Mod. Phys.* 41 (1969) 531.
- [313] V.E. Zakharov, A.V. Mikhailov, *Zh. Eksp. Teor. Fiz.* 74 (1978) 1953; *Comm. Math. Phys.* 74 (1980) 21;
H. Eichenherr, *Nucl. Phys.* B146 (1978) 215;
H. Eichenherr, M. Forger, *Nucl. Phys.* B155 (1979) 381;
E. Brézin, C. Itzykson, J. Zinn-Justin, J.-B. Zuber, *Phys. Lett.* 82B (1979) 442;
H.J. de Vega, *Phys. Lett.* 87B (1979) 233.
- [314] M. Lüscher, *Nucl. Phys.* B135 (1978) 1;
D. Bernard, A. LeClair, *Comm. Math. Phys.* 142 (1991) 99, and references therein.
- [315] R.D. Pisarski, *Phys. Rev.* D20 (1979) 3358;
E. Brézin, S. Hikami, J. Zinn-Justin, *Nucl. Phys.* B165 (1980) 528.
- [316] S. Hikami, *Phys. Lett.* 98B (1981) 208;
W. Bernreuther, F.J. Wegner, *Phys. Rev. Lett.* 57 (1986) 1383;
F. Wegner, *Nucl. Phys.* B316 (1989) 663.

- [317] For the application of these models to the metal–insulator transition of independent electrons in a random potential, see
 F. Wegner, *Z. Phys.* B35 (1979) 207;
 L. Schäfer, F. Wegner, *Z. Phys.* B38 (1980) 113;
 S. Hikami, *Prog. Theor. Phys. Suppl.* 84 (1985) 120.
- [318] D.H. Friedan, *Ann. Phys. (NY)* 163 (1985) 318.
- [319] J. Schwinger, *Phys. Rev.* 128 (1962) 2425.
- [320] L.D. Faddeev, L.A. Takhtajan, *Teor. Mat. Fiz.* 21 (1974) 160.
- [321] R.F. Dashen, B. Hasslacher, A. Neveu, *Phys. Rev.* D11 (1975) 3424; *Phys. Rev.* D12 (1975) 2443.
- [322] W. Thirring, *Ann. Phys. (N.Y.)* 3 (1958) 91.
- [323] A. Luther, I. Peschel, *Phys. Rev.* B9 (1974) 2911;
 S. Coleman, *Phys. Rev.* D11 (1975) 2088.
- [324] Ph. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal Field Theory*, (Springer, New-York 1997).
- [325] J.H. Lowenstein, J.A. Swieca, *Ann. Phys. (NY)* 68 (1971) 172.
- [326] S. Coleman, R. Jackiw, L. Susskind, *Ann. Phys. (NY)* 93 (1975) 267;
 S. Coleman, *Ann. Phys. (NY)* 101 (1976) 239.
- [327] T. Banks, D. Horn, H. Neuberger, *Nucl. Phys.* B108 (1976) 119.
- [328] V.L. Berezinskii, *Zh. Eksp. Teor. Fiz.* 59 (1970) 907 (*Sov. Phys. JETP* 32 (1970) 493).
- [329] J.M. Kosterlitz, D.J. Thouless, *J. Phys. C: Solid State Phys.* 6 (1973) L97;
 J.M. Kosterlitz, *J. Phys. C: Solid State Phys.* 7 (1974) 1046.
- [330] P. Minnhagen, *Rev. Mod. Phys.* 59 (1987) 1001;
 B. Nienhuis, *J. Stat. Phys.* 34 (1984) 731.
- [331] J. Villain, *J. Phys. (France)* 6 (1975) 581;
 J.V. José, L.P. Kadanoff, S. Kirkpatrick, D.R. Nelson, *Phys. Rev.* B16 (1977) 1217.
- [332] P.B. Wiegmann, *J. Phys. C: Solid State Phys.* 11 (1978) 1583;
 D.J. Amit, Y.Y. Goldschmidt, G. Grinstein, *J. Phys. A: Math. Gen.* 13 (1980) 585.
- [333] M.E. Fisher in *Critical Phenomena*, School of Physics Enrico Fermi, Course No 51, Varenna 1970, M.S. Green ed. (Academic Press, New York 1971);
 M.E. Fisher, M.N. Barber, *Phys. Rev. Lett.* 28 (1972) 1516;
 Y. Imry, D. Bergman, *Phys. Rev. A3* (1971) 1416.
- [334] M.N. Barber in *Phase Transitions and Critical Phenomena* vol. 8, C. Domb, J. Lebowitz eds. (Academic Press, New York 1983);
Finite Size Scaling, J.L. Cardy ed. (North-Holland, Amsterdam 1988);
Finite Size Scaling and Numerical Simulation of Statistical Systems, V. Privman ed., World Scientific Singapore (1990).
- [335] V. Privman, M.E. Fisher, *J. Stat. Phys.* 33 (1983) 385.
- [336] E. Brézin, *J. Physique (Paris)* 43 (1982) 15;
 M. Lüscher, *Phys. Lett.* 118B (1982) 391; *Nucl. Phys.* B219 (1983) 233;

- E. Brézin, J. Zinn-Justin, *Nucl. Phys.* B257 [FS14] (1985) 867;
 J. Rudnick, H. Guo, D. Jasnow, *J. Stat. Phys.* 41 (1985) 353.
- [337] E.G. Floratos, D. Petscher, *Nucl. Phys.* B252 (1985) 689.
- [338] D.J. Wallace, R.K.P. Zia, *Phys. Rev. Lett.* 43 (1979) 808.
- [339] Among the various reviews and textbooks see, for example,
 N.P. Landsman, C. van Weert, *Phys. Rep.* 145 (1987) 141;
 J. I. Kapusta, *Finite Temperature Field Theory* (Cambridge Univ. Press 1989);
 M. Le Bellac, *Thermal Field Theory* (Cambridge Univ. Press 1996, 2000);
 H. Meyer-Ortmanns, *Rev. Mod. Phys.* 68 (1996) 473;
 J.P. Blaizot, E. Iancu, *Phys. Rep.* 359 (2002) 355.
- [340] D.A. Kirzniits *JETP Lett.* 15 (1972) 529;
 D.A. Kirzniits, A.D. Linde, *Phys. Lett.* B42 (1972) 471; *Ann. Phys.* 101 (1976) 195;
 A.D. Linde *Rep. Prog. Phys.* 42 (1979) 389.
- [341] C.W. Bernard, *Phys. Rev.* D9 (1974) 3312;
 S. Weinberg, *Phys. Rev.* D9 (1974) 3357;
 L. Dolan, R. Jackiw, *Phys. Rev.* D9 (1974) 3320;
 D.J. Gross, R.D. Pisarski, L.G. Yaffe, *Rev. Mod. Phys.* 53 (1981) 43.
- [342] K.-C. Chou, Z.-B. Su, B.-L. Hao, L. Yu, *Phys. Rep.* 118 (1985) 1.
- [343] P. Ginsparg, *Nucl. Phys.* B170 (1980) 388;
 T. Appelquist, R.D. Pisarski, *Phys. Rev.* D23 (1981) 2305;
 S. Nadkarni, *Phys. Rev.* D27 (1983) 917, *ibidem* D38 (1988) 3287;
 E. Braaten, A. Nieto, *Phys. Rev.* D51 (1995) 6990;
 K. Kajantie, K. Rummukainen, M.E. Shaposhnikov, *Nucl. Phys.* B407 (1993) 356.
- [344] S. Sachdev, *Quantum Phase Transitions*, (Wiley 2007).
- [345] N.P. Landsman, *Nucl. Phys.* B322 (1989) 498.
- [346] E. Braaten, A. Nieto, *Phys. Rev.* D51 (1995) 6990;
 K. Kajantie, M. Laine, K. Rummukainen, M. Shaposhnikov, *Nucl. Phys.* B458 (1996) 90.
- [347] J.I. Kapusta, C. Gale, *Finite-temperature field theory: Principles and applications* (Cambridge monographs on mathematical physics, 1996, 2006).
- [348] K. Farakos, K. Kajantie, K. Rummukainen, M. Shaposhnikov, *Nucl. Phys.* B425 (1994) 67; *Nucl. Phys.* B442 (1995) 317.
- [349] P. Arnold, C. Zhai, *Phys. Rev.* D51 (1995) 1906.
- [350] P. Arnold, O. Espinosa, *Phys. Rev.* D47 (1993) 3546;
 Z. Fodor, A. Hebecker, *Nucl. Phys.* B432 (1994) 127.
- [351] J.-P. Blaizot, E. Iancu, A. Rebhan, *Phys. Rev.* D63 (2001) 065003;
 J.O. Andersen, E. Braaten, M. Strickland, *Phys. Rev.* D63 (2001) 105008.
- [352] P. Langevin, *Comptes Rendus Acad. Sci. Paris* 146 (1908) 530.
- [353] A.D. Fokker, *Ann. der Physik* 348 (1914) 810;
 M. Planck, *Sitzungberichte der preuss. Akademie der Wissenschaften* 24 (1917) 324.
 For an early reference, see S. Chandrasekhar, *Rev. Mod. Phys.* 15 (1943) 1.
- [354] P.C. Martin, E.D. Siggia, H.A. Rose, *Phys. Rev.* A8 (1973) 423.

- [355] H. Nicolai, *Phys. Lett.* 89B (1980) 341; *Nucl. Phys.* B176 (1980) 419.
- [356] E. Witten, *Nucl. Phys.* B188 (1981) 513;
M.V. Feigel'man, A.M. Tsvelik, *Sov. Phys.-JETP* 56 (1982) 823;
H. Nakazato, M. Namiki, I. Ohba, K. Okano, *Prog. Theor. Phys.* 70 (1983) 298;
E. Egorian, S. Kalitzin, *Phys. Lett.* 129B (1983) 320.
- [357] C. Jarzynski, *Phys. Rev. Lett.* 78 (1997) 2690.
- [358] G. Parisi, Chapters 19 and 20 of Ref. [134].
- [359] N.G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam 1981, 1992);
A selection of early papers can be found in
Selected Papers on Noise and Stochastic Processes, N. Wax ed. (Dover, New York 1954).
- [360] C. De Dominicis, L. Peliti, *Phys. Rev.* B18 (1978) 353, and references therein.
- [361] J.L. Cardy, *Phys. Lett.* B125 (1983) 470.
- [362] R. Bausch, H.K. Janssen, Y. Yamazaki, *Z. Phys.* B37 (1980) 163.
- [363] K.D. Elworthy, *Stochastic Differential Equations on Manifolds*, London Math. Soc. *Lectures Notes* 70 (Cambridge Univ. Press 1982).
- [364] G. Parisi, N. Sourlas, *Phys. Rev. Lett.* 43 (1979) 744.
- [365] Y. Imry, S.K. Ma, *Phys. Rev. Lett.* 35 (1975).
- [366] J. Imbrie, *Comm. Math. Phys.* 98 (1985) 145;
J. Bricmont, A. Kupiainen, *Phys. Rev. Lett.* 59 (1987) 1829.
- [367] P.C. Hohenberg, B.I. Halperin, *Rev. Mod. Phys.* 49 (1977) 435.
- [368] C. De Dominicis, *Nuovo Cimento Lett.* 12 (1975) 567;
C. De Dominicis, E. Brézin, J. Zinn-Justin, *Phys. Rev.* B12 (1975) 4945;
E. Brézin, C. De Dominicis, *Phys. Rev.* B12 (1975) 4954;
H.K. Janssen, *Z. Phys.* B23 (1976) 377.
- [369] H. K. Janssen, B. Schaub, B. Schmittmann, *Z. Phys.* B73 (1989) 539.
- [370] N.V. Antonov, A.N. Vasil'ev, *Theor. Math. Phys.* 60 (1984) 671.
- [371] R. Bausch, V. Dohm, H.K. Janssen, R.K.P. Zia, *Phys. Rev. Lett.* 47 (1981) 1837;
J.C. Niel, *Phys. Rev.* B37 (1988) 2156;
N. Breuer, H.K. Janssen, *Z. Phys.* B41 (1981) 55;
H.K. Janssen, B. Schmittmann, *Z. Phys.* B64 (1986) 503;
K. Leung, J.L. Cardy, *J. Stat. Phys.* 44 (1986) 567.
- [372] M. Suzuki, *Prog. Theor. Phys.* 58 (1977) 1142;
J.C. Niel, J. Zinn-Justin, *Nucl. Phys.* B280 [FS18] (1987) 335;
Y.Y. Goldschmidt, *Nucl. Phys.* B280[FS18] (1987) 340.
- [373] In Chapters 37–40, we follow J. Zinn-Justin, in *Recent Advances in Field Theory and Statistical Mechanics* p.41-172, Les Houches school 1982, R. Stora, J.-B. Zuber eds. (Elsevier, Amsterdam 1984); *Phys. Rep.* 70 (1981) 109.
- [374] A pedagogical presentation of instanton calculus can also be found in
S. Coleman, *The Whys of Subnuclear Physics*, Erice 1977, A. Zichichi ed. (Plenum, New York 1979).

- [375] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev.* D15 (1977) 1544.
- [376] J.J. Loeffel, A. Martin, B. Simon, A.S. Wightman, *Phys. Lett.* 30B (1969) 656; S. Graffi, V. Grecchi, B. Simon, *Phys. Lett.* 32B (1970) 631; B. Simon, *Functional Integration and Quantum Physics* (Academic Press, New York 1979).
- [377] C.M. Bender, T.T. Wu, *Phys. Rev.* 184 (1969) 1231.
- [378] J. Zittartz, J.S. Langer, *Phys. Rev.* 148 (1966) 741; J.S. Langer, *Ann. Phys. (NY)* 41 (1967) 108.
- [379] R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York 1966, 2nd edn. Springer, New York 1982) p.433.
- [380] E. Brézin, J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev.* D15 (1977) 1558.
- [381] I.M. Gel'fand, A.M. Yaglom, *J. Math. Phys.* 1 (1960) 48; R.F. Dashen, B. Hasslacher, A. Neveu, *Phys. Rev.* D10 (1974) 4114, 4130 and 4138.
- [382] J.H. Van Vleck, *PNAS* 14 (1928) 178.
- [383] U.D. Jentschura, J. Zinn-Justin, *Ann. Phys.* 326 (2011) 2186; L.T. Giorgini *et al*, *Phys. Rev.* D101 (2020) 125001.
- [384] L.N. Lipatov, *JETP Lett.* 24 (1976) 157; *Sov. Phys. JETP* 44 (1976) 1055; *JETP Lett.* 25 (1977) 104; *Sov. Phys. JETP* 45 (1977) 216.
- [385] S. Coleman, V. Glaser, A. Martin, *Comm. Math. Phys.* 58 (1978) 211.
- [386] G.H. Derrick, *J. Math. Phys.* 5 (1964) 1252.
- [387] J. Zinn-Justin (1977) unpublished and Ref. [375].
- [388] J.-L. Gervais, B. Sakita, *Phys. Rev.* D11 (1975) 2943; V.E. Korepin, P.P. Kulish, L.D. Faddeev, *JETP Lett.* 21 (1975) 138; N.H. Christ, T.D. Lee, *Phys. Rev.* D12 (1975) 1606.
- [389] Y. Frishman, S. Yankielowicz, *Phys. Rev.* D19 (1979) 540.
- [390] Early references on vacuum decay are, for example,
Y.B. Zeldovich, I.Yu. Kobzarev, L.B. Okun, *Sov. Phys. JETP* 40 (1974) 1;
M. Stone, *Phys. Rev.* D14 (1976) 3568; *Phys. Lett.* 67B (1977) 186.
- [391] S. Coleman, *Phys. Rev.* D15 (1977) 2929; Erratum *Phys. Rev.* D16 (1977) 1248; C.G. Callan, S. Coleman, *Phys. Rev.* D16 (1977) 1762; S. Coleman, F. De Luccia, *Phys. Rev.* D21 (1980) 3305.
- [392] G. Parisi, *Phys. Lett.* 66B (1977) 167; *Phys. Lett.* 68B (1977) 361.
- [393] A.M. Polyakov, *Nucl. Phys.* B120 (1977) 429.
- [394] J. Zinn-Justin, *Introduction to spontaneous symmetry breaking for pedestrians, Basko Polje summer school 1976*, Saclay preprint T/76/99, unpublished.
- [395] J. Zinn-Justin, U.D. Jentschura, *Ann. Phys. (NY)* 313 (2004) 197, 269.
- [396] E.B. Bogomolnyi, *Sov. J. Nucl. Phys.* 24 (1976) 449.
- [397] H. Eichenherr, *Nucl. Phys.* B146 (1978) 215; Erratum *Nucl. Phys.* B155 (1979) 544; A. D'Adda, P. Di Vecchia, M. Lüscher, *Nucl. Phys.* B146 (1978) 63; *Nucl. Phys.* B152 (1979) 125; V.L. Golo, A. Perelomov, *Phys. Lett.* 79B (1978) 112.

- [398] V.A. Fateev, I.V. Frolov, A.S. Schwarz, *Nucl. Phys.* B154 (1979) 1;
 B. Berg, M. Lüscher, *Comm. Math. Phys.* 69 (1979) 57;
 A.P. Bukhvostov, L.N. Lipatov, *Nucl. Phys.* B180 [FS2] (1981) 116.
- [399] A.A. Belavin, A.M. Polyakov, A.S. Schwartz, Yu S. Tyupkin, *Phys. Lett.* 59B (1975) 85;
 E. Witten, *Phys. Rev. Lett.* 38 (1977) 121;
 G. 't Hooft, *Phys. Rev.* D14 (1976) 3432; Erratum *Phys. Rev.* D18 (1978) 2199;
 A. Actor, *Rev. Mod. Phys.* 51 (1979) 461.
- [400] F.J. Dyson, *Phys. Rev.* 85 (1952) 631.
- [401] C.M. Bender, T.T. Wu, *Phys. Rev.* 184 (1969) 1231; *Phys. Rev.* D7 (1973) 1620;
Phys. Rev. Lett. 37 (1976) 117;
 T. Banks, C.M. Bender, T.T. Wu, *Phys. Rev.* D8 (1973) 3346;
 T. Banks, C.M. Bender, *Phys. Rev.* D8 (1973) 3366.
- [402] C.S. Lam, *Nuovo Cimento*, 55A (1968) 258.
- [403] See the review
 E.B. Bogomolny, V.A. Fateyev, L.N. Lipatov, *Sov. Sci. Rev.* section A: *Phys. Rev.* 2 (1980) 247, I.M. Khalatnikov ed.
 and the extensive reprint volume
Large Order Behaviour of Perturbation Theory, Current Physics vol. 7, J.C. Le Guillou, J. Zinn-Justin eds. (North-Holland, Amsterdam 1990).
- [404] G. Parisi, *Phys. Lett.* 66B (1977) 167.
- [405] E.B. Bogolomny, *Phys. Lett.* 67B (1977) 193.
- [406] E. Brézin, G. Parisi, J. Zinn-Justin, *Phys. Rev.* D16 (1977) 408.
- [407] E.B. Bogomolny, *Phys. Lett.* 91B (1980) 431.
- [408] J. Zinn-Justin, *Nucl. Phys.* B192 (1981) 125; *Nucl. Phys.* B218 (1983) 333, *J. Math. Phys.* 22 (1981) 511; *J. Math. Phys.* 25 (1984) 549.
- [409] E. Brézin, G. Parisi, *J. Stat. Phys.* 19 (1978) 269.
- [410] M. Beneke, *Phys. Rep.*, 317 (1999) 1.
- [411] G. Parisi in *Cargèse Lectures 1977*, vol. B39 (Plenum, New York 1979); *Phys. Lett.* 76B (1978) 65;
 B. Lautrup, *Phys. Lett.* 69B (1977) 109;
 G. 't Hooft, *Erice 1977*, A. Zichichi ed. (Plenum, New York 1979).
- [412] G. Parisi, *Nucl. Phys.* B150 (1979) 163;
 F. David, *Nucl. Phys.* B209 (1982) 433; *Nucl. Phys.* B234 (1984) 237; *Nucl. Phys.* B263 (1986) 637.
- [413] G. Parisi, *Phys. Lett.* 66B (1977) 382.
- [414] P. Renouard, *Annales de l'I.H.P.* A27 (1977) 237.
- [415] E.B. Bogomolny, V.A. Fateyev, *Phys. Lett.* 71B (1977) 93;
 L.N. Lipatov, A.P. Bukhvostov, E.I. Malkov, *Phys. Rev.* D19 (1979) 2974.
- [416] C. Itzykson, G. Parisi, J.-B. Zuber, *Phys. Rev.* D16 (1977) 996;
 R. Balian, C. Itzykson, J.-B. Zuber, G. Parisi, *Phys. Rev.* D17 (1978) 1041.

- [417] E.B. Bogomolny, V.A. Fateyev, *Phys. Lett.* 76B (1978) 210;
 E.B. Bogomolny, Yu A. Kubyshin, *Yad. Fiz.* 34 (1981) 1535, *Yad. Fiz.* 35 (1982) 202 (*Sov. J. Nucl. Phys.* 34 (1981) 853; *Sov. J. Nucl. Phys.* 35 (1982) 114).
- [418] The calculation of the series expansion for the RG functions has been initiated by Nickel, and reported (together with an estimate of $N = 1$ exponents) in
 G.A. Baker, B.G. Nickel, M.S. Green, D.I. Meiron, *Phys. Rev. Lett.* 36 (1976) 1351;
 B.G. Nickel, D.I. Meiron, G.B. Baker, *Univ. of Guelph Report* 1977;
 G.A. Baker, B.G. Nickel, D.I. Meiron, *Phys. Rev. B* 17 (1978) 1365.
- [419] D.B. Murray, B.G. Nickel, unpublished Guelph University report (1991).
- [420] S.A. Antonenko, A.I. Sokolov, *Phys. Rev. E* 51 (1995) 1894.
- [421] E. Borel, *Leçon sur les Séries Divergentes*, 2nd edn (Gauthier-Villars, Paris 1928);
 G. Hardy, *Divergent Series* (Oxford Univ. Press 1949).
- [422] J.P. Eckmann, J. Magnen, R. Sénéor, *Comm. Math. Phys.* 39 (1975) 251;
 J.S. Feldman, K. Osterwalder, *Ann. Phys. (NY)* 97 (1976) 80;
 J. Magnen, R. Sénéor, *Comm. Math. Phys.* 56 (1977) 237;
 J.-P. Eckmann, H. Epstein, *Comm. Math. Phys.* 68 (1979) 245.
- [423] Summation of series by Borel transformation and mapping was proposed by
 J.J. Loeffel, *Saclay Report*, DPh-T/76/20 unpublished.
- [424] J.C. Le Guillou, J. Zinn-Justin, *J. Physique Lett., EDP Sciences* 46 (1985) L137; *J. Physique, EDP Sciences* 48 (1987) 19; *J. Physique, EDP Sciences* 50 (1989) 1365.
- [425] J.C. Le Guillou, J. Zinn-Justin, *Phys. Rev. Lett.* 39 (1977) 95; *Phys. Rev. B* 21 (1980) 3976.
- [426] B. Nienhuis, *Phys. Rev. Lett.* 49 (1982) 1062.
- [427] A.J. Guttmann, in *Phase Transitions and Critical Phenomena* vol. 13, C. Domb, J. Lebowitz eds. (Academic Press, New York 1989).
- [428] G.A. Baker, *Phys. Rev. B* 15 (1977) 1552.
- [429] S. Zinn, S. Lai, M. Fisher, *Phys. Rev. E* 54 (1996) 1176.
- [430] R. Guida, J. Zinn-Justin, *J. Phys. A* 31 (1998) 8103.
- [431] J. Zinn-Justin, *J. Physique, EDP Sciences* 42 (1981) 783;
 A.J. Guttmann, *J. Phys. A* 20 (1987) 1855;
 A.J. Liu, M.E. Fisher, *Physica* A156 (1989) 35;
 A.J. Guttmann, I.G. Enting, *J. Phys. A* 26 (1993) 807, *J. Phys. A* 27 (1994) 8007;
 G. Bhanot, M. Creutz, U. Glässner, K. Schilling, *Phys. Rev. B* 49 (1994) 12909;
 P. Butera, M. Comi, *Phys. Rev. E* 55 (1997) 6391; *Phys. Rev. B* 56 (1997) 8212.
- [432] G.S. Pawley, R.H. Swendsen, D.J. Wallace, K.G. Wilson, *Phys. Rev. B* 29 (1984) 4030;
 C.F. Baillie, R. Gupta, K.A. Hawick, G.S. Pawley, *Phys. Rev. B* 45 (1992) 10438;
 B. Li, N. Madras, A.D. Sokal, *J. Stat. Phys.* 80 (1995) 661;
 S. Caracciolo, M.S. Causo, A. Pelissetto, *Phys. Rev. E* 57 (1998) 1215;
 M. Campostrini *et al.*, *Phys. Rev. B* 74 (2006) 144506.
- [433] N. Clisby, B. Dünweg, *Phys. Rev. E* 94 (2016) 052102;
 N. Clisby, *J. Phys. A* 50 (2017) 264003.

- [434] H. Shimada, S. Hikami, *J. Stat. Phys.* 165 (2016) 1006.
- [435] F. Kos, D. Poland, D. Simmons-Duffin, A. Vichi, *JHEP* 2016 (2016) 36.
- [436] *Phase Transitions*, Cargèse Summer Institute 1980, vol. B72, M. Lévy, J.C. Le Guillou, J. Zinn-Justin eds. (Plenum, New York 1982).
- [437] J.P. Cotton, *J. Physique Lett. (Paris)* 41 (1980) L231.
- [438] D. Beysens, A. Bourgou, P. Calmettes, *Phys. Rev. A* 26 (1982) 3589;
A. Kumar, H.R. Krishnamurthy, E.S.R. Gopal, *Phys. Rep.* 98 (1983) 57.
- [439] D. Beysens, A. Bourgou, *Phys. Rev. A* 19 (1979) 2407.
- [440] J.V. Sengers, J.G. Shanks, *J. Stat. Phys.* 137 (2009) 857.
- [441] P. Damay, F. Leclercq, P. Chieux, *Phys. Rev. B* 40 (1989) 4696.
- [442] A. Lytle, D.T. Jacobs, *J. Chem. Phys.* 120 (2004) 5709.
- [443] V. Privman, P.C. Hohenberg and A. Aharony, in *Phase Transitions and Critical Phenomena* vol. 14, C. Domb, J. Lebowitz eds. (Academic Press, New York 1991).
- [444] R. Guida, J. Zinn-Justin, *Nucl. Phys.* B489 (1997) 626.
- [445] C. Bagnuls, C. Bervillier, *Phys. Rev. B* 24 (1981) 1226; *J. Phys. A* 19 (1986) L85;
C. Bervillier, *Phys. Rev. B* 34 (1986) 8141;
C. Bagnuls, C. Bervillier, D.I. Meiron, B.G. Nickel, *Phys. Rev. B* 35 (1987) 3585;
R. Schloms, V. Dohm, *Phys. Rev. B* 42 (1990) 6142;
H. J. Krause, R. Schloms, V. Dohm, *Z. Phys.* B79 (1990) 287;
F.J. Halfkann, V. Dohm, *Z. Phys.* B89 (1992) 79;
G. Münster, J. Heitger, *Nucl. Phys.* B424 [FS] (1994) 582;
S.A. Larin, M. Moenigmann, M. Stroesser, V. Dohm, *Phys. Rev. B* 58 (1998) 3394;
M. Stroesser, S.A. Larin, V. Dohm, *Nucl. Phys.* B540 (1999) 654.
- [446] J. Nutall, *J. Math. Anal. Appl.* 31 (1970) 147;
J. Zinn-Justin, *Rocky Mountain J. Math.* 4 (1974) 325.
- [447] R. Seznec, J. Zinn-Justin, *J. Math. Phys.* 20 (1979) 1398.
- [448] R. Guida, K. Konishi, H. Suzuki, *Ann. Phys. (NY)* 241 (1995) 152.
- [449] U.D. Jentschura, J. Zinn-Justin, *J. Phys. A: Math. Theor.* 43 (2010) 425301.
- [450] J.C. Le Guillou, J. Zinn-Justin, *Ann. Phys. (NY)* 147 (1983) 57.
- [451] A.J. Guttmann, G.S. Joyce, *J. Phys. A: Math. Gen.* 5 (1972) L81;
M.E. Fisher, J.H. Chen, in Ref. [436].
- [452] G.V. Dunne, M. Ünsal, *JHEP* 2012 (2012) 170.
- [453] A.A. Andrianov, *Ann. Phys. (NY)* 140 (1982) 82.
- [454] R. Damburg, R. Propin, V. Martyshchenko, *J. Phys. A* 17 (1984) 3493.
- [455] G.V. Dunne, M. Ünsal, *Phys. Rev. D* 89 (2014) 105009.
- [456] A. Voros, *Ann. Inst. Henri Poincaré* A39 (1983) 211.
- [457] F. Pham, *Resurgence, Quantized Canonical Transformation and Multi-Instanton*, Algebraic Analysis, vol. II (1988), *C.R. Acad. Sci. Paris* 309 (1989) 999;
E. Delabaere, H. Dillinger, F. Pham, *J. Math. Phys.* 38 (1997) 6126.

Index

- Abelian anomaly 582
Abelian gauge theory: finite temperature 817
action, effective 166
adjoint representation 294
almost planar interface model 785
amplitude ratios 416
amputated correlation functions 142
anharmonic oscillator 960, 1007
annihilation operator 64
anomalies 567, 573
anomalous dimension 372
anomaly calculation 584
anomaly cancellation 586, 591
anticommutation relations 9
anti-periodic boundary conditions 84, 86, 787
area law 614
Arrhenius law 949
asymptotic freedom 459, 488 489, 567, 582, 593, 611, 708, 735
asymptotic series 975
axial current 582, 586, 589
partial conservation (PCAC) 304
axial gauge symmetry 586
axion 577
background field 114, 122, 262, 709
band structure 945
bare action 189
bare correlation functions 188
bare parameters 188
bare RG equations 211
Bargmann potential 907
barrier penetration 899, 900, 908, 924, 942, 951
Berezinian 12
Bianchi identity 680
binary fluids 987
Bohr–Sommerfeld quantization 37
Borel
 summability 452, 976, 977, 980, 997
 summation 976, 978, 983, 999
 transformation 976
Borel–Leroy transformation 983
Bose gas 72, 102
Bose–Einstein condensation 77, 390
bosonization 721
boundary conditions: twisted 765
Brillouin zone 176
Brownian chain 377
Brownian motion 23, 831, 850
BRST
 operator 634
 symmetry 556, 638, 701, 843, 857, 863
Cabibbo angle 574
Cabibbo–Kobayashi–Maskawa (CKM)
 matrix 568
Callan–Symanzik (CS) equations 197
canonical
 dimension 162
 invariance 44, 646
 transformations 43, 58
Casimir effect 544
Casimir operator 294
characteristics: method of 212, 370
charge conjugation 261, 267, 288
chemical potential 73, 269
chiral
 anomaly 276, 585, 729, 733
 components 287
 gauge invariance 746
 models 703, 868
 symmetry 268, 304, 489, 738, 740
 discrete 490
 spontaneous breaking 733, 734
Christoffel symbol 674
classical action: Hamiltonian formalism 42
classical field equation 552
classical gas: field integrals 382
Clifford algebra 9, 280
cluster properties 140, 330
cohomology operator 649
Coleman–Gross theorem 593
collective coordinate 905, 909, 921

- colour
 - confinement 567
 - gauge group 567
 - quantum number 573, 576, 602, 607
- commutation relations 106
- compatibility conditions 294, 630
- composite operators 173
- confinement 576, 607, 611, 733, 734
- conformal
 - bootstrap 420
 - group 321
 - invariance 322, 721, 940, 985
 - transformations 321
- connected contributions 4
- connected correlation functions 131
 - decay of 157
- connection 549, 674
- conserved current 510
 - non-local 705
- continuum limit 28, 50
- convexity 144
- correlation functions 126, 176, 836
- correlation length 158, 324, 339
- correlation time 880, 888
- coset space 693
- cosine potential 945, 993, 1000
- cosmological constant 670
- cosmological term 684
- Coulomb gas 747, 751
- Coulomb potential 99
- Coulomb's gauge 517
- counter-terms 189
- coupling constant renormalization 189
- covariance of RG functions 209
- covariant
 - curl 673
 - derivative 514, 521, 549, 675, 704
 - gauge 554
 - measure 49
- CP violation: strong 958
- CP(N-1) model 953
 - semi-classical vacuum 955
- creation operator 64
- critical
 - dimension: lower- 459, 602
 - upper- 423
- domain 368, 391
- exponent: see exponent 338
- slowing down 875
- surface 360
 - temperature 337, 797
- crossing symmetry 115
- crossover exponent 408
- current 317
- curvature tensor 522, 550, 609, 679
- cut-off 178
- cyclic identity 680
- cylindrical geometry 772
- dark energy 690
- dark matter 577, 690
- decay of the false vacuum 934
- decay rate 900
- deep inelastic scattering 604
- degenerate minima 942, 963, 992
 - general potential 1004
- density matrix at thermal equilibrium 18, 64
- detailed balance 839, 854
- differential approximants: linear 991
- differential forms 672
 - closed, exact 7
- dilatation current 319
- dilute gas approximation 996
- dimension of
 - vertex 163
 - composite operators 174
 - fields 162
- dimensional reduction 36, 786, 789, 804
- dimensional regularization 223, 465
 - fermions 276
- dipolar forces, strong 431
- Dirac γ matrices 280
- Dirac fermion 103
- Dirac neutrinos 567
- Dirac operator: index 588
- dissipative couplings 875
- divergence from operator ordering 50
- divergences, ultraviolet (UV) 169
- double-well potential 993
- dressed skeleton diagram 194
- driving force 838, 875

- dynamic
 - action 842, 857, 858
 - correlation functions 843, 858
 - critical exponent 879
 - scaling 879, 888
- effective action at one-loop 806
- effective field theory (EFT) 213
- eikonal approximation 97
- Einstein's action 684
- electromagnetic π_0 decay 592
- electron–positron annihilation 604
- Elitzur's conjecture 464
- Elitzur's theorem 617
- emergent symmetry 334, 408
- energy–momentum tensor 318
- epsilon expansion 368, 404
- equation of state 338, 441, 484
 - scaling 401
- equilibrium distribution 835, 840, 853
- Euclidean
 - action 22, 118
 - magnetic field 47
 - fermion 280
 - formulation 18
 - time 20
- Euler–Lagrange equations 43
- evolution operator 19, 106
- exceptional momenta 204, 248
- exponent
 - α 399
 - β 401, 439, 485
 - δ 338, 401
 - Δ_1 419
 - η 339, 372, 456, 488
 - γ 338, 398, 456
 - ν 340, 374, 484, 488
 - ω 370, 456
 - θ 419, 456
 - classical value 340
- Faddeev–Popov ghosts 554
- Faddeev–Popov quantization 541
- Fermi constant 571
- Fermi gas: non-relativistic 88
- fermion 83
 - large-order behaviour 968
- loops 272
- mass generation 504, 743
- number violation 268
- propagator 270
- relativistic 258
- vertex functions 270
- Feynman
 - diagrams 132, 141, 558
 - parameters 239
- Feynman–Kac formula 18
- Feynman's gauge 511, 558
- field integral 75, 106, 118
 - perturbative definition 134
- field renormalization 122, 124, 189
- field RG function: positivity 234
- field theory: holomorphic representation 108
- Fierz transformation 289
- fine-tuning problem 172, 178, 193, 491, 567, 575, 577, 656
- finite size: hypercubic geometry 766
- finite temperature
 - ϕ^4 field theory 790
 - Dirac fermions 269
 - quantum field theory (QFT) 786
 - renormalization group (RG) 788
- finite-size scaling 760
- fixed point Hamiltonian 359
- flavour quantum number 576, 602
- Flory's approximation 379
- fluctuation–dissipation theorem 877
- flux tube 617
- Fock space 74, 108
- Fokker–Planck (FP) equation 835, 850, 857, 864
- Fokker–Planck (FP) Hamiltonian 889
- Fourier representation 144
- free energy 327, 335
- free field theory 127
- functional
 - δ -function 138
 - (field) integral 75
 - differentiation 29, 85, 126
- Furry's theorem 534
- gamma matrices 284
- γ_S 281
- traces 291

- gap equation 498, 812
- gauge
 - coupling to
 - leptons 570
 - quarks 573
 - covariant 511, 553
 - dependence 528, 545
 - field 514, 549, 683
 - fermion vertex 580
 - fixing 511, 516
 - invariance 47, 683
 - invariant action 551
 - invariant operators 530, 647
 - linear 646
 - pure 704
 - section 552
 - symmetry 514
 - theory 507 548
 - semi-classical vacuum 957
 - transformation 47, 511, 514, 548
 - infinitesimal 550
- Gaussian
 - fixed point 364
 - integral 1, 4, 24
 - complex 5
 - Grassmann variables 13, 84
 - relativistic fermions 269
 - random walk 377
 - renormalization 178, 367
 - two-point function 46
 - white noise 831
- Gauss's law 518, 553
- General Relativity (GR) 684
- generating functional 29, 126
- ghost equation of motion 647
- Ginsparg–Wilson relation 278
- gluonium 607
- gluons 567, 576
- Goldberger–Treiman relation 308, 312
- Goldstone bosons 303, 458
- grand canonical formalism 73
- Grassmann algebra 6, 269, 280
 - δ -function 81
 - differentiation 8
 - Gaussian integration 13
- integration 9
- coordinate 632
- gravity
 - classical equations 685
 - quantization 687
- Gribov's ambiguity 555, 874
- Gross–Neveu model 489, 504
 - finite temperature 810
- Gross–Neveu–Yukawa model 489
- Gross–Pitaevski equation 102
- Hamiltonian formalism 551
 - separable 44
- Hamilton–Jacobi equation 58
- harmonic oscillator 24
 - quantum 45
- Heisenberg representation 41
- helium superfluid transition 987
- Hermitian conjugation 266
- hierarchy problem 575
- Higgs
 - boson 559, 569
 - mass: bound 596
 - mechanism 559, 569
 - model: finite temperature 821
 - high temperature 796
 - effective theory 789
 - series 347
 - higher spins: renormalizability 513
 - holomorphic
 - formalism: many-body interpretation 72
 - representation 64
 - field theory 106
 - holonomy variables, group 681
 - homogeneous CS equations 205
 - homogeneous RG equations 208
 - homogeneous spaces 692, 715
 - non-compact 711
 - homotopy classes 954
 - hyperscaling relations 422
 - improved action 430
 - infinitesimal change of variables 136
 - instanton 588, 592, 899, 920
 - CP(N-1) model 953
 - general scalar field theories 924
 - interaction 995, 1012

- renormalization 923
- solution 908
- SU(N) model 956
- integrable field theories: classical 60
- integrable Hamiltonian systems: classical 58
- interaction representation 91, 104
- infrared (IR) divergences 463
- infrared (IR) regulator 558
- irrelevant operator 362, 426
- Ising model 326
- Itô calculus 848
- Jacobi identity 643, 680
- Jacobian 911
- Jacobi's θ -functions 829
- Jacobi's elliptic functions 829
- Jarzynski's relation 847
- Josephson's parametrization 414
- Källen–Lehmann representation 40, 123
- Kosterlitz–Thouless (KT) phase transition 743, 747
- Landau ghost 448, 965
- Landau theory of phase transitions 341
- Landau–Ginzburg–Higgs model: Abelian 535
- Landau–Ginzburg–Wilson's Hamiltonian 367
- Landau's gauge 511, 558
- Langevin equation 831, 857, 864, 875, 948
 - discretized 848
 - dissipative 838, 844
- large
 - field behaviour 971
 - momentum behaviour 593
 - N expansion 449, 453, 800
 - N limit: σ -model 449
 - order behaviour 961, 964, 972, 975, 999, 1006
- lattice
 - fermions: doubling 277
 - gas model 382
 - gauge theories 607
 - QED 616
 - regularization 175, 465, 525
- Legendre transformation 16, 43, 49, 141
- leptons 570
- Lie algebras: conventions 293
- light-cone expansion 256
- Liouville measure 43
- Lippman–Schwinger's equation 94
- liquid–vapour transition 382, 987
- little group 459
- local expansion 426, 477, 789, 814
- locality 20, 105, 167
- localization 711
- logarithmic corrections 423
- loop expansion 150
- loop factor 224
- low temperature expansion 459
- Luttinger liquid 89
- magnetic
 - exponent 338
 - Hamiltonian 47
 - monopole 63, 616
 - susceptibility 338
- Majorana particles 268
- Majorana spinors 268, 657
- many-body quantum mechanics (QM) 117
- marginal operator 362
- Markov property 19
- Markov's process 831, 848, 852
- mass
 - insertion 204
 - operator 144
 - renormalization 189
- massless field theories 204
- master equation 468
- matching conditions 808
- mean-field
 - approximation 332
 - expansion: general formalism 348
 - theory: gauge model 617
- Mermin–Wagner–Coleman theorem 331
- metric tensor 49, 677, 850
- minimal subtraction (MS) 171, 227
 - modified 227
- mode expansion 787
- momentum cut-off 166
 - regularization, gauge theories 523
- Monte-Carlo methods, simulations 619, 875
- multi-instanton 992
- Nambu–Goldstone modes 304

- Nambu–Jona-Lasinio model 743
 neutral currents 570
 neutrino oscillations 572
 Noether’s theorem 317
 noise field 857
 non decoupling of scales 177
 non-Abelian
 anomaly 589
 gauge theory: finite temperature 824
 non-Borel summability 1014
 non-linear σ -model 476
 finite-size effects 776
 $O(2)$ 747
 quantum, finite temperature 799
 non-relativistic limit 116
 non-relaxational model 886
 non-renormalizable interactions 180
 non-renormalizable theories 165
 normal order 66, 70, 86
 normal-ordered vertex 217
 $O(2)$ non-linear σ -model 747
 one-line irreducibility 151
 one-particle irreducible (1PI)
 diagrams 126, 151
 functional: physical interpretation 152
 generating functional 151
 operator
 insertion
 CS equations 243
 renormalization 240, 244
 ordering 48
 power counting 173
 product expansion (OPE) 256
 order-dependent mapping (ODM) 990
 order parameter 324
 Ornstein and Zernike form 339
 overlap fermions 278
 overlapping divergences 196
 Padé approximants 978
 parallel transport 548, 608, 677
 partition function 25, 126
 path integral 22
 Grassmann algebra 86
 holomorphic 67
 phase space 42
 representation 20, 842
 Pauli matrices 284
 Pauli–Villars’s regularization 168, 464
 gauge theories 523
 perimeter law 614
 periodic boundary conditions 25, 30, 760,
 764, 787
 perturbation theory 3, 15, 34, 35, 86, 104,
 128, 270
 quad compact form 128
 perturbative assumption 177
 perturbative expansion 46, 843
 Pfaffian 15
 phase transition
 breaking of ergodicity 324, 856
 first order 337
 second order 337
 physical scale 177
 plaquette action 609
 Poisson brackets 43
 Poisson’s formula 782
 polymer chains 987
 polymers: statistical properties 377
 PMNS (Pontecorvo–Maki–Nakagawa–Sakata)
 matrix 568
 positivity of the field RG function 234
 power counting 163, 179
 power-law behaviour 371
 probability conservation 835
 propagator 128
 proper-time regularization 182
 proper-time representation 181
 quantization 849
 quantum
 chromodynamics (QCD) 567, 576, 607
 electrodynamics (QED) 507
 $Z_A = 1/Z_e$ 528
 compact 616
 field equation 245
 field theory (QFT) 74, 105
 free 107
 non-relativistic 74, 76
 large-order behaviour 972
 one-loop β -function 532
 gravity 670, 687, 689

- jump 934
- particles:
 - bosons 73
 - fermions 87
- partition function 18
- statistical operator 18
- statistical physics:
 - fixed number of particles 72
- quark generations 573
- quark masses 573
- quartic double-well potential 942
- quartic ghost terms 646
- quasi-Gaussian approximation 342
- quenched approximation 620
- quenched average 354
- random magnetic field 874
- random walk 831
- ratio of amplitudes 340
- real-time quantum mechanics 92
- reduced action at one-loop 792
- redundant operator 362
- reflection Hermiticity 266
- reflection transformations 287
- regularization 166
 - dimensional 223
 - lattice 175
 - non-Abelian gauge 559
 - non-linear σ -model 462
 - Pauli–Villars 168
- regulator fields 168
- relaxation rates 875
- relevant operator 361
- renormalizable
 - action 179
 - gauges: general 643
 - theories 166
- renormalization
 - conditions 192
 - constants 189
 - group: see RG 185
 - local monomials 180
 - of gauge theories 623
- renormalization group (RG) 185, 357
 - β -function 578, 708
 - equations 577
- bare 211
- microscopic theory 369
- fixed trajectories 497
- functions 234, 238
 - covariance 209
- transformation 178
- renormalized
 - action 189
 - form of the action 564
 - gauge action 642
 - parameters 188
- renormalon 448, 456, 967
- replica trick 355, 380
- response functions 876
- Ricci tensor 680
- Riemann tensor 679
- Riemannian manifolds 49, 121, 156, 692, 711, 850, 871, 952
- rigid rotator 779
 - $O(2)$ 51
 - $O(N)$ 52
- rigid symmetry 295
- roughening transition 616
- saddle point 4
- scalar curvature 680, 715
- scalar superfield 656
- scale invariance 919
- scaling equation of state 338, 414
- scaling laws 396, 399, 400
- Schwinger model 733
- Schwinger–Dyson equations 134
- screening length 752
- second quantization 73, 74, 80
- self-avoiding random walk (SAW) 377
- self-duality equations 954
- semi-classical expansion 146
- semi-classical limit 899
- short-distance expansion (SDE) 249, 410
 - RG equations 253
- sigma model 306, 458
 - linear 306
- sine-Gordon model (sG) 731, 746, 751
- skeleton diagram 194
- Slavnov–Taylor identities 637
- Slavnov–Taylor symmetry 860

- S-matrix 91, 104, 118
 - functional integral 90
 - gauge independence 648
 - perturbative expansion 92
- Sobolev inequalities 937
- soft symmetry breaking 292
- soliton 746, 899
 - scattering 114
- specific heat exponent α 399
- spin connection 683
- spin group 259, 284
- spin quantization 61
- spinors 259, 266, 286
- spin-statistics connection 262, 272
- spontaneous fermion mass generation 489
- spontaneous magnetization 338
- spontaneous symmetry breaking 302, 462, 490, 943
- stabilizer subgroup 459
- staggered fermions 278
- Standard Model
 - particle physics 591
 - weak-electromagnetic interactions 568
- statistical operator 83, 118
 - phase-space path integral 44
- statistical physics: quantum, classical 26
- steepest descent method 4, 146
- stereographic mapping 940
- stochastic
 - differential equation 831
 - distribution 554
 - process: phase transition 856
 - variables 630
- Stratanovich convention 849
- streaming term 875
- stress tensor 323
- string tension 616
- strong coupling expansion 614
- strong CP violation 958
- strong scaling 394
- super-diagrams 894
- superficial degree of divergence 163
- superfield 844
 - scalar 664
 - vector 666
- super-propagator 660
- super-renormalizable theories 165, 180, 216
- supersymmetric
 - action 665
 - curvature tensor 668
 - non-linear σ -model 661
 - structure 950
 - Ward-Takahashi identities 658
- supersymmetry 262, 655, 843, 845, 864, 877, 882, 889
 - four dimensions 662
 - surface tension 785
 - symmetric space 561, 703, 708
- symmetry
 - breaking
 - linear 298
 - quadratic 313
 - soft 292, 316
 - continuous 763
 - linear 292
 - discrete 316
 - global 295
 - linear realization 292
 - local, gauge 521
 - restoration 947
- symplectic form 56
- 't Hooft's gauge 563
- 't Hooft's or decoupling gauge 538
- temporal gauge 517, 520, 552, 612
- tensors 671
- thermal mass 787
- thermal wavelength 36
- thermodynamic potential: variational principle 153
- theta vacuum 958, 1001
- Thirring model 736
 - SU(2)-invariant 739
 - SU(N)-invariant 742
- time ordering 19
- time-ordered product 41, 104
- topological 52
 - action 60
 - charge 954, 956
 - topology and anomalies 585
 - torsion 675

- T-product: see time-ordered 104
transfer matrix 20, 26, 325
transport coefficients 875
tree order action 189
tricritical behaviour 180
tricritical point 346
triviality issue 213, 309, 375, 496, 593, 595
twist 257
twisted boundary conditions 765
two-instanton configuration 994
two-point correlation function 39
two-point vertex function 492
U(1) problem 592, 958
ultraviolet (UV) divergences 160, 169
ultraviolet (UV) fixed points 593
uniaxial ferromagnet 431
unitarity 111, 262
 relations 530
unitary gauge 511, 536, 561
universality 28, 360, 394
 classes 360
universality 28, 357, 391
 classes 357
unstable vacua in field theory 919
vacuum 118
 energy 542, 543
Van Vleck's formula 913, 918
vector field: massive 508, 561, 565
vertex functions 141
vielbein 682, 851
volume element: invariant 678
vortex excitations 755
Ward–Takahashi (WT) identities 292, 466, 564, 697, 700, 865
weak angle θ_W 569
Weinberg's theorem 205
Wess–Zumino consistency conditions 590
Weyl gauge 520
W-gauge boson 569
Wheeler–DeWitt equation 688
Wick's theorem 3, 31, 34, 128
 complex variables 6
 fermions 14
 Grassmann variables 14
Wilson's fermions 277
Wilson's loop 607
winding number 954, 957
WKB expansion 1016
XY model: classical, two-dimensional 758
Yang–Lee edge singularity 172
Yang–Mills field 549
Yukawa couplings 573
Yukawa-type interaction 489
zero mode 765, 788
Z-gauge boson 569
Zinn-Justin (ZJ) equation 640