Quantum Field Theory and Critical Phenomena

Fourth Edition

JEAN ZINN-JUSTIN

Commissariat à l'Energie Atomique
Direction des Sciences de la matière
Gif sur Yvette, France

CLARENDON PRESS • OXFORD
2002

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

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

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

with an associated company in Berlin

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

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

First edition 1989
Second edition 1993
Third edition 1996
Fourth edition 2002

A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data

(Data available)

ISBN 0 19 850923 5

1 3 5 7 9 10 8 6 4 2

Typeset by the author using TeX and private macros
Printed in Great Britain
on acid-free paper by T.J. International Ltd., Padstow

Preface

The last thirty years have witnessed the spectacular progress of quantum field theory (QFT). Originally introduced to describe quantum electrodynamics (QED), QFT has become the framework for the discussion of all fundamental interactions except gravity. Much more surprisingly, it has also provided the framework for the understanding of second order phase transitions in statistical mechanics. In fact, as will hopefully become clear in this work, QFT is the natural framework for the discussion of most systems in which an infinite number of degrees of freedom are coupled. 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 200 GeV, altogether more than *twenty orders of magnitude* in the energy scale.

Therefore, although several good textbooks about QFT have already been published, I thought that it might not be completely worthless to present a work in which the common aspects of 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 path and functional integrals has been adopted to study the properties of QFT. Less important, the space–time metric has been chosen euclidean, as is natural for statistical mechanics and convenient in general for perturbative calculations even in particle physics. The language of partition and correlation functions has been used even in applications of field theory to particle physics. Renormalization and renormalization group properties have been systematically discussed, whereas little space has been devoted to scattering theory. Only formal aspects of QED have been considered since excellent textbooks already cover this subject.

The idea of renormalizable quantum field theories first appeared empirically in Quantum electrodynamics. QED, as well as all more complete field theories describing particle physics, is plagued by a serious disease. In a straightforward calculation all physical quantities are infinite, due to the short distance singularities of the theory. This situation has to be contrasted with what happens in Classical or non-relativistic Quantum Mechanics; 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. This can be summarized by saying that in the latter theories phenomena of very different scales, to a good approximation, decouple.

A strange remedy to this disease has been found empirically: one artificially modifies the theory at short distance (in a way which, in general, leads to unphysical short distance properties), at a scale characterized by a short distance cut-off. Inspired by methods of condensed matter physics, one then re-expresses all physical quantities in terms of a small number of physical constants, like the physical masses and charges, instead of the original parameters of the lagrangian. After this change of parametrization, the cut-off is removed, and somewhat miraculously all other physical quantities have a finite

Preface

The last thirty years have witnessed the spectacular progress of quantum field theory (QFT). Originally introduced to describe quantum electrodynamics (QED), QFT has become the framework for the discussion of all fundamental interactions except gravity. Much more surprisingly, it has also provided the framework for the understanding of second order phase transitions in statistical mechanics. In fact, as will hopefully become clear in this work, QFT is the natural framework for the discussion of most systems in which an infinite number of degrees of freedom are coupled. 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 200 GeV, altogether more than *twenty orders of magnitude* in the energy scale.

Therefore, although several good textbooks about QFT have already been published, I thought that it might not be completely worthless to present a work in which the common aspects of 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 path and functional integrals has been adopted to study the properties of QFT. Less important, the space–time metric has been chosen euclidean, as is natural for statistical mechanics and convenient in general for perturbative calculations even in particle physics. The language of partition and correlation functions has been used even in applications of field theory to particle physics. Renormalization and renormalization group properties have been systematically discussed, whereas little space has been devoted to scattering theory. Only formal aspects of QED have been considered since excellent textbooks already cover this subject.

The idea of renormalizable quantum field theories first appeared empirically in Quantum electrodynamics. QED, as well as all more complete field theories describing particle physics, is plagued by a serious disease. In a straightforward calculation all physical quantities are infinite, due to the short distance singularities of the theory. This situation has to be contrasted with what happens in Classical or non-relativistic Quantum Mechanics; 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. This can be summarized by saying that in the latter theories phenomena of very different scales, to a good approximation, decouple.

A strange remedy to this disease has been found empirically: one artificially modifies the theory at short distance (in a way which, in general, leads to unphysical short distance properties), at a scale characterized by a short distance cut-off. Inspired by methods of condensed matter physics, one then re-expresses all physical quantities in terms of a small number of physical constants, like the physical masses and charges, instead of the original parameters of the lagrangian. After this change of parametrization, the cut-off is removed, and somewhat miraculously all other physical quantities have a finite limit, when the theory is so-called renormalizable. Moreover, this limit is independent of the precise form of the short distance modification. Applied to QED this strategy led to predictions of extraordinary accuracy. Therefore, it was natural to try also to construct a renormalizable field theory for all other interactions. This led to another great achievement: a model for all three strong, weak and electromagnetic interactions. The so-called Standard Model has now successfully confronted all experimental data, for more than twenty-five years.

As the consequence of these truly remarkable results, renormalizability was then slowly promoted to a kind of additional law of nature. In particular, once the standard model of weak, electromagnetic and strong interactions was established, much effort was devoted to cast gravity in the same framework. Despite many ingenious attempts, no renormalizable form of quantum gravity has been found yet.

Note that it was realized early, first as a mathematical curiosity, that in massless renormalizable theories a renormalization group could be associated with transformation properties under space dilatations. Later it was realized that this property could be used to discuss the short distance structure of some physical processes. The basic idea was to introduce a set of scale-dependent coupling constants. In *asymptotically free* field theories, these effective couplings become small at large euclidean momenta and, therefore, perturbation theory, improved by renormalization group, can be used. Only the theory of Strong interactions, an $SU(3)$ gauge theory, shares this property. However, most of the field theories proposed to describe strong, electromagnetic and weak interactions are not asymptotically free.

More generally, it was suggested by Weinberg that the existence of UV fixed points, that is, the existence of limits for the effective short distance couplings, was a necessary condition for the consistency of a field theory. 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 allow for non-perturbative explorations, have failed to discover non-trivial fixed points. Therefore, it seems that the Standard Model, which describes so precisely Particle Physics at present scale (except for possible neutrino oscillations), is not consistent on all scales and has to be modified at short distance. This is a second indication that maybe the property of renormalizability has a different origin.

Somewhat surprisingly, QFT has also become an essential tool for the understanding of some critical phenomena: second order phase transitions in condensed matter physics. Near the critical temperature, cooperative phenomena generate a large scale, associated with the so-called correlation length, although the fundamental interactions are short range. Moerover, 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 classical ideas: a description in terms only of macroscopic degrees of freedom adapted to the scale of large distance physics. 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. Unfortunately, it became slowly clear that the too universal predictions of such a theory were in conflict with numerical calculations of critical exponents, experimental data and, finally, exact results in two dimensions. All these data still supported the concept of *universality* in the sense that 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 ... Actually, an analysis of leading corrections to mean field or gaussian approximations indeed reveals that, at least in low space dimensions, degrees of freedom associated with shorter distances never completely decouple.

To explain this remarkable situation, that is, that large distance properties of second order phase are to a large extent short distance insensitive, although the degrees of freedom on all scales seem to be coupled, Wilson, partially inspired by some prior attempts of Kadanoff, introduced the renormalization group idea: starting from a microscopic hamiltonian one integrates out the degrees of freedom corresponding to short distance fluctuations and generates a scale-dependent effective hamiltonian. Universality relies then upon the existence of IR fixed points in hamiltonian space. One of the spectacular implications was that the universal properties of a large class of critical phenomena could be accurately predicted by the same field theory methods that had been invented for particle physics. The appearance of renormalizable field theories was there related to the fixed point structure, and the property that the effective theory relevant for long distance physics depended only on a small number of parameters.

Predictions obtained from a renormalization group analysis of simple field theories like the $(\phi^2)^2$ field theory have been successfully compared to experiments as well as numerical data from lattice models. The same field theory 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 quantum field theory has led to an understanding of the concept of universality and allowed the calculation of many universal physical quantities, conversely, critical phenomena have shed a new light on the mysterious role of renormalization and renormalizable field theories in particle physics.

For many years the renormalization procedure has been considered as an *ad hoc* method introduced only in order to calculate physical quantities in perturbation theory in terms of a small number of parameters. One sometimes even tried to put the blame on the perturbative treatment of field theory, and one tried to hide the renormalization procedure as much as possible (like the BPHZ effort). After all it worked and still works.

However, directly inspired from the theory of critical phenomena, another interpretation, originally also proposed by Wilson, has gained strength over the years. New physics should be expected at very short distances (maybe at the Planck's scale). At this scale the familiar notion of renormalizable local field theories probably loses its meaning. However, possible non-local effects are limited to this short scale (the equivalent of the condition of short range forces in statistical systems). In addition, dynamical effects, of a nature which at present can only be guessed, generate long distance physics associated with the appearance of almost massless particles (compared to the Planck mass, for example, all known particles are essentially massless). Nevertheless, because degrees of freedom on all scales remain coupled, the short distance cut-off can never be eliminated from the theory and this explains the impossibility of constructing a finite hamiltonian formalism. Fortunately, some renormalization group is also at work here, in such a way that observations at present energies can be accurately described in terms of an effective long distance renormalizable field theory. Large distance physics is short distance insensitive but in a way that is thus much more subtle than in classical physics.

The classification of interactions in terms of renormalization group properties becomes relevant. The necessity of a UV fixed point disappears, because Green's functions no longer need to be well defined on all scales, and marginal or irrelevant interactions have to be considered. Weak and Electromagnetic interactions are probably marginal and, the free field theory being an infrared fixed point, their strength is proportional to an inverse power of the logarithm of the cut-off: Green's functions are consistent on all scales only for vanishing coupling constants (the *triviality* problem). This would explain why they can be described by a renormalizable theory with small (although much larger than in gravity) coupling constants. For Strong Interactions, the situation seems to be more complicated. One must imagine that the effective interaction first decreases at shorter distances and then increases again because the free field theory is an ultraviolet fixed point. Finally, Gravitation is presumably irrelevant in the sense of critical phenomena, which means that it, is non-renormalizable and, therefore, very weak because, for dimensional reasons, its strength is proportional to a power of the short distance cut-off.

However, 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 massless particles has to be explained from general properties of the unknown fundamental theory. This is the famous *hierarchy* problem. Spontaneous breaking of a continuous symmetry, gauge principle and chiral invariance 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 one of the remaining fundamental problems of quantum field theory. An intriguing question among many is the following: if a theory contains a light vector boson, does the effective theory automatically take the form of a Higgs model?

On the other hand, since the large distance physics is short distance insensitive, the real nature of fundamental interactions may remain elusive in the foreseeable future, in the same way that a precise knowledge of the critical exponents of the liquid–vapour phase transition gives very little information about the molecular 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. It can be roughly divided into four parts. Chapters 1–12 deal with general field theory, functional integrals and functional methods. An introduction to renormalization theory is provided on the simple example of the ϕ^4 field theory and renormalization group (Callan–Symanzik) equations are derived. In Chapters 13–21 renormalization properties of theories with symmetries are studied and specific applications to particle physics are emphasized. Chapter 22 gathers the few elements of classical and quantum gravity needed elsewhere in the work. Chapters 23–37 are mainly devoted to critical phenomena. A brief introduction to lattice gauge theories is included and asymptotic freedom in four dimensions (at large momentum, a problem relevant to Particle Physics) is discussed (Chapters 34, 35). Chapter 38 provides an introduction to finite temperature relativistic quantum field theory. Chapters 39–43 describe the role of instantons in quantum mechanics and field theory, the application of instanton calculus to the analysis of large order behaviour of perturbation theory and the problem of summation of the perturbative expansion.

Note that, for lack of space, the exercises of the third edition have been removed. They will be available under the address **www-spht.cea.fr/articles/T02/001**.

I am perfectly 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. However, I also believe that a complete survey of quantum field theory and its applications is now beyond the scope of a single physicist and can only be produced by a more collective effort.

This work incorporates notes of lectures delivered in numerous summer schools, most notably, Cargèse 1973, Bonn 1974, Karpacz 1975, Basko Polje 1976 and Les Houches 1982, as well as notes prepared for graduate courses in Princeton, Louvain-la-Neuve, Berlin, Lausanne, Cambridge (Harvard) and universities in the Paris area.

Finally remarks, comments, corrections... are welcome and can be sent to the Email address: zinn@spht.saclay.cea.fr .

Acknowledgments

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. However, I wish to specially thank my masters M. Froissart and D. Bessis who guided my first steps in physics, E. Brézin and J.C. Le Guillou who have collaborated with me for more than fifteen years and without whom obviously this work would never have been produced. I also think with a deep emotion of the late 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 renormalization group ideas are a major source of inspiration for this work. S. Coleman and the late K. Symanzik played a very important role in my understanding of several aspects of physics through their articles and lecture notes as well as in 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.

Several colleagues agreed to read part of the manuscript before publication and I have benefited from their criticisms, remarks and wisdom, in particular E. Brézin who directly collaborated in the first chapters of the initial edition, R. Stora, C. Bervillier, E. Iancu, O. Napoly, R. Guida, A.N. Vasil'ev, J.-B. Zuber and my son Paul.

More generally, my colleagues of the Saclay theory group, with whom I have had so many discussions, in particular C. de Dominicis and the late C. Itzykson, have directly influenced this work. Finally, I also wish to thank all the other colleagues with whom I have collaborated over the years, and more specifically P. Ginsparg, 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 recent stay at MIT where the notes on finite temperature quantum field theory were prepared is gratefully acknowledged.

S. Zaffanella typed a large part of the first edition with care and competence. M. Porneuf tracked in it with persistence the innumerable misprints.

All deserve my deepest gratitude.

4$^{\text{th}}$ edition, Saclay, 24 March 2002

General References

Several textbooks or reviews have been a direct source of inspiration or complement naturally the present work (or both):

N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, New York 1959);

J. Bjorken and S. Drell, *Relativistic Quantum Mechanics, and Relativistic Quantum Fields* (McGraw-Hill, New York 1964, 1965);

F.A. Berezin, *The Method of Second Quantization* (Academic Press, New York 1966);

S. Weinberg, *Gravitation and Cosmology* (John Wiley & Sons, New York 1972) and *The Quantum Theory of Fields*, 2 volumes, (Cambridge University Press, Cambridge 1995, 1996);

E. S. Abers and B.W. Lee, *Phys. Rep.* 9C (1973) 141;

K.G. Wilson and J.B. Kogut, *Phys. Rep.* 12C (1974) 75;

Methods in Field Theory, R. Balian and J. Zinn-Justin eds., Proceedings of Les Houches Summer School 1975 (North-Holland, Amsterdam 1976);

Phase Transitions and Critical Phenomena, vol. 6, C. Domb and M.S. Green eds. (Academic Press, London 1976), in particular the review *Field Theoretical Approach to Critical Phenomena* by E. Brézin, J.C. Le Guillou and J. Zinn-Justin;

S.K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA 1976);

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

C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York 1980);

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

P. Ramond, *Field Theory: A Modern Primer* (Benjamin, Reading, MA 1980);

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

Recent Advances in Field Theory and Statistical Mechanics, R. Stora and J.-B. Zuber eds., Proceedings of Les Houches Summer School 1982, (Elsevier, Amsterdam 1984);

S. Coleman, *Aspects of Symmetry* (Cambridge University Press, Cambridge 1985);

G. Parisi, *Statistical Field Theory* (Addison-Wesley, New York 1988);

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 University Press, Cambridge 1989);

M. Le Bellac, *Des Phénomènes Critiques aux Champs de Jauge* (Editions du CNRS, 1988), English version: *Quantum and Statistical Field Theory* (Oxford Univ. Press, Oxford 1992);

M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, USA 1995);

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);

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

Contents

1 ALGEBRAIC PRELIMINARIES

It is somewhat unusual to begin a physics textbook with algebraic identities, which are in general hidden in appendices. However, our discussion of perturbative aspects of quantum mechanics and quantum field theory is entirely based on path or functional integrals and more generally functional techniques. Therefore, a reader not familiar with these concepts may find it difficult to follow the algebraic manipulations that enter into the derivation of many results. Moreover, we want to indicate by such a choice that the various technical difficulties that we shall meet will in general be directly confronted rather than carefully hidden.

Therefore, in this first chapter, we recall a few algebraic identities about gaussian integrals, in particular Wick's theorem, a result also relevant for gaussian probability distributions. We discuss the steepest descent method, which reduces a certain type of integrals to gaussian expectation values.

We then define and discuss a few properties of differentiation and integration in a Grassmann, that is, antisymmetric algebra, relevant for theories with fermion particles. In particular, we calculate gaussian integrals and again reduce general integrals to gaussian expectation values.

Throughout the chapter, all expressions are given for a finite but arbitrary number of variables, because the focus is mainly on algebraic properties. However, the generalization to an infinite number of variables is simple, as will be discussed in the following chapters.

We also recall the concept of Legendre transformation, generating functional, functional differentiation and the algebraic definition of the determinant of an operator.

Notation. In this chapter, as well as in the whole work, we will use the convention of ***summation over repeated indices*** when necessary. Exceptions to this rule will be stated explicitly.

Partial derivatives of a function of several variables x_i will be denoted either explicitly by $\partial/\partial x_i$ or when the notation is not ambiguous simply by ∂_i.

Finally, boldface will denote a matrix or a vector in its entirety and the corresponding italics with indices will denote elements.

1.1 Gaussian Integrals

In this section, we briefly review a few algebraic properties of gaussian integrals in the case of a finite number of integration variables.

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

$$\mathcal{Z}(\mathbf{A}) = \int \mathrm{d}^n x \, \mathrm{e}^{-A_2(\mathbf{x})} \tag{1.1}$$

with

$$A_2(\mathbf{x}) \equiv \tfrac{1}{2} \sum_{i,j=1}^{n} x_i A_{ij} x_j , \tag{1.2}$$

where the matrix $\mathbf{A}$ of elements A_{ij} is complex symmetric with a non-negative real part, and non-vanishing eigenvalues a_i:

$$\mathrm{Re}\,\mathbf{A} \geq 0\,, \quad a_i \neq 0\,.$$

When the matrix $\mathbf{A}$ is real, it can be diagonalized by an orthogonal transformation matrix $\mathbf{O}$. Changing variables $x_i \mapsto x'_i$:

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

a transformation of jacobian unity, we obtain a product of independent x'_i integrals. Each integral yields a factor $\sqrt{2\pi/a_i}$. The result thus involves the product of all eigenvalues, that is, the determinant

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

Moreover, since both the initial integral and the determinant are analytic functions of the coefficients of the matrix $\mathbf{A}$, the identity can be extended by analytic continuation to the complex case.

We now consider the more general integral

$$\mathcal{Z}(\mathbf{A},\mathbf{b}) = \int \mathrm{d}^n x\; \mathrm{e}^{-A_2(\mathbf{x})+\mathbf{b}\cdot\mathbf{x}}, \quad \mathbf{b}\cdot\mathbf{x} \equiv \sum_{i=1}^{n} b_i x_i\,, \tag{1.4}$$

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

$$\frac{\partial}{\partial x_i}\left(A_2(\mathbf{x}) - \mathbf{b}\cdot\mathbf{x}\right) = 0 \quad \Rightarrow \quad \sum_j A_{ij} x_j = b_i\,.$$

The solution is

$$x_i = \sum_j \left(A^{-1}\right)_{ij} b_j\,. \tag{1.5}$$

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

$$x_i = \sum_j \left(A^{-1}\right)_{ij} b_j + y_i \quad \Rightarrow \quad -A_2(\mathbf{x}) + \mathbf{b}\cdot\mathbf{x} = w_2(\mathbf{b}) - A_2(\mathbf{y}) \tag{1.6}$$

with

$$w_2(\mathbf{b}) = \frac{1}{2}\sum_{i,j=1}^{n} b_i \left(A^{-1}\right)_{ij} b_j\,. \tag{1.7}$$

The integral becomes

$$\mathcal{Z}(\mathbf{A},\mathbf{b}) = \mathrm{e}^{w_2(\mathbf{b})} \int \mathrm{d}^n y\; \mathrm{e}^{-A_2(\mathbf{y})} = (2\pi)^{n/2} (\det \mathbf{A})^{-1/2}\, \mathrm{e}^{w_2(\mathbf{b})}\,. \tag{1.8}$$

Gaussian expectation values. We now consider expectation values of polynomials with a gaussian distribution:

$$\langle x_{k_1} x_{k_2} \dots x_{k_\ell} \rangle \equiv \mathcal{Z}^{-1}(\mathbf{A},0) \int \mathrm{d}^n x\, x_{k_1} x_{k_2} \dots x_{k_\ell}\, \mathrm{e}^{-A_2(\mathbf{x})}, \tag{1.9}$$

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

From expression (1.4), one derives

$$\frac{\partial}{\partial b_k} \mathcal{Z}(\mathbf{A}, \mathbf{b}) = \int \mathrm{d}^n x \, x_k \, \mathrm{e}^{-A_2(\mathbf{x}) + \mathbf{b}\cdot\mathbf{x}} . \tag{1.10}$$

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

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

and replacing the integral $\mathcal{Z}(\mathbf{A}, \mathbf{b})$ by its explicit form (1.8),

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

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

$$\langle F(x) \rangle = \left\{ F\left(\frac{\partial}{\partial b} \right) \mathrm{e}^{w_2(\mathbf{b})} \right\} \Bigg|_{\mathbf{b}=0} . \tag{1.12}$$

Wick's theorem. From the identity (1.11) follows one form of Wick's theorem. Each time a differential operator acts on the exponential in the r.h.s. it generates a factor b. A second differential operator has to act on the same factor, otherwise the corresponding contribution vanishes when one sets $\mathbf{b} = 0$. Therefore, the expectation value of the product $x_{k_1} \dots x_{k_\ell}$ with the gaussian weight $\mathrm{e}^{-A_2(\mathbf{x})}$ is obtained in the following way: one considers all possible pairings of the indices $k_1, \dots, k_\ell$ (ℓ must thus be even). To each pair $k_p k_q$ is associated the matrix element $(A^{-1})_{k_p k_q}$ of the matrix $\mathbf{A}^{-1}$. Then,

$$\langle x_{k_1} \dots x_{k_\ell} \rangle = \sum_{\substack{\text{all possible pairings} \\ P \text{ of } \{k_1 \dots k_\ell\}}} A^{-1}_{k_{P_1} k_{P_2}} \dots A^{-1}_{k_{P_{\ell-1}} k_{P_\ell}} , \tag{1.13}$$

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

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

1.2 Perturbation Theory. Connected Contributions. Steepest Descent

1.2.1 Perturbation theory

We now consider a more general integral:

$$\mathcal{Z}(\lambda) = \int \mathrm{d}^n x \, \exp\left(-A_2(\mathbf{x}) - \lambda V(x) \right) , \tag{1.15}$$

in which $A_2(\mathbf{x})$ is the quadratic form (1.2), $V(x)$ a polynomial in the variables x_i and λ a parameter. To calculate the integral, we can expand the integrand in powers of λ:

$$\mathcal{Z}(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int \mathrm{d}^n x \, \mathrm{e}^{-A_2(\mathbf{x})} \, V^k(x).$$

The successive terms in the expansion are gaussian expectation values of polynomials, which can be evaluated by Wick's theorem (1.13):

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

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 (1.7)):

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

1.2.2 Connected contributions

In the expansion (1.16) at order λ we find $\langle V(x) \rangle$. At next order appears the expectation value $\langle V(x) V(x) \rangle$. Using Wick's theorem, we see 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 $\langle \bullet \rangle_c$:

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

($\langle V(x) \rangle_c = \langle V(x) \rangle$). 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 $k = k_1 + k_2 + \cdots + k_p$. The corresponding contribution

$$\left\langle V^{k_1}(x) \right\rangle_c \left\langle V^{k_2}(x) \right\rangle_c \ldots \left\langle V^{k_p}(x) \right\rangle_c ,$$

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 in clusters of $k_1 + k_2 \cdots$,

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

If, instead, k_i appears m times in the decomposition we have the same contribution $m!$ times and we have, therefore, to divide by $m!$. If we sum these contributions, we find

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

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

1.2.3 Steepest descent

In the case of contour integrals in the complex domain, one sometimes uses a method, steepest descent, which reduces their evaluation to gaussian integrals. Let us consider the integral

$$\mathcal{I}(\lambda)=\int \mathrm{d}^n x\ \mathrm{e}^{-A(\mathbf{x})/\lambda}, \tag{1.19}$$

where $A(x)$ is an analytic function of the variables x_i. In the limit $\lambda\to 0$, the integral is dominated by saddle points $\mathbf{x}^c$:

$$\left.\frac{\partial}{\partial x_i}A(\mathbf{x})\right|_{\mathbf{x}=\mathbf{x}^c}=0\,. \tag{1.20}$$

To calculate the contribution of the leading saddle point $\mathbf{x}^c$, we change variables, setting

$$\mathbf{x}=\mathbf{x}^c+\mathbf{y}\sqrt{\lambda}\,. \tag{1.21}$$

We then expand $A(\mathbf{x})$ in powers of λ (and thus y):

$$\frac{1}{\lambda}A(\mathbf{x})=\frac{1}{\lambda}A(\mathbf{x}^c)+\frac{1}{2!}\frac{\partial^2 A(\mathbf{x}^c)}{\partial x_i x_j}y_i y_j+\sum_{k=3}^{\infty}\frac{\lambda^{k/2-1}}{k!}\frac{\partial^k A(\mathbf{x}^c)}{\partial x_{i_1}\ldots\partial x_{i_k}}y_{i_1}\ldots y_{i_k}\,. \tag{1.22}$$

(Summation over repeated indices is implied here.) The change of variables is such that the term quadratic in $\mathbf{y}$ is independent of λ. The integral becomes

$$\mathcal{I}(\lambda)=\lambda^{n/2}\,\mathrm{e}^{-A(\mathbf{x}^c)/\lambda}\int \mathrm{d}^n y\ \exp\left[-\frac{1}{2!}\frac{\partial^2 A(\mathbf{x}^c)}{\partial x_i\partial x_j}y_i y_j-R(\mathbf{y})\right], \tag{1.23}$$

$$R(\mathbf{y})=\sum_{k=3}^{\infty}\frac{\lambda^{k/2-1}}{k!}\frac{\partial^k A(\mathbf{x}^c)}{\partial x_{i_1}\ldots\partial x_{i_k}}y_{i_1}\ldots y_{i_k}\,. \tag{1.24}$$

We then expand the integrand in powers of $\sqrt{\lambda}$. At leading order, we find

$$\mathcal{I}(\lambda)\underset{\lambda\to 0}{\sim}(2\pi\lambda)^{n/2}\left[\det\frac{\partial^2 A(\mathbf{x}^c)}{\partial x_i\partial x_j}\right]^{-1/2}\mathrm{e}^{-A(\mathbf{x}^c)/\lambda}\,. \tag{1.25}$$

At higher orders, the calculation involves expectation values of polynomials with a gaussian weight.

1.3 Complex Structures

We shall often meet complex structures: we have $2n$ integration variables $\{x_i\}$ and $\{y_i\}$, $i=1,\ldots,n$, and the integrand is invariant under a simultaneous identical rotation in all (x_i,y_i) planes. It is then natural to introduce formal complex variables z_i and $\bar z_i$ which, for normalization purposes, we define by

$$z_i=(x_i+iy_i)/\sqrt{2}\,,\qquad \bar z_i=(x_i-iy_i)/\sqrt{2}\,. \tag{1.26}$$

Note, however, that z_i and $\bar z_i$ are *independent integration variables* and only formally complex conjugates since x_i and y_i could correspond to contour integrals and themselves be complex.

The generic gaussian integral now is

$$\mathcal{Z}(\mathbf{A};\mathbf{b},\bar{\mathbf{b}}) = \int \left(\prod_{i=1}^{n} \frac{\mathrm{d}z_i \mathrm{d}\bar{z}_i}{2i\pi}\right) \exp\left[-\sum_{i,j=1}^{n} \bar{z}_i A_{ij} z_j + \sum_{i=1}^{n} \left(\bar{b}_i z_i + b_i \bar{z}_i\right)\right], \tag{1.27}$$

in which $\mathbf{A}$ is a complex matrix with non-vanishing determinant.

The calculation proceeds as in the real case. We first eliminate the terms linear in z_i and $\bar{z}_i$ by a change of variables $z_i \mapsto v_i$, $\bar{z}_i \mapsto \bar{v}_i$,

$$z_i = v_i + \sum_j \left(A^{-1}\right)_{ij} b_j \,, \qquad \bar{z}_i = \bar{v}_i + \sum_j \bar{b}_j \left(A^{-1}\right)_{ji} . \tag{1.28}$$

The gaussian integral can then be calculated either by returning to the "real" variables (1.26) or by a change of variables like $\sum_j A_{ij} v_j = v'_i$. The result is

$$\mathcal{Z}(\mathbf{A};\mathbf{b},\bar{\mathbf{b}}) = (\det \mathbf{A})^{-1} \exp\left[\sum_{i,j=1}^{n} \bar{b}_i \left(A^{-1}\right)_{ij} b_j\right]. \tag{1.29}$$

By systematically differentiating with respect to b_i and $\bar{b}_j$, one establishes Wick's theorem for expectation values with the gaussian weight $\exp(-\bar{z}_i A_{ij} z_j)$. Each derivative with respect to b has to be paired with a derivative with respect to $\bar{b}$, otherwise, the contribution vanishes. Only monomials with an equal number of factors z and $\bar{z}$ have a non-vanishing expectation value:

$$\begin{aligned}\langle z_{i_1}\bar{z}_{j_1}\ldots z_{i_\ell}\bar{z}_{j_\ell}\rangle &= \sum_{\substack{\text{all permutations}\\ P \text{ of } \{j_1,\ldots,j_\ell\}}} A^{-1}_{i_1 j_{P_1}} A^{-1}_{i_2 j_{P_2}} \ldots A^{-1}_{i_\ell j_{P_\ell}} \\ &= \sum_{\substack{\text{all permutations}\\ P \text{ of } \{j_1,\ldots,j_\ell\}}} \langle z_{i_1}\,\bar{z}_{j_{P_1}}\rangle \langle z_{i_2}\,\bar{z}_{j_{P_2}}\rangle \ldots \langle z_{i_\ell}\,\bar{z}_{j_{P_\ell}}\rangle .\end{aligned} \tag{1.30}$$

1.4 Grassmann Algebras. Differential Forms

We shall also deal with theories containing fermions. Since fermion wave functions, field correlation functions (or Green's 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.

Grassmann algebra. A Grassmann (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

$$\theta_i\theta_j + \theta_j\theta_i = 0\,, \qquad \forall i,j\,. \tag{1.31}$$

As a consequence:

(i) all elements in a Grassmann algebra are first degree polynomials in each generator;

(ii) if the algebra has a finite number n of generators, the elements of the algebra form a finite-dimensional vector space on $\mathbb{R}$ or $\mathbb{C}$ of dimension 2^n, spanned by the unit 1 and the products $\theta_{i_1}\theta_{i_2}\ldots\theta_{i_p}$ with $i_1 < i_2 < \cdots < i_p$.

$\mathfrak{A}$ is also a graded algebra in the sense that to any monomial $\theta_{i_1}\theta_{i_2}\dots\theta_{i_p}$ can be associated an integer p counting the number of generators in the product.

Finally, the elements of $\mathfrak{A}$ are invertible if and only if their expansion as a sum of products of generators contains a term of degree zero which is invertible. For example, the element $1+\theta$ is invertible, and has $1-\theta$ as inverse, but θ is not invertible.

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

$$\mathrm{P}(\theta_i) = -\theta_i \;\Rightarrow\; \mathrm{P}^2 = \mathbf{1}\,. \tag{1.32}$$

Then on a monomial of degree p, P acts like

$$\mathrm{P}\left(\theta_{i_1}\dots\theta_{i_p}\right) = (-1)^p\theta_{i_1}\dots\theta_{i_p}\,. \tag{1.33}$$

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

$$\mathrm{P}\left(\mathfrak{A}^{\pm}\right) = \pm\mathfrak{A}^{\pm}. \tag{1.34}$$

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

Differential forms. An application of Grassmann algebras is the representation of differential forms. The language of differential forms will not often be used in this work. However, it is interesting to recall, here, one concept, the exterior derivative of forms, whose generalization will appear in the context of BRS symmetry (see Chapter 16). We consider totally antisymmetric tensors $\Omega_{\mu_1,\dots,\mu_l}(x)$, 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

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

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

One can define a differential operator d acting on forms

$$\mathrm{d} \equiv \theta^\mu \frac{\partial}{\partial x^\mu}\,. \tag{1.36}$$

We note that if Ω is an l-form, $\mathrm{d}\Omega$ is an $(l+1)$-form (see Chapter 22 for details). One immediately verifies that d is *nilpotent* of vanishing square:

$$\mathrm{d}^2 = \theta^\mu \frac{\partial}{\partial x^\mu}\theta^\nu \frac{\partial}{\partial x^\nu} = 0\,, \tag{1.37}$$

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

We also recall that a form Ω which satisfies $\mathrm{d}\Omega = 0$ is called *closed* and a form Ω which can be written as $\Omega = \mathrm{d}\Omega'$ 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 algebra dx^μ instead of θ^μ and then uses the notation $\wedge$ for the product to indicate that it is antisymmetric.

1.5 Differentiation in Grassmann Algebras

It is useful to define differentiation in Grassmann algebras. A naive definition would be inconsistent due to the non-commutative character of the algebra. The problem can be solved in the following way: considered as functions of one 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

$$\frac{\partial A}{\partial \theta_i} = A_2 . \tag{1.38}$$

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

Remark. 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

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

For the second term in the r.h.s. the order between factors matters.

Formal construction. To show the consistency of the definition (1.38) and exhibit some properties, we now define differentiation in a Grassmann algebra more generally by some formal rules, similar to but slightly different from those used in commutative algebras. A Grassmann differential operator D (also called an anti-derivation) acting on $\mathfrak{A}$ is defined by the following two properties:

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

$$\mathrm{D}\left(\lambda_1 A_1 + \lambda_2 A_2\right) = \lambda_1 \mathrm{D}\left(A_1\right) + \lambda_2 \mathrm{D}\left(A_2\right) \qquad \text{for } \lambda_1, \lambda_2 \in \mathbb{R} \text{ or } \mathbb{C} , \tag{1.40}$$

(ii) It satisfies the condition

$$\mathrm{D}\left(A_1 A_2\right) = \mathrm{P}\left(A_1\right) \mathrm{D}\left(A_2\right) + \mathrm{D}\left(A_1\right) A_2 . \tag{1.41}$$

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

$$\mathrm{DP} + \mathrm{PD} = 0 , \tag{1.42}$$

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,

$$\mathrm{D}\left[F(A)\right] = \mathrm{D}(A) F'(A) \qquad \text{for } A \in \mathfrak{A}^+ . \tag{1.43}$$

Note finally 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 (1.40,1.41), then the anticommutator

$$\Delta = \mathrm{D}\mathrm{D}' + \mathrm{D}'\mathrm{D}\,, \tag{1.44}$$

is a usual differential operator:

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

Furthermore,

$$\Delta \mathrm{P} = \mathrm{P}\Delta\,. \tag{1.46}$$

These properties, which are the consequence of the addition of relation (1.42) to the definitions (1.40,1.41), allow to extend the notion of Lie algebra and are directly relevant to the discussion of supersymmetries.

A basis. Since a differential operator satisfies conditions (1.40,1.41), 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 (1.40,1.41). We conclude that any differential operator can be expanded on a basis of operators $\partial/\partial\theta_i$ defined by

$$\frac{\partial}{\partial\theta_i}\theta_j = \delta_{ij} \tag{1.47}$$

with left coefficients in $\mathfrak{A}^+$. One verifies 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

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

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

$$\mathrm{D}_i^{\pm} = \frac{\partial}{\partial\theta_i} \pm \theta_i\,,$$

which satisfy

$$\{\mathrm{D}_i^{\pm}, \mathrm{D}_j^{\pm}\} = \pm 2\delta_{ij}\,, \quad \{\mathrm{D}_i^{+}, \mathrm{D}_j^{-}\} = 0\,. \tag{1.49}$$

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

1.6 Integration in Grassmann Algebras

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

$$\int \mathrm{d}\theta_i\, A \equiv \frac{\partial}{\partial\theta_i} A, \qquad \forall A \in \mathfrak{A}\,. \tag{1.50}$$

The integral or derivative symbols will be used depending on the context.

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

$$\Delta = \mathrm{D}\mathrm{D}' + \mathrm{D}'\mathrm{D}\,, \tag{1.44}$$

is a usual differential operator:

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

Furthermore,

$$\Delta \mathrm{P} = \mathrm{P}\Delta\,. \tag{1.46}$$

These properties, which are the consequence of the addition of relation (1.42) to the definitions (1.40,1.41), allow to extend the notion of Lie algebra and are directly relevant to the discussion of supersymmetries.

A basis. Since a differential operator satisfies conditions (1.40,1.41), 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 (1.40,1.41). We conclude that any differential operator can be expanded on a basis of operators $\partial/\partial\theta_i$ defined by

$$\frac{\partial}{\partial\theta_i}\theta_j = \delta_{ij} \tag{1.47}$$

with left coefficients in $\mathfrak{A}^+$. One verifies 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

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

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

$$\mathrm{D}_i^{\pm} = \frac{\partial}{\partial\theta_i} \pm \theta_i\,,$$

which satisfy

$$\{\mathrm{D}_i^{\pm}, \mathrm{D}_j^{\pm}\} = \pm 2\delta_{ij}\,, \quad \{\mathrm{D}_i^{+}, \mathrm{D}_j^{-}\} = 0\,. \tag{1.49}$$

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

1.6 Integration in Grassmann Algebras

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

$$\int \mathrm{d}\theta_i\, A \equiv \frac{\partial}{\partial\theta_i}A, \qquad \forall A \in \mathfrak{A}\,. \tag{1.50}$$

The integral or derivative symbols will be used depending on the context.

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

$$\mathrm{I}\left(\lambda_1 A_1 + \lambda_2 A_2\right) = \lambda_1 \mathrm{I}\left(A_1\right) + \lambda_2 \mathrm{I}\left(A_2\right), \tag{1.51}$$

which satisfies the three properties:

$$\mathrm{ID} = 0\,, \tag{1.52}$$

$$\mathrm{DI} = 0\,, \tag{1.53}$$

and

$$\mathrm{D}(A) = 0 \Longrightarrow \mathrm{I}(BA) = \mathrm{I}(B)A\,. \tag{1.54}$$

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

$$\mathrm{PI} + \mathrm{IP} = 0\,.$$

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 $\mathrm{D}^2 = 0$, D itself satisfies all conditions. The differential operators $\partial/\partial\theta_i$ indeed have a vanishing square.

1.6.1 Change of variables in a Grassmann integral

We consider the integral

$$I = \int \mathrm{d}\theta\, f(\theta), \tag{1.55}$$

and perform the (necessarily) affine change of variables:

$$\theta = a\theta' + b\,, \tag{1.56}$$

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) we find

$$\int \mathrm{d}\theta\, f(\theta) = a^{-1} \int \mathrm{d}\theta'\, f(a\theta' + b). \tag{1.57}$$

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. This difference reflects the identity between differentiation and integration for Grassmann variables.

Generalization. More generally, a change of variables

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

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 $\partial\theta_i/\partial\theta'_j$:

$$\mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n = \mathrm{d}\theta'_1 \ldots \mathrm{d}\theta'_n J(\theta') \tag{1.58}$$

with

$$J^{-1} = \det \frac{\partial\theta_i}{\partial\theta'_j}\,. \tag{1.59}$$

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

To prove the result, we again start from the identity between differentiation and integration:

$$\int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n\, f(\boldsymbol{\theta}) \equiv \prod_i \frac{\partial}{\partial\theta_i} f(\boldsymbol{\theta}),$$

in which the product on the l.h.s. is ordered. We then assume that f is a function of variables θ'_i and, therefore, using the chain rule (1.39):

$$\prod_i \frac{\partial}{\partial\theta_i} f(\boldsymbol{\theta}) = \prod_i \frac{\partial\theta'_{j_i}}{\partial\theta_i}\frac{\partial}{\partial\theta'_{j_i}} f(\boldsymbol{\theta}).$$

We now factorize the elements $\partial\theta'_{j_i}/\partial\theta_i$ which commute. The differential operators $\partial/\partial\theta'_{j_i}$ anticommute (see equations (1.48)) and are thus all proportional to the product ordered from 1 to n. A sign is generated which is the signature of the permutation $j_1, j_2, \ldots, j_n$. We then recognize the determinant of the matrix $\partial\theta'_j/\partial\theta_i$:

$$\prod_i \frac{\partial}{\partial\theta_i} = \det \frac{\partial\theta'_j}{\partial\theta_i} \prod_k \frac{\partial}{\partial\theta'_k}.$$

The identity between differentiation and integration then immediately leads to the equations (1.58,1.59).

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

$$1 = \int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n\, \theta_n \ldots \theta_1\,.$$

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

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

the result relies upon the identity

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

1.6.2 Mixed change of variables

In this work, we shall 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):

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

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:

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

(ii) One then goes from (θ, x') to (θ', x'). The second step just gives, as explained above, the jacobian J_2:

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

The complete jacobian J, also called the ***berezinian*** of the matrix of derivatives, is thus

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

For the jacobian to be non-singular, the matrices $\mathbf{A}$ and $\mathbf{D}$ have to be invertible (and, therefore, their contributions of degree zero in θ').

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

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

then from identity (1.101), we see that the jacobian has the form

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

We now consider the mixed case:

$$x_a = x'_a + \varepsilon f_a(x', \theta'), \qquad \theta_i = \theta'_i + \varepsilon \varphi_i(x', \theta'). \tag{1.64}$$

Then, setting

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

we find, as a consequence of identity (1.63):

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

To maintain the connection between jacobian and trace, we are, therefore, led to define the supertrace of a mixed matrix, for which the notation Str will be used, as the difference of traces

$$\operatorname{Str}\mathbf{M}_1 = \operatorname{tr}\mathbf{A}_1 - \operatorname{tr}\mathbf{D}_1. \tag{1.66}$$

One verifies that the supertrace, like the usual trace, has a cyclic property, $\operatorname{Str}\mathbf{M}_1\mathbf{M}_2 = \operatorname{Str}\mathbf{M}_2\mathbf{M}_1$.

1.7 Gaussian Integrals with Grassmann Variables

In this work, we shall mainly deal with Grassmann algebras in which the generators can be separated into two conjugated sets. We then denote by θ_i and $\bar{\theta}_i$, $i = 1, \ldots, n$, these generators. In many examples, a complex conjugation can be defined which exchanges θ_i and $\bar{\theta}_i$.

As in the case of commuting variables, we now calculate gaussian integrals, with the same motivation: we will try to reduce more general integrals to a finite or formal infinite sum of gaussian integrals.

We first consider

$$\mathcal{Z}(\mathbf{M}) = \int \mathrm{d}\theta_1 \mathrm{d}\bar{\theta}_1 \mathrm{d}\theta_2 \mathrm{d}\bar{\theta}_2 \ldots \mathrm{d}\theta_n \mathrm{d}\bar{\theta}_n \exp\left(\sum_{i,j=1}^{n} \bar{\theta}_i M_{ij} \theta_j\right). \tag{1.67}$$

According to the rules of Grassmann integration, the result is simply the coefficient of the product $\bar{\theta}_n \theta_n \ldots \bar{\theta}_1 \theta_1$ in the expansion of the integrand. The integrand can be rewritten as (no implicit summation over repeated indices)

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

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

$$\sum_{\substack{\text{permutations} \\ \{j_1 \ldots j_n\}}} M_{nj_n} M_{n-1 j_{n-1}} \ldots M_{1j_1} \bar{\theta}_n \theta_{j_n} \ldots \bar{\theta}_1 \theta_{j_1}.$$

Commuting the generators to put them in the standard order $\bar{\theta}_n \theta_n \ldots \bar{\theta}_1 \theta_1$, we find a sign which is the signature of the permutation, and recognize the coefficient as the determinant of M_{ij}:

$$\mathcal{Z}(\mathbf{M}) = \det \mathbf{M}. \tag{1.68}$$

The result is the inverse of the one obtained with complex commuting variables. This will eventually lead in perturbation theory to a sign $(-1)^L$ in front of the Feynman diagrams with L fermion loops. The calculation above is mainly a verification since we could have changed variables $\theta_i \mapsto \theta'_i$ (provided $\det \mathbf{M} \neq 0$),

$$\sum_j M_{ij} \theta_j = \theta'_i, \tag{1.69}$$

and used the form (1.58,1.59) of the jacobian (no summation over repeated indices)

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

General gaussian integrals. We now introduce another copy of the Grassmann algebra $\mathfrak{A} \equiv \{\theta, \bar{\theta}\}$ and we denote its generators by η_i and $\bar{\eta}_i$. Following the strategy of Section 1.1, we first calculate the integral

$$\mathcal{Z}_{\mathrm{G}}(\eta, \bar{\eta}) = \int \prod_i \mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i \exp\left[\sum_{i,j=1}^{n} M_{ij}\bar{\theta}_i\theta_j + \sum_{i=1}^{n} \left(\bar{\eta}_i\theta_i + \bar{\theta}_i\eta_i\right)\right], \tag{1.70}$$

in which the integrand is an element of the direct sum of the two Grassmann algebras.

The calculation, as before, relies on a change of variables $\theta \mapsto \theta'$, $\bar{\theta} \mapsto \bar{\theta}'$:

$$\theta_i = \theta'_i - \sum_j \left(M^{-1}\right)_{ij} \eta_j, \qquad \bar{\theta}_i = \bar{\theta}'_i - \sum_j \bar{\eta}_j \left(M^{-1}\right)_{ji},$$

and leads to the result

$$\mathcal{Z}_{\mathrm{G}}(\eta, \bar{\eta}) = \det \mathbf{M} \exp\left[-\sum_{i,j=1}^{n} \bar{\eta}_i \left(M^{-1}\right)_{ij} \eta_j\right]. \tag{1.71}$$

Using the notation $\langle \bullet \rangle$ for expectation values with respect to the gaussian weight of equation (1.70), with our definition of $\mathcal{Z}_{\mathrm{G}}$ we find

$$\frac{\partial}{\partial \bar{\eta}_i} \mathcal{Z}_{\mathrm{G}} = \det \mathbf{M} \left\langle \theta_i \right\rangle, \tag{1.72}$$

$$\frac{\partial}{\partial \eta_i} \mathcal{Z}_{\mathrm{G}} = \det \mathbf{M} \left\langle -\bar{\theta}_i \right\rangle. \tag{1.73}$$

Note the sign in equation (1.73).

Wick's theorem for Grassmann integrals. Following the same lines, we derive Wick's theorem for Grassmann gaussian expectation values. Gaussian expectation values are defined by

$$\det \mathbf{M} \left\langle \bar{\theta}_{i_1}\theta_{j_1}\bar{\theta}_{i_2}\theta_{j_2}\ldots\bar{\theta}_{i_n}\theta_{j_n} \right\rangle = \int \left(\prod_i \mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i\right) \bar{\theta}_{i_1}\theta_{j_1}\ldots\bar{\theta}_{i_n}\theta_{j_n} \exp\left(\sum_{i,j=1}^{n} \bar{\theta}_i M_{ij}\theta_j\right), \tag{1.74}$$

From equations (1.72,1.73) it follows that

$$\det \mathbf{M} \left\langle \bar{\theta}_{i_1}\theta_{j_1}\bar{\theta}_{i_2}\theta_{j_2}\ldots\bar{\theta}_{i_n}\theta_{j_n} \right\rangle = \left[\frac{\partial}{\partial \bar{\eta}_{j_1}} \frac{\partial}{\partial \eta_{i_1}} \cdots \frac{\partial}{\partial \bar{\eta}_{j_n}} \frac{\partial}{\partial \eta_{i_n}} \mathcal{Z}_{\mathrm{G}}(\eta, \bar{\eta})\right]\Bigg|_{\eta=\bar{\eta}=0}, \tag{1.75}$$

and using the result (1.71)

$$\begin{aligned}&\left\langle \bar{\theta}_{i_1}\theta_{j_1}\bar{\theta}_{i_2}\theta_{j_2}\ldots\bar{\theta}_{i_n}\theta_{j_n} \right\rangle \\ &= \left\{\frac{\partial}{\partial \bar{\eta}_{j_1}} \frac{\partial}{\partial \eta_{i_1}} \cdots \frac{\partial}{\partial \bar{\eta}_{j_n}} \frac{\partial}{\partial \eta_{i_n}} \exp\left[-\sum_{i,j=1}^{n} \bar{\eta}_j \left(M^{-1}\right)_{ji} \eta_i\right]\right\}\Bigg|_{\eta=\bar{\eta}=0}.\end{aligned} \tag{1.76}$$

After explicit differentiation (which is the same as integration) we obtain

$$\begin{aligned}\langle\bar\theta_{i_1}\theta_{j_1}\ldots\bar\theta_{i_n}\theta_{j_n}\rangle &= \det M^{-1}_{j_l i_k} = \det\langle\bar\theta_{i_k}\theta_{j_l}\rangle\\ &= \sum_{\substack{\text{permutations}\\ P \text{ of } \{j_1\ldots j_n\}}} \epsilon(P)\left(M^{-1}\right)_{j_{P_1} i_1}\left(M^{-1}\right)_{j_{P_2} i_2}\cdots\left(M^{-1}\right)_{j_{P_n} i_n}, \end{aligned}\qquad (1.77)$$

where $\epsilon(P)$ is the signature of the permutation P. This result differs from the expression (1.30), obtained in the case of complex commuting variables, only by the sign $\epsilon(P)$.

Perturbative expansion. Expressions (1.72,1.73) form the basis of perturbation theory. To calculate the integral

$$\mathcal{Z}(\eta,\bar\eta)=\int\prod_i \mathrm{d}\bar\theta_i\mathrm{d}\theta_i\,\exp\left[\sum_{i,j=1}^{n} M_{ij}\bar\theta_i\theta_j+V(\bar\theta,\theta)+\sum_{i=1}^{n}\left(\bar\eta_i\theta_i+\bar\theta_i\eta_i\right)\right],\qquad (1.78)$$

we can formally expand in a power series of V and integrate term by term. We then find

$$\mathcal{Z}(\eta,\bar\eta)=\exp\left[V\left(-\frac{\partial}{\partial\eta},\frac{\partial}{\partial\bar\eta}\right)\right]\mathcal{Z}_{\mathrm{G}}(\eta,\bar\eta).\qquad (1.79)$$

The pfaffian. More generally, we can also calculate gaussian integrals of the form

$$\mathcal{Z}(\mathbf{A})=\int\mathrm{d}\theta_{2n}\ldots\mathrm{d}\theta_2\mathrm{d}\theta_1\,\exp\left(\frac{1}{2}\sum_{i,j=1}^{2n}\theta_i A_{ij}\theta_j\right),\qquad (1.80)$$

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

$$A_{ij}+A_{ji}=0\,.\qquad (1.81)$$

Expanding the exponential in a power series, we observe that only the term of order n which contains all products of degree $2n$ in θ gives a non-zero contribution:

$$\mathcal{Z}(\mathbf{A})=\frac{1}{2^n n!}\int\mathrm{d}\theta_{2n}\ldots\mathrm{d}\theta_1\left(\sum_{ij}\theta_i A_{ij}\theta_j\right)^n.\qquad (1.82)$$

In the expansion of the product only the terms containing a permutation of $\theta_1\ldots\theta_{2n}$ do not vanish. Ordering, then, all terms to factorize the product $\theta_1\theta_2\ldots\theta_{2n}$, we find

$$\mathcal{Z}(\mathbf{A})=\frac{1}{2^n n!}\sum_{\substack{\text{permutations } P\\ \text{of } \{i_1\ldots i_{2n}\}}}\epsilon(P)A_{i_1 i_2}A_{i_3 i_4}\ldots A_{i_{2n-1}i_{2n}},\qquad (1.83)$$

where $\epsilon(P)=\pm1$ is the signature of the permutation P. The quantity in the r.h.s. is called the ***pfaffian*** of the antisymmetric matrix $\mathbf{A}$:

$$\mathcal{Z}(\mathbf{A})=\mathrm{Pf}\,(\mathbf{A}).\qquad (1.84)$$

Grassmann integral techniques allow proving a classical algebraic identity:

$$\mathrm{Pf}^2(\mathbf{A}) = \det \mathbf{A} \,. \tag{1.85}$$

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

$$\mathcal{Z}^2(\mathbf{A}) = \int \mathrm{d}\theta_{2n} \ldots \mathrm{d}\theta_1 \, \mathrm{d}\theta'_{2n} \ldots \mathrm{d}\theta'_1 \, \exp\left[\tfrac{1}{2}\sum_{ij}\left(\theta_i A_{ij}\theta_j + \theta'_i A_{ij}\theta'_j\right)\right]. \tag{1.86}$$

We change variables, setting

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

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

$$\theta_i\theta_j + \theta'_i\theta'_j = \bar\eta_i\eta_j - \bar\eta_j\eta_i \,, \tag{1.87}$$

$$\mathrm{d}\eta_{2n} \ldots \mathrm{d}\eta_1 \mathrm{d}\bar\eta_{2n} \ldots \mathrm{d}\bar\eta_1 = (-1)^{n^2} \prod_i \mathrm{d}\eta_i \mathrm{d}\bar\eta_i \,. \tag{1.88}$$

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

$$\mathrm{Pf}^2(\mathbf{A}) = \int \mathrm{d}\eta_1 \mathrm{d}\bar\eta_1 \ldots \mathrm{d}\eta_{2n} \mathrm{d}\bar\eta_{2n} \, \exp\left(\sum_{ij} \bar\eta_i A_{ij} \eta_j\right) = \det \mathbf{A} \,.$$

Wick's theorem. One can then 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

$$\left\langle \theta_{i_1}\theta_{i_2}\ldots\theta_{i_{2p}}\right\rangle = \sum_{\substack{\text{all possible pairings}\\ \text{of } (1,2,\ldots,2p)}} \epsilon(P) \left\langle \theta_{i_{P_1}}\theta_{i_{P_2}}\right\rangle \ldots \left\langle \theta_{i_{P_{2p-1}}}\theta_{i_{P_{2p}}}\right\rangle, \tag{1.89}$$

in which $\epsilon(P)$ 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, the generating functionals of connected and one-line irreducible (in the sense of Feynman diagrams) correlation functions. We recall here the definition and explain a few simple properties.

We consider a real function $W(\mathbf{h})$ of n variables h_i that has first and second partial derivatives and such that the matrix of elements $\partial_i\partial_j W$ is positive:

$$\forall \mathbf{v}, \quad |\mathbf{v}| > 0 \,, \quad \sum_{ij} v_i \frac{\partial^2 W}{\partial h_i \partial h_j} v_j > 0 \,.$$

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

$$W(\mathbf{h}) + \Gamma(\mathbf{m}) = \sum_i h_i m_i\,, \tag{1.90a}$$

$$m_i = \frac{\partial W}{\partial h_i}\,. \tag{1.90b}$$

The positivity of the matrix of second derivatives ensures that when equation (1.90b) can be inverted it has a unique solution $\mathbf{h}(\mathbf{m})$.

The Legendre transformation is involutive. Indeed differentiating equation (1.90a) with respect to $\mathbf{m}$ one obtains

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

and thus

$$h_i = \frac{\partial \Gamma}{\partial m_i}\,. \tag{1.91}$$

Then, comparing the derivatives of equation (1.91) with respect to $\mathbf{m}$ with the derivatives of (1.90b) with respect to $\mathbf{h}$, one finds that the matrix $\partial_i\partial_j\Gamma(\mathbf{m})$ is the inverse of the matrix $\partial_i\partial_j W$:

$$\sum_k \frac{\partial^2 W}{\partial h_i \partial h_k}\frac{\partial^2 \Gamma}{\partial m_k \partial m_j} = \delta_{ij}\,, \tag{1.92}$$

and thus the matrix $\partial_i\partial_j\Gamma(\mathbf{m})$ is also positive.

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

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

and thus

$$\left.\frac{\partial \Gamma}{\partial \mu}\right|_{\mathbf{m}\ \mathrm{fixed}} + \left.\frac{\partial W}{\partial \mu}\right|_{\mathbf{h}\ \mathrm{fixed}} = 0\,. \tag{1.93}$$

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

$$\mathrm{e}^{W(\mathbf{b})} = \int \mathrm{d}^n x\, \mathrm{e}^{-A(\mathbf{x}) + \mathbf{b}\cdot\mathbf{x}}\,. \tag{1.94}$$

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

$$-\partial_i A(\mathbf{x}) + b_i = 0\,, \tag{1.95}$$

and in the leading approximation

$$W(\mathbf{b}) + A(\mathbf{x}) - \mathbf{b}\cdot\mathbf{x} = 0\,. \tag{1.96}$$

Moreover, the matrix of second derivatives must be positive when the saddle point is real. Therefore, at leading order $W(\mathbf{b})$ and $A(\mathbf{x})$ are related by a Legendre transformation. 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.

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

$$W(\mathbf{h}) + \Gamma(\mathbf{m}) = \sum_i h_i m_i \,, \tag{1.90a}$$

$$m_i = \frac{\partial W}{\partial h_i} \,. \tag{1.90b}$$

The positivity of the matrix of second derivatives ensures that when equation (1.90*b*) can be inverted it has a unique solution $\mathbf{h}(\mathbf{m})$.

The Legendre transformation is involutive. Indeed differentiating equation (1.90*a*) with respect to $\mathbf{m}$ one obtains

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

and thus

$$h_i = \frac{\partial \Gamma}{\partial m_i} \,. \tag{1.91}$$

Then, comparing the derivatives of equation (1.91) with respect to $\mathbf{m}$ with the derivatives of (1.90*b*) with respect to $\mathbf{h}$, one finds that the matrix $\partial_i \partial_j \Gamma(\mathbf{m})$ is the inverse of the matrix $\partial_i \partial_j W$:

$$\sum_k \frac{\partial^2 W}{\partial h_i \partial h_k} \frac{\partial^2 \Gamma}{\partial m_k \partial m_j} = \delta_{ij} \,, \tag{1.92}$$

and thus the matrix $\partial_i \partial_j \Gamma(\mathbf{m})$ is also positive.

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

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

and thus

$$\frac{\partial \Gamma}{\partial \mu}\bigg|_{\mathbf{m}\,\text{fixed}} + \frac{\partial W}{\partial \mu}\bigg|_{\mathbf{h}\,\text{fixed}} = 0 \,. \tag{1.93}$$

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

$$\mathrm{e}^{W(\mathbf{b})} = \int \mathrm{d}^n x \, \mathrm{e}^{-A(\mathbf{x}) + \mathbf{b} \cdot \mathbf{x}} \,. \tag{1.94}$$

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

$$-\partial_i A(\mathbf{x}) + b_i = 0 \,, \tag{1.95}$$

and in the leading approximation

$$W(\mathbf{b}) + A(\mathbf{x}) - \mathbf{b} \cdot \mathbf{x} = 0 \,. \tag{1.96}$$

Moreover, the matrix of second derivatives must be positive when the saddle point is real. Therefore, at leading order $W(\mathbf{b})$ and $A(\mathbf{x})$ are related by a Legendre transformation. 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.

1.9 Generating Functionals. Functional Derivatives. Determinants

We have gathered in this section a few other useful algebraic techniques.

Generating functional. Functional differentiation. In the discussion of algebraic properties of correlation functions the concept of generating functional will be extremely useful. Let $\{F^{(n)}(x_1,\ldots,x_n)\}$, $n=0,1,\ldots$, be a set of symmetric functions of their arguments. We introduce a function of one variable $f(x)$ and consider the following formal series in f:

$$\mathcal{F}(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int \mathrm{d}x_1\ldots\mathrm{d}x_n\,F^{(n)}(x_1,\ldots,x_n)f(x_1)\ldots f(x_n). \tag{1.97}$$

$\mathcal{F}(f)$ is called the *generating functional* of the functions $F^{(n)}$.

To recover the functions $F^{(n)}$ from $\mathcal{F}(f)$, we then need the concept of *functional derivative* $\delta/\delta f(x)$. Functional differentiation obeys the standard algebraic rules (linearity and Leibnitz's rule):

$$\begin{aligned}\frac{\delta}{\delta f(x)}\left[\mathcal{F}_1(f)+\mathcal{F}_2(f)\right]&=\frac{\delta}{\delta f(x)}\mathcal{F}_1(f)+\frac{\delta}{\delta f(x)}\mathcal{F}_2(f),\\ \frac{\delta}{\delta f(x)}\left[\mathcal{F}_1(f)\mathcal{F}_2(f)\right]&=\mathcal{F}_1(f)\frac{\delta}{\delta f(x)}\mathcal{F}_2(f)+\mathcal{F}_2(f)\frac{\delta}{\delta f(x)}\mathcal{F}_1(f),\end{aligned} \tag{1.98}$$

and in addition

$$\frac{\delta}{\delta f(y)}f(x)=\delta(x-y), \tag{1.99}$$

where $\delta(x)$ is Dirac's δ-function. Differentiating $\mathcal{F}(f)$ for example, one finds

$$\frac{\delta}{\delta f(y)}\mathcal{F}(f)=\sum_{n=0}^{\infty}\frac{1}{n!}\int \mathrm{d}x_1\ldots\mathrm{d}x_n\,F^{(n+1)}(y,x_1,\ldots,x_n)f(x_1)\ldots f(x_n), \tag{1.100}$$

and, therefore,

$$F^{(p)}(x_1,\ldots,x_p)=\left.\left\{\left(\prod_{i=1}^{p}\frac{\delta}{\delta f(x_i)}\right)\mathcal{F}\right\}\right|_{f\equiv 0}.$$

Determinants of operators. We have seen that gaussian integrals generate determinants. When the number of integration variables diverges, matrices become operators. We shall, therefore, face the problem of calculating determinants of operators. Let us assume that an operator is represented by a kernel $M(x,y)$ which, after some transformations, can be cast into the form $\delta(x-y)+K(x,y)$. Provided the traces of all powers of $\mathbf{K}$ exist, the following identity, valid for any matrix $\mathbf{M}$,

$$\ln\det\mathbf{M}\equiv\operatorname{tr}\ln\mathbf{M}\,, \tag{1.101}$$

expanded in powers of the kernel $\mathbf{K}$, will often be useful:

$$\begin{aligned}\ln\det\left[\mathbf{1}+\mathbf{K}\right]&=\int \mathrm{d}x\,K(x,x)-\frac{1}{2}\int \mathrm{d}x_1\mathrm{d}x_2\,K(x_1,x_2)K(x_2,x_1)+\cdots\\ &\quad+\frac{(-1)^{n+1}}{n}\int \mathrm{d}x_1\ldots\mathrm{d}x_n\,K(x_1,x_2)K(x_2,x_3)\ldots K(x_n,x_1)+\cdots\,.\end{aligned} \tag{1.102}$$

Bibliographical Notes

Integration over anticommuting variables was introduced in

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

EUCLIDEAN PATH INTEGRALS IN QUANTUM MECHANICS

In most of this work, we study Quantum Mechanics and Quantum Field Theory in a euclidean formulation. This means that, in general, we discuss matrix elements of the quantum statistical operator $e^{-\beta H}$ (the density matrix at thermal equilibrium), where H is the hamiltonian and β the inverse temperature, rather than those of the quantum evolution operator $e^{-iHt/\hbar}$.

The quantum statistical operator, whose trace is the quantum partition function $\mathcal{Z}(\beta)$,

$$\mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta H}, \tag{2.1}$$

describes "evolution" in imaginary time, and in this sense most of the algebraic properties which will be derived, also apply to the real time evolution operator, explicit expressions being obtained by analytic continuation $\beta \mapsto it/\hbar$. Therefore, in this chapter, to keep track of the $\hbar$ factors of real-time evolution, we first calculate $e^{-tH/\hbar}$.

Our basic tools to study first Quantum Mechanics and then Quantum Field Theory are *path integrals* and *functional integrals*. We shall see that the path integral formulation of quantum mechanics is well suited to the study of systems with an arbitrary number of degrees of freedom. It, therefore, allows a smooth transition between quantum mechanics and quantum field theory.

Another feature will play an essential role: the euclidean path and functional integral formulation emphasizes the deep connection between Quantum Field Theory and the Statistical Physics of critical systems and phase transitions.

Let us also mention a specific property. The operator $e^{-\beta H}$ provides a tool to determine the structure of the quantum ground state. For example, if H is bounded from below, the ground state energy E_0 is given by

$$E_0 = \lim_{\beta\to\infty}\left(-\frac{1}{\beta}\ln \operatorname{tr} e^{-\beta H}\right). \tag{2.2}$$

If the ground state is in addition unique and isolated, $e^{-\beta H}$ projects, for β large, onto the ground state vector $|0\rangle$:

$$e^{-\beta H} \underset{\beta\to\infty}{\sim} e^{-\beta E_0}\,|0\rangle\,\langle 0|\,. \tag{2.3}$$

Therefore, the euclidean functional integral often leads to a simple and intuitive understanding of the structure of the ground state of systems with an infinite number of degrees of freedom. In particular, it gives a natural interpretation to barrier penetration effects in the semi-classical approximation.

Finally, it is generally easier to define properly 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 quantum mechanics 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 we first derive the path integral representation of the matrix elements of the quantum statistical operator for hamiltonians of the simple form $p^2/2m + V(q)$.

EUCLIDEAN PATH INTEGRALS IN QUANTUM MECHANICS

In most of this work, we study Quantum Mechanics and Quantum Field Theory in a euclidean formulation. This means that, in general, we discuss matrix elements of the quantum statistical operator $e^{-\beta H}$ (the density matrix at thermal equilibrium), where H is the hamiltonian and β the inverse temperature, rather than those of the quantum evolution operator $e^{-iHt/\hbar}$.

The quantum statistical operator, whose trace is the quantum partition function $\mathcal{Z}(\beta)$,

$$\mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta H}, \tag{2.1}$$

describes "evolution" in imaginary time, and in this sense most of the algebraic properties which will be derived, also apply to the real time evolution operator, explicit expressions being obtained by analytic continuation $\beta \mapsto it/\hbar$. Therefore, in this chapter, to keep track of the $\hbar$ factors of real-time evolution, we first calculate $e^{-tH/\hbar}$.

Our basic tools to study first Quantum Mechanics and then Quantum Field Theory are *path integrals* and *functional integrals*. We shall see that the path integral formulation of quantum mechanics is well suited to the study of systems with an arbitrary number of degrees of freedom. It, therefore, allows a smooth transition between quantum mechanics and quantum field theory.

Another feature will play an essential role: the euclidean path and functional integral formulation emphasizes the deep connection between Quantum Field Theory and the Statistical Physics of critical systems and phase transitions.

Let us also mention a specific property. The operator $e^{-\beta H}$ provides a tool to determine the structure of the quantum ground state. For example, if H is bounded from below, the ground state energy E_0 is given by

$$E_0 = \lim_{\beta\to\infty}\left(-\frac{1}{\beta}\ln \operatorname{tr} e^{-\beta H}\right). \tag{2.2}$$

If the ground state is in addition unique and isolated, $e^{-\beta H}$ projects, for β large, onto the ground state vector $|0\rangle$:

$$e^{-\beta H} \underset{\beta\to\infty}{\sim} e^{-\beta E_0}\,|0\rangle\,\langle 0|\,. \tag{2.3}$$

Therefore, the euclidean functional integral often leads to a simple and intuitive understanding of the structure of the ground state of systems with an infinite number of degrees of freedom. In particular, it gives a natural interpretation to barrier penetration effects in the semi-classical approximation.

Finally, it is generally easier to define properly 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 quantum mechanics 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 we first derive the path integral representation of the matrix elements of the quantum statistical operator for hamiltonians of the simple form $p^2/2m + V(q)$.

Comparing classical statistical physics in one space dimension and quantum statistical physics of the particle, we introduce statistical correlation functions and discuss their quantum interpretation. We then explicitly calculate the path integral corresponding to a harmonic oscillator in a time-dependent external force. This result can be used to reduce the evaluation of path integrals in the case of analytic potentials to perturbation theory.

We show on a first example that path integrals are especially well suited to the study of the classical limit, by relating a quantum and classical partition function. In the appendix, we explain some general properties of the two-point function, and use the semi-classical approximation of the partition function to derive Bohr–Sommerfeld's quantization condition.

2.1 Path Integrals: The General Idea

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

$$U(t,t'')U(t'',t') = U(t,t'), \qquad t \geq t'' \geq t' . \tag{2.4}$$

This is a property also characteristic of the kind of stochastic processes we examine in Chapter 4, 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 previous evolution.

In the limit $t'' = t'$ we find

$$U(t',t') = \mathbf{1} .$$

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

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

Differentiating equation (2.4) with respect to t we obtain in the limit $t'' = t$:

$$\hbar\frac{\partial U}{\partial t}(t,t') = -H(t)U(t,t'), \qquad U(t',t') = \mathbf{1} . \tag{2.5}$$

When the operator H is time-independent it is the generator of time translations and the equation has the formal solution $U(t,t') = \mathrm{e}^{-(t-t')H/\hbar}$.

The semi-group property (2.4) allows to write the operator $U(t'',t')$ as a product of operators corresponding to arbitrarily small time intervals $\varepsilon = (t'' - t')/n$:

$$U(t'',t') = \prod_{m=1}^{n} U[t' + m\varepsilon, t' + (m-1)\varepsilon], \qquad n\varepsilon = t'' - t' , \tag{2.6}$$

where the product is time-ordered according to the rule (2.4).

Note that the operator $U(t+\varepsilon,t)$ is the analogue of the transfer matrix in classical statistical physics (see Section 2.4).

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

In terms of matrix elements, equation (2.6) becomes

$$\langle q''|U(t'',t')|q'\rangle = \int \prod_{k=1}^{n-1} \mathrm{d}q_k \prod_{k=1}^{n} \langle q_k|U(t_k,t_{k-1})|q_{k-1}\rangle\,, \tag{2.7}$$

with the conventions

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

In this expression, we can take the limit $n \to \infty$ 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 $\varepsilon \to 0$.

Locality of short time evolution. If the *hamiltonian H is local in the basis in which the position operator $\hat{q}$ is diagonal*, which means that its matrix elements $\langle q_1|H(t)|q_2\rangle$ have a support restricted to $q_1 = q_2$, then for $\varepsilon \to 0$ only the matrix elements of U with $|q-q'|$ small contribute significantly to expression (2.7), and we can construct a **path integral** representation of $\langle q''|U(t'',t')|q'\rangle$.

This property holds for operators $H(t)$ that can be expressed in terms of $\hat{p}$ and $\hat{q}$, the momentum and position operators of quantum mechanics, $H(t) \equiv H(\hat{p},\hat{q};t)$, and that are polynomials in $\hat{p}$.

The operator H. Up to now the arguments are rather general. We can now specialize the operator H. If we take H anti-hermitian, $H = i\tilde{H}$, where $\tilde{H}$ has the form of a quantum hamiltonian, then U is unitary and has the form of a quantum evolution operator. In this chapter instead, we choose H hermitian. If H is time-independent and if we set $t - t' = \hbar\beta$ we find for U the quantum statistical operator or matrix density at thermal equilibrium at temperature $T = 1/\beta$. We nevertheless call the variable t time (or euclidean time), though 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. The same continuation allows also to transpose the algebraic part of the calculations that follow to real-time evolution.

Finally note that in Chapter 4 other operators H will appear that, in general, are neither hermitian nor anti-hermitian.

2.2 Path Integral Representation: Special Hamiltonians

In this chapter, we evaluate the expression (2.7) for the special class of hamiltonians of the separable form

$$H = \hat{\mathbf{p}}^2/2m + V(\hat{\mathbf{q}},t), \qquad [\hat{q}_\alpha,\hat{p}_\beta] = i\hbar\delta_{\alpha\beta}\,, \tag{2.8}$$

($\mathbf{p},\mathbf{q}$ now being d-dimensional vectors). More general hamiltonians will be discussed starting with Chapter 3.

In the $|q\rangle$ basis equation (2.5), expressed in terms of the matrix elements $\langle\mathbf{q}|U|\mathbf{q}'\rangle$ then takes the form of a Schrödinger equation in imaginary time

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

with the boundary condition:

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

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

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

To solve equation (2.9) in the small $\varepsilon = t - t'$ limit, it is convenient to set

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

The Schrödinger equation becomes

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

We expect the function σ to be dominated by the free contribution. We thus expand

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

where $\sigma_1 = O(\varepsilon)$. Neglecting terms of order ε^2, we obtain the equation

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

We introduce the linear trajectory $\mathbf{q}(\tau)$ which goes from $\mathbf{q}'$ to $\mathbf{q}$ at constant velocity,

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

and note that it is a solution of the homogeneous form 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

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

Moreover, the free contribution can also be expressed in terms of the trajectory (2.13) ($\dot{q} \equiv \mathrm{d}q/\mathrm{d}t$)

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

This leads to

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

We verify that the normalization is such that the limit for $\varepsilon \to 0$ is indeed $\delta(\mathbf{q}-\mathbf{q}')$.

From equation (2.7), we thus derive

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

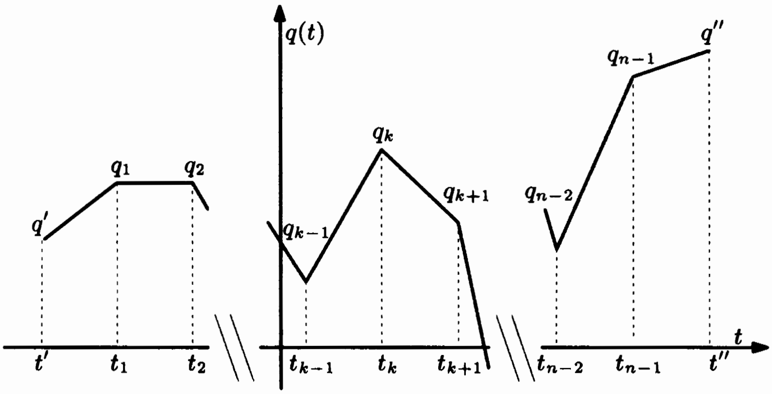


Fig. 2.1 A path contributing to the integral (2.16).

with

$$\mathcal{S}(\mathbf{q},\varepsilon)=\int_{t'}^{t''}\mathrm{d}t\left[\tfrac{1}{2}m\dot{\mathbf{q}}^2(t)+V\bigl(\mathbf{q}(t),t\bigr)\right]+O(n\varepsilon^2), \tag{2.17}$$

where $\mathbf{q}(t)$ now is the piece-wise linear continuous trajectory defined by (figure 2.1)

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

In terms of this 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 $\mathbf{q}(t)$.

We then observe that higher orders in ε in (2.17) give vanishing contributions in the small ε, $n\varepsilon$ fixed, limit. The function $\mathcal{S}(\mathbf{q},\varepsilon)$ has the limit

$$\lim_{\varepsilon\to0}\mathcal{S}(\mathbf{q},\varepsilon)=\mathcal{S}(\mathbf{q})\equiv\int_{t'}^{t''}\mathrm{d}t\left[\tfrac{1}{2}m\dot{\mathbf{q}}^2(t)+V\bigl(\mathbf{q}(t),t\bigr)\right]. \tag{2.18}$$

The functional $\mathcal{S}(\mathbf{q})$ is the *euclidean action*, the integral of the "euclidean lagrangian" associated with the original hamiltonian.

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

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

The r.h.s. is called *path integral* since the integral involves a summation over all paths satisfying the prescribed boundary conditions, with a weight $\exp\left[-\mathcal{S}/\hbar\right]$.

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

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

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

Therefore, in explicit calculations we shall always normalize the result by dividing the path integral by a reference path integral for which the result is exactly known (the free motion $V \equiv 0$, for example).

Fundamental remarks. The most singular term in σ (equation (2.11)) for $\varepsilon \to 0$ 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 $|\mathbf{q}'-\mathbf{q}| = O(\sqrt{\varepsilon})$,

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

if the potential is differentiable. Hence, replacing in the expression (2.14) σ_1, for instance, by $\varepsilon V(\mathbf{q},t)$ modifies σ by terms of order $\varepsilon^{3/2}$ which are negligible. More generally, for the three terms to be equivalent, the potential has to be at least continuous, other potentials require a special analysis.

This singular term also implies that in the action (2.18) the two terms play quite different roles. The potential weights paths according to the value of $\mathbf{q}(t)$ at each time, and determines the physical properties of the theory. The kinetic term $\int \mathrm{d}t\, \dot{\mathbf{q}}^2$ instead selects paths regular enough, those for which $\left[\mathbf{q}(t+\varepsilon)-\mathbf{q}(t)\right]^2/\varepsilon$ remains finite when ε goes to zero. More precisely one finds that the expectation value of $[\mathbf{q}(t+\varepsilon)-\mathbf{q}(t)]^2$ is proportional to $|\varepsilon|\hbar/m$ for $\varepsilon \to 0$ (see Section A2.1 for a proof in the continuum limit). The kinetic term really is part of the functional measure and determines the functional space over which to integrate. It is essential to the very existence of the path integral.

The relevant paths are typical of the brownian motion (for details see Chapter 4); 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 it immediately shows that the paths which give the largest contributions to the path integral (2.19) are in the neighbourhood of the paths which minimize the action (2.18),

$$\frac{\delta \mathcal{S}}{\delta q_i(t)} = 0\,, \quad \text{and in the sense of operators} \quad \frac{\delta^2 \mathcal{S}}{\delta q_i(t_1)\delta q_j(t_2)} \ge 0\,,$$

that is, classical (differentiable) euclidean paths. This observation is at the basis of semi-classical approximations.

Many particle systems. The generalization of the previous construction to several particles is straightforward and leads to a path integral involving the corresponding euclidean action.

2.3 Explicit Evaluation of a Path Integral: The Harmonic Oscillator

At this point in the discussion it may not be very clear whether the concept of path integral is really useful. In particular, it is important to understand to what extent path integrals can be calculated in the continuum, without returning to the definition as a limit of integrals involving discrete time intervals. Thus we work out explicitly a simple example: we calculate a gaussian path integral associated with the hamiltonian of the harmonic oscillator.

If we write the hamiltonian of the one-dimensional harmonic oscillator

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2\,, \tag{2.21}$$

(ω constant), the corresponding euclidean action $\mathcal{S}_0(q)$ is

$$\mathcal{S}_0(q) = \int_{t'}^{t''} \left[\tfrac{1}{2}m\dot{q}^2(t) + \tfrac{1}{2}m\omega^2 q^2(t)\right] \mathrm{d}t, \tag{2.22}$$

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

$$\langle q''|\, U_0(t'',t')\, |q'\rangle = \int_{q(t')=q'}^{q(t'')=q''} [\mathrm{d}q(t)] \exp\left[-\mathcal{S}_0(q)/\hbar\right]. \tag{2.23}$$

The simplest part of the calculation is the determination of the dependence on the boundary conditions, that is, on q', q''. We change variables $q(t) \mapsto r(t)$,

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

in which q_c is the solution of the classical equation of motion:

$$m\ddot{q}_c(t) - m\omega^2 q_c(t) = 0 \quad \text{with} \quad q_c(t') = q'\,,\ q_c(t'') = q''. \tag{2.24}$$

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

$$r(t') = r(t'') = 0\,. \tag{2.25}$$

The action becomes

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

Integrating by parts, $\int \dot{q}_c\dot{r} = r\dot{q}_c - \int r\ddot{q}_c$, taking into account equation (2.24) and the boundary conditions (2.25), we see that the terms linear in r cancel and the action (2.22) reduces to:

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

Then, the remaining integral over $r(t)$ no longer depends on q', q'', and yields a normalization factor, function only of ω and $t'' - t'$.

The classical action. We now calculate $\mathcal{S}_0(q_c)$ explicitly. Setting $\tau = t'' - t'$, we find

$$q_c(t) = \frac{1}{\sinh(\omega\tau)} \left[q' \sinh\left(\omega(t''-t)\right) + q'' \sinh\left(\omega(t-t')\right)\right]. \tag{2.26}$$

To calculate $\mathcal{S}_0(q_c)$ one can again integrate by parts, $\int \dot{q}^2\, \mathrm{d}t = q\dot{q} - \int q\ddot{q}\, \mathrm{d}t$, and use the equation of motion (2.24). The result is

$$\mathcal{S}_0(q_c) = \frac{m\omega}{2\sinh\omega\tau} \left[\left(q'^2 + q''^2\right)\cosh\omega\tau - 2q'q''\right]. \tag{2.27}$$

The path integral. We thus obtain

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

where

$$\mathcal{N}(\omega,\tau) = \int [\mathrm{d}r(t)] \exp\left[-\frac{1}{\hbar}\int_{t'}^{t''} \mathrm{d}t \left(\frac{1}{2}m\dot{r}^2(t) + \frac{1}{2}m\omega^2 r^2(t)\right)\right], \tag{2.29}$$

with $r(t') = r(t'') = 0$. To complete the calculation, one still has to evaluate this last gaussian integral over $r(t)$. Since this involves an infinite normalization factor we postpone the calculation until Section 2.5.2. The final result is

$$\langle q''|\, U_0(t'',t)\, |q'\rangle = \left(\frac{m\omega}{2\pi\hbar\sinh\omega\tau}\right)^{1/2} \times \exp\left\{-\frac{m\omega}{2\hbar\sinh\omega\tau}\left[\left(q'^2+q''^2\right)\cosh\omega\tau - 2q'q''\right]\right\}. \tag{2.30}$$

2.4 Partition Function. Correlation Functions

In this section, the potentials will be assumed *time-independent* and with discrete spectrum.

The quantum partition function. The quantum partition function $\mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta H}$ has a path integral representation which is immediately deduced from the representation of the density matrix

$$\mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta H} \equiv \operatorname{tr} U(\hbar\beta, 0) = \int \mathrm{d}q \, \langle q| U(\hbar\beta, 0) |q\rangle = \int [\mathrm{d}q(t)] \exp\left[-\mathcal{S}(q)/\hbar\right], \quad (2.31)$$

where the paths now satisfy periodic boundary conditions: $q(0) = q(\hbar\beta)$, and one integrates over all values of $q(0)$. It is actually convenient to rescale time $t \mapsto t/\hbar$. The action then reads

$$\mathcal{S}(q)/\hbar = \int_0^\beta \mathrm{d}t \left[\tfrac{1}{2} m \dot{\mathbf{q}}^2(t)/\hbar^2 + V\big(\mathbf{q}(t)\big)\right]. \quad (2.32)$$

The harmonic oscillator. Equation (2.28) allows to relate the normalization of the path integral to the partition function $\mathcal{Z}_0(\beta)$ of the harmonic oscillator. Taking the trace of $U_0(\hbar\beta, 0)$ one finds

$$\begin{aligned} \mathcal{Z}_0(\beta) &= \operatorname{tr} U_0(\hbar\beta, 0) \equiv \int \mathrm{d}q \, \langle q| U_0(\hbar\beta, 0) |q\rangle \\ &= \mathcal{N}(\omega, \beta) \left(\frac{\pi\hbar}{m\omega \tanh(\beta\hbar\omega/2)}\right)^{1/2}. \end{aligned} \quad (2.33)$$

The large β limit. One can also set the boundary conditions at $t = \pm\beta/2$. 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 is led to integrate over paths on the positive real line with the boundary condition $q(0) = 0$, while in the second case one obtains an explicitly time translation invariant formalism on the whole real line. The latter formalism is clearly simpler.

2.4.1 Classical and quantum statistical physics

The time-discretized form of the path integral has an interpretation in terms of a classical statistical model on a one-dimensional lattice.

Let us consider the quantity

$$\mathcal{Z}(n, \varepsilon) = \int \prod_{k=1}^{n} \mathrm{d}q_k \, \exp\left[-\mathcal{S}(q, \varepsilon)\right] \quad (2.34)$$

with

$$\mathcal{S}(q, \varepsilon) = \sum_{k=1}^{n} \left[\frac{(q_k - q_{k-1})^2}{2\varepsilon} + \varepsilon V(q_k)\right],$$

and $q_0 = q_n$. It can also be considered as the classical partition function of a one-dimensional lattice model. The variable q_k characterizes the configuration on the site k of a 1D lattice, n is the size of the lattice, ε plays, somewhat, the role of the temperature, and $V(q)$ determines the distribution of q on each site. Finally, since $q_n = q_0$ the model satisfies periodic boundary conditions.

Introducing the function

$$S(q,q') = \frac{(q-q')^2}{2\varepsilon} + \frac{1}{2}\varepsilon V(q) + \frac{1}{2}\varepsilon V(q'), \tag{2.35}$$

we can write

$$\mathcal{S}(q,\varepsilon) = \sum_{k=1}^{n} S(q_{k-1}, q_k).$$

The partition function (2.34) can then be expressed in terms of the *transfer matrix* $\mathbf{T}$ which is represented by the kernel $\mathcal{T}(q,q')$:

$$\mathcal{Z}(n,\varepsilon) = \operatorname{tr} \mathbf{T}^n$$

with

$$\mathcal{T}(q,q') \equiv \langle q' | \mathbf{T} | q \rangle = \mathrm{e}^{-S(q,q')} . \tag{2.36}$$

In one dimension, one expects a non-trivial collective behaviour only at low temperature, where $\varepsilon \to 0$. Then, one finds (with $\hbar = 1$)

$$\mathcal{T}(q,q') \underset{\varepsilon\to 0}{\sim} \sqrt{2\pi\varepsilon}\, \langle q' | \mathrm{e}^{-\varepsilon H} | q \rangle , \quad H = \tfrac{1}{2}\hat{p}^2 + V(\hat{q}).$$

For ε small, the eigenvalues and eigenvectors of the transfer matrix are simply related to those of H. In the thermodynamic limit $n \to \infty$, the transfer matrix is dominated by its largest eigenvalue which corresponds to the ground state energy E_0 of H. The free energy $\mathcal{W} = \ln \mathcal{Z}$ per unit "volume" 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, according to the preceding analysis, when $\varepsilon \to 0$ with $n\varepsilon = \beta$ fixed, $\mathcal{Z}(n,\varepsilon)$ becomes (up to an irrelevant factor) the quantum partition function $\mathcal{Z}(\beta)$

$$(2\pi\varepsilon)^{n/2} \mathcal{Z}(n,\varepsilon) \to \mathcal{Z}(\beta) = \operatorname{tr} \mathrm{e}^{-\beta H} .$$

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

2.4.2 Correlation functions

Correlation functions in the classical statistical model are given by

$$\langle q_{i_1} q_{i_2} \ldots q_{i_m} \rangle = \mathcal{Z}^{-1}(n,\varepsilon) \int \left(\prod_{k=1}^{n} \mathrm{d}q_k \right) q_{i_1} q_{i_2} \ldots q_{i_m} \exp\left[-\mathcal{S}(q,\varepsilon)\right] . \tag{2.37}$$

and $V(q)$ determines the distribution of q on each site. Finally, since $q_n = q_0$ the model satisfies periodic boundary conditions.

Introducing the function

$$S(q,q') = \frac{(q-q')^2}{2\varepsilon} + \frac{1}{2}\varepsilon V(q) + \frac{1}{2}\varepsilon V(q'), \tag{2.35}$$

we can write

$$\mathcal{S}(q,\varepsilon) = \sum_{k=1}^{n} S(q_{k-1}, q_k).$$

The partition function (2.34) can then be expressed in terms of the *transfer matrix* $\mathbf{T}$ which is represented by the kernel $\mathcal{T}(q,q')$:

$$\mathcal{Z}(n,\varepsilon) = \operatorname{tr} \mathbf{T}^n$$

with

$$\mathcal{T}(q,q') \equiv \langle q' | \mathbf{T} | q \rangle = \mathrm{e}^{-S(q,q')} . \tag{2.36}$$

In one dimension, one expects a non-trivial collective behaviour only at low temperature, where $\varepsilon \to 0$. Then, one finds (with $\hbar = 1$)

$$\mathcal{T}(q,q') \underset{\varepsilon\to 0}{\sim} \sqrt{2\pi\varepsilon}\, \langle q' | \mathrm{e}^{-\varepsilon H} | q \rangle , \quad H = \tfrac{1}{2}\hat{p}^2 + V(\hat{q}).$$

For ε small, the eigenvalues and eigenvectors of the transfer matrix are simply related to those of H. In the thermodynamic limit $n \to \infty$, the transfer matrix is dominated by its largest eigenvalue which corresponds to the ground state energy E_0 of H. The free energy $\mathcal{W} = \ln \mathcal{Z}$ per unit "volume" 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, according to the preceding analysis, when $\varepsilon \to 0$ with $n\varepsilon = \beta$ fixed, $\mathcal{Z}(n,\varepsilon)$ becomes (up to an irrelevant factor) the quantum partition function $\mathcal{Z}(\beta)$

$$(2\pi\varepsilon)^{n/2} \mathcal{Z}(n,\varepsilon) \to \mathcal{Z}(\beta) = \operatorname{tr} \mathrm{e}^{-\beta H} .$$

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

2.4.2 Correlation functions

Correlation functions in the classical statistical model are given by

$$\langle q_{i_1} q_{i_2} \ldots q_{i_m} \rangle = \mathcal{Z}^{-1}(n,\varepsilon) \int \left(\prod_{k=1}^{n} \mathrm{d}q_k \right) q_{i_1} q_{i_2} \ldots q_{i_m} \exp\left[-\mathcal{S}(q,\varepsilon)\right] . \tag{2.37}$$

If we assume the order $0 < i_1 \le i_2 \le \cdots \le i_m < n$ we can express the m-point correlation function in terms of the transfer matrix (2.36). We decompose $\mathcal{S}(q,\varepsilon)$ into

$$\mathcal{S}(q,\varepsilon) = \sum_{s=0}^{m} \sum_{k=i_s+1}^{i_s+1} S(q_{k-1}, q_k),$$

with $i_0 = 0$, $i_{m+1} = n$. We integrate over all q variables except $q_{i_1}, q_{i_2}, \ldots, q_{i_m}$ and find

$$\langle q_{i_1} q_{i_2} \ldots q_{i_m} \rangle = \mathcal{Z}^{-1}(n,\varepsilon) \int \prod_{s=1}^{m} \mathrm{d}q_{i_s} \, \langle q_{i_1} | \mathbf{T}^{n-i_m+i_1} | q_{i_m} \rangle \, q_{i_m}$$
$$\times \langle q_{i_m} | \mathbf{T}^{i_m - i_{m-1}} | q_{i_{m-1}} \rangle \, q_{i_{m-1}} \ldots \langle q_{i_2} | \mathbf{T}^{i_2 - i_1} | q_{i_1} \rangle \, q_{i_1} .$$

We then use the property that the position operator $\hat{q}$ is diagonal in the $|q\rangle$ basis

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

and obtain

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

For ε small we can introduce the hamiltonian H. It is also convenient to change notation, associating discrete values of a continuous variable t to site positions

$$t_k = \varepsilon i_k .$$

One then finds

$$Z^{(m)}(t_1, t_2, \ldots, t_m) \equiv \langle q(t_1) q(t_2) \ldots q(t_m) \rangle_\beta$$
$$\sim \mathcal{Z}^{-1}(\beta) \operatorname{tr} \left[\mathrm{e}^{-(\beta - t_n + t_1)H} \, \hat{q} \, \mathrm{e}^{-(t_m - t_{m-1})H} \, \hat{q} \ldots \mathrm{e}^{-(t_2 - t_1)H} \, \hat{q} \right] . \quad (2.38)$$

It also follows from the direct definition (2.37) that correlation functions have, for $\varepsilon \to 0$, $n\varepsilon = \beta$, $\varepsilon i_k = t_k$ fixed, a path integral representation

$$\langle q(t_1) q(t_2) \ldots q(t_m) \rangle_\beta = \mathcal{Z}^{-1}(\beta) \int [\mathrm{d}q] \, q(t_1) \ldots q(t_m) \, \mathrm{e}^{-\mathcal{S}(q)} . \quad (2.39)$$

Notice 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. This does not prevent, however, 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 continuation of the classical correlation functions $t_k \mapsto it_k$ are the dynamic correlation functions of the quantum theory.

Thermodynamic limit. In the classical statistical model the limit $\beta \to \infty$ (zero temperature of the quantum model) is the thermodynamic limit. Assuming that the ground state is unique (this is always true in quantum mechanics) we find

$$\langle q(t_1) q(t_2) \ldots q(t_m) \rangle = \langle 0 | \hat{q} \, \mathrm{e}^{-(t_m - t_{m-1})(H - E_0)} \, \hat{q} \ldots \mathrm{e}^{-(t_2 - t_1)(H - E_0)} \, \hat{q} \, | 0 \rangle . \quad (2.40)$$

If we assume the order $0 < i_1 \le i_2 \le \cdots \le i_m < n$ we can express the m-point correlation function in terms of the transfer matrix (2.36). We decompose $\mathcal{S}(q,\varepsilon)$ into

$$\mathcal{S}(q,\varepsilon) = \sum_{s=0}^{m} \sum_{k=i_s+1}^{i_s+1} S(q_{k-1}, q_k),$$

with $i_0 = 0$, $i_{m+1} = n$. We integrate over all q variables except $q_{i_1}, q_{i_2}, \ldots, q_{i_m}$ and find

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

We then use the property that the position operator $\hat{q}$ is diagonal in the $|q\rangle$ basis

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

and obtain

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

For ε small we can introduce the hamiltonian H. It is also convenient to change notation, associating discrete values of a continuous variable t to site positions

$$t_k = \varepsilon i_k \, .$$

One then finds

$$\begin{aligned} Z^{(m)}(t_1, t_2, \ldots, t_m) &\equiv \langle q(t_1) q(t_2) \ldots q(t_m) \rangle_\beta \\ &\sim \mathcal{Z}^{-1}(\beta) \operatorname{tr} \left[\mathrm{e}^{-(\beta - t_n + t_1)H} \, \hat{q} \, \mathrm{e}^{-(t_m - t_{m-1})H} \, \hat{q} \ldots \mathrm{e}^{-(t_2 - t_1)H} \, \hat{q} \right] . \end{aligned} \tag{2.38}$$

It also follows from the direct definition (2.37) that correlation functions have, for $\varepsilon \to 0$, $n\varepsilon = \beta$, $\varepsilon i_k = t_k$ fixed, a path integral representation

$$\langle q(t_1) q(t_2) \ldots q(t_m) \rangle_\beta = \mathcal{Z}^{-1}(\beta) \int [\mathrm{d}q] \, q(t_1) \ldots q(t_m) \, \mathrm{e}^{-\mathcal{S}(q)} \, . \tag{2.39}$$

Notice 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. This does not prevent, however, 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 continuation of the classical correlation functions $t_k \mapsto it_k$ are the dynamic correlation functions of the quantum theory.

Thermodynamic limit. In the classical statistical model the limit $\beta \to \infty$ (zero temperature of the quantum model) is the thermodynamic limit. Assuming that the ground state is unique (this is always true in quantum mechanics) we find

$$\langle q(t_1) q(t_2) \ldots q(t_m) \rangle = \langle 0 | \, \hat{q} \, \mathrm{e}^{-(t_m - t_{m-1})(H - E_0)} \, \hat{q} \ldots \mathrm{e}^{-(t_2 - t_1)(H - E_0)} \, \hat{q} \, | 0 \rangle \, . \tag{2.40}$$

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

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

The expectation value of $q(t)$ has no special meaning here and can always be eliminated by the constant shift $q(t) \mapsto q(t) - \langle q(t)\rangle$.

The two-point function is given by

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

Calling E_k, $E_0 < E_1 < E_2 \ldots$ the successive eigenvalues of H and $|k\rangle$ the corresponding eigenvectors, we find for large separations $|t_2 - t_1| \to \infty$ (with $\langle q\rangle = 0$),

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

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

$$\langle q_{i_1} q_{i_2}\rangle \propto \mathrm{e}^{-\varepsilon(E_1-E_0)|i_2-i_1|}\,.$$

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

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

We conclude that in the continuum limit $\varepsilon \to 0$, the correlation length, in lattice units, diverges. Keeping the variables $t_1, t_2, \ldots, \beta$ fixed when $\varepsilon \to 0$, 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 the details of the time discretization. This is the first example of a situation we shall again encounter in the context of second order phase transitions.

Remark. We have assumed the ground state to be unique. The existence of phase transitions, as will be discussed in Chapter 23, is related to a possible ground state degeneracy.

2.5 Generating Functional of Correlation Functions. Perturbative Expansion

We first calculate the path integral of the harmonic oscillator coupled to an external force, and with *periodic boundary conditions*. This quantity is also the generating functional of statistical correlation functions of the gaussian model of Section 2.3. We then use it to generate perturbative expansions of more general partition functions.

2.5.1 Harmonic oscillator coupled to an external force

We consider the hamiltonian

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 - qb(t), \tag{2.41}$$

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

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

The expectation value of $q(t)$ has no special meaning here and can always be eliminated by the constant shift $q(t) \mapsto q(t) - \langle q(t)\rangle$.

The two-point function is given by

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

Calling E_k, $E_0 < E_1 < E_2 \ldots$ the successive eigenvalues of H and $|k\rangle$ the corresponding eigenvectors, we find for large separations $|t_2 - t_1| \to \infty$ (with $\langle q\rangle = 0$),

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

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

$$\langle q_{i_1} q_{i_2}\rangle \propto \mathrm{e}^{-\varepsilon(E_1-E_0)|i_2-i_1|}\,.$$

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

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

We conclude that in the continuum limit $\varepsilon \to 0$, the correlation length, in lattice units, diverges. Keeping the variables $t_1, t_2, \ldots, \beta$ fixed when $\varepsilon \to 0$, 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 the details of the time discretization. This is the first example of a situation we shall again encounter in the context of second order phase transitions.

Remark. We have assumed the ground state to be unique. The existence of phase transitions, as will be discussed in Chapter 23, is related to a possible ground state degeneracy.

2.5 Generating Functional of Correlation Functions. Perturbative Expansion

We first calculate the path integral of the harmonic oscillator coupled to an external force, and with *periodic boundary conditions*. This quantity is also the generating functional of statistical correlation functions of the gaussian model of Section 2.3. We then use it to generate perturbative expansions of more general partition functions.

2.5.1 Harmonic oscillator coupled to an external force

We consider the hamiltonian

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2 - qb(t), \tag{2.41}$$

where $b(t)$ is a fixed arbitrary function. The corresponding action is

$$\mathcal{S}_{\mathrm{G}}(q,b)=\int_{-\tau/2}^{\tau/2}\mathrm{d}t\left[\tfrac{1}{2}m\dot{q}^2(t)+\tfrac{1}{2}m\omega^2q^2(t)-b(t)q(t)\right]. \tag{2.42}$$

To obtain the partition function, we have to calculate the more general gaussian path integral

$$\operatorname{tr}U_{\mathrm{G}}(\tau/2,-\tau/2)=\int_{q(\tau/2)=q(-\tau/2)}[\mathrm{d}q(t)]\exp[-\mathcal{S}_{\mathrm{G}}(q,b)/\hbar]. \tag{2.43}$$

To eliminate the linear term in the action, we follow the strategy explained in Section 1.1. We change variables $q(t)\mapsto r(t)$:

$$q(t)=q_c(t)+r(t),\quad q_c(\tau/2)=q_c(-\tau/2)\ \Rightarrow\ r(\tau/2)=r(-\tau/2). \tag{2.44}$$

Then,

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

In the term linear in r, we integrate by parts and use the periodic boundary conditions (2.44)

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

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

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

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

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

where the function $\Delta(t)$ is the solution of the equation

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

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

$$\Delta(t)=\frac{1}{2\omega\,\sinh(\omega\tau/2)}\cosh\bigl(\omega(\tau/2-|t|)\bigr). \tag{2.45}$$

In the limit $\tau\to\infty$, we find

$$\Delta(t)=\frac{1}{2\omega}\mathrm{e}^{-\omega|t|}. \tag{2.46}$$

We then obtain for $\mathcal{S}_{\mathrm{G}}(q_c, b)$:

$$\begin{aligned}\mathcal{S}_{\mathrm{G}}(q_c, b) &= \int_{-\tau/2}^{\tau/2} \mathrm{d}t \left[\tfrac{1}{2}m\dot{q}_c^2(t) + \tfrac{1}{2}m\omega^2 q_c^2(t) - b(t)q_c(t)\right] \\ &= \int_{-\tau/2}^{\tau/2} \mathrm{d}t\, q_c(t) \left[-\tfrac{1}{2}m\ddot{q}_c(t) + \tfrac{1}{2}m\omega^2 q_c(t) - b(t)\right] = -\tfrac{1}{2}\int_{-\tau/2}^{\tau/2} \mathrm{d}t\, q_c(t)b(t) \\ &= -\frac{1}{2m}\int_{-\tau/2}^{\tau/2} \mathrm{d}t\, \mathrm{d}u\, b(t)\Delta(t-u)b(u).\end{aligned}$$

The remaining integral over $r(t)$ just yields $\operatorname{tr} U_0(\tau/2, -\tau/2)$, and, therefore,

$$\operatorname{tr} U_{\mathrm{G}}(\tau/2, -\tau/2) = \operatorname{tr} U_0(\tau/2, -\tau/2)\, \mathrm{e}^{-\mathcal{S}_{\mathrm{G}}(q_c, b)/\hbar}\,.$$

The partition function follows

$$\begin{aligned}\mathcal{Z}_{\mathrm{G}}(b,\beta) &= \operatorname{tr} U_{\mathrm{G}}(\hbar\beta/2, -\hbar\beta/2) \\ &= \mathcal{Z}_0(\beta) \exp\left[\frac{1}{2m\hbar}\int_{-\hbar\beta/2}^{\hbar\beta/2} \mathrm{d}u\, \mathrm{d}v\, \Delta(v-u)b(v)b(u)\right],\end{aligned} \tag{2.47}$$

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

2.5.2 Correlation functions, Wick's theorem

If we differentiate with respect to $b(t)$ we obtain (for a definition of functional differentiation see Section 1.9):

$$\hbar \frac{\delta}{\delta b(t_1)} \mathcal{Z}_{\mathrm{G}}(b,\beta) = \int [\mathrm{d}q]\, q(t_1) \exp\left[-\mathcal{S}_{\mathrm{G}}(q,b)/\hbar\right].$$

Therefore, by differentiating p times with respect to $b(t)$ we can generate any product $q(t_1)\ldots q(t_p)$. Taking the $b \equiv 0$ limit, we obtain correlation functions corresponding to the quadratic action (2.22):

$$\hbar^p \prod_{j=1}^{p} \frac{\delta}{\delta b(t_j)} \mathcal{Z}_{\mathrm{G}}(b,\beta)\bigg|_{b\equiv 0} = \int [\mathrm{d}q] \prod_{j=1}^{p} q(t_j) \exp\left[-\mathcal{S}_0(q)/\hbar\right] \tag{2.48a}$$

$$\equiv \mathcal{Z}_0(\beta) \left\langle q(t_1)q(t_2)\ldots q(t_p)\right\rangle_0, \tag{2.48b}$$

and, therefore, a differential operator $\mathcal{F}(\hbar\delta/\delta b(t))$ generates in the r.h.s. the expectation value of the functional $\mathcal{F}(q)$.

Replacing $\mathcal{Z}_{\mathrm{G}}(b,\beta)$ by the explicit expression (2.47), and differentiating twice we obtain the two-point correlation function

$$\langle q(t)q(u)\rangle_0 = \mathcal{Z}_0^{-1}(\beta)\hbar^2 \frac{\delta^2}{\delta b(t)\delta b(u)} \mathcal{Z}_{\mathrm{G}}(b,\beta)\bigg|_{b\equiv 0} = \frac{\hbar}{m}\Delta(t-u). \tag{2.49}$$

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

$$\langle q(t_1)q(t_2)\ldots q(t_\ell)\rangle_0 = \sum_{\substack{\text{all possible pairings}\\ P \text{ of } \{1,2,\ldots,\ell\}}} \langle q(t_{P_1})q(t_{P_2})\rangle_0 \cdots \left\langle q(t_{P_{\ell-1}})q(t_{P_\ell})\right\rangle_0 . \tag{2.50}$$

Harmonic oscillator: the partition function. We are now in a position to determine the dependence of the partition function $\mathcal{Z}_0(\beta)$ on the parameter ω. Indeed, differentiating the path integral we obtain

$$\frac{\partial}{\partial\omega}\ln\mathcal{Z}_0(\beta)=-\frac{m\omega}{\hbar}\int_{-\hbar\beta/2}^{\hbar\beta/2}\mathrm{d}t\,\left\langle q^2(t)\right\rangle_0=-\frac{\hbar\beta}{2}\frac{\cosh(\omega\hbar\beta/2)}{\sinh(\omega\hbar\beta/2)}. \tag{2.51}$$

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 $\beta\to\infty$, where one should find $\mathrm{e}^{-\beta E_0}$. The complete result, thus, is

$$\mathcal{Z}_0(\beta)=\frac{1}{2\sinh(\beta\hbar\omega/2)}=\frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1-\mathrm{e}^{-\beta\hbar\omega}}, \tag{2.52}$$

where we indeed recognize the partition function of the harmonic oscillator. This also completes the calculation of the normalization in expression (2.30), as a consequence of the relation (2.33).

2.5.3 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 again 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,

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

and $q_0=q_n$.

It is easy to diagonalize the quadratic form in the variables q_k because, due to the periodic boundary conditions, the system is translation invariant. We thus introduce a discrete Fourier representation, setting

$$q_k=\frac{1}{\sqrt{n}}\sum_{\ell=0}^{n-1}\mathrm{e}^{2i\pi k\ell/n}\,c_\ell\,, \tag{2.53}$$

with the reality conditions

$$c_0=\bar{c}_0\,,\quad \bar{c}_{n-\ell}=c_\ell\,. \tag{2.54}$$

Then,

$$\mathcal{S}_0(q)=\sum_{\ell=0}^{n-1}\bar{c}_\ell\left[(1-\cos(2\pi\ell/n))/\varepsilon+\tfrac{1}{2}\omega^2\varepsilon\right]c_\ell\,, \tag{2.55}$$

where the orthogonality relations have been used

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

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 (2.54) implies that only about half of the complex variables are independent. We find

$$\mathcal{Z}_0 = (2\varepsilon)^{-n/2}\left[\prod_{\ell=0}^{n-1}\left(1-\cos(2\pi\ell/n)\right)/\varepsilon + \frac{1}{2}\omega^2\varepsilon\right]^{-1/2}. \tag{2.56}$$

The product can be calculated explicitly. Setting

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

one obtains

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

In the $\varepsilon \to 0$ limit, with $n\varepsilon = \beta$ fixed, $n\theta \to \beta\omega$, and thus

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

We recover the partition function of the harmonic oscillator.

Continuum calculation. The calculation with discrete variables suggests how to perform a continuum calculation. In the continuum limit the change of variables (2.53) becomes an expansion of a function on a basis of normalized periodic *square integrable functions*:

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

The jacobian is unity and the measure simply

$$[\mathrm{d}q(t)] \mapsto \mathrm{d}c_0 \prod_{\ell>0}\mathrm{d}c_\ell\,\mathrm{d}\bar{c}_\ell\,. \tag{2.57}$$

The function $\mathcal{S}_0$ becomes

$$\mathcal{S}_0 = \tfrac{1}{2}\omega^2 c_0^2 + \sum_{\ell\ge1}\bar{c}_\ell(\omega^2 + 4\pi^2\ell^2/\beta^2)c_\ell\,.$$

The integration then is straightforward

$$\mathcal{Z}_0(\beta) \propto \frac{1}{\omega}\prod_{\ell\ge1}\left[(\omega^2 + 4\pi^2\ell^2/\beta^2)\right]^{-1}. \tag{2.58}$$

The infinite product diverges and has to be normalized. The free hamiltonian is not available because the partition function does not exist. It is possible, however, to compare different values of ω, or take the derivative of $\ln\mathcal{Z}_0$

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

which is the result (2.51).

2.5.4 Perturbed harmonic oscillator

We now consider the hamiltonian

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 q^2 + V_{\mathrm{I}}(q), \tag{2.59}$$

where we only assume that $V_{\mathrm{I}}(q)$ is expandable in powers of q:

$$V_{\mathrm{I}}(q) = \sum_n v_n q^n .$$

The corresponding partition function is given by (in this section we set $\hbar = 1$):

$$\mathcal{Z}(\beta) = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q] \exp\left\{-\int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}^2(t) + \tfrac{1}{2}\omega^2 q^2(t) + V_{\mathrm{I}}\big(q(t)\big)\right]\mathrm{d}t\right\}. \tag{2.60}$$

To evaluate the path integral (2.60), we can generalize the identities (1.15,1.17). We apply the identity (2.48a) to the path integral (2.43):

$$\mathcal{Z}(\beta) = \left\{\exp\left[-\int_{-\beta/2}^{\beta/2} \mathrm{d}t\, V_{\mathrm{I}}\left(\frac{\delta}{\delta b(t)}\right)\right] \mathcal{Z}_{\mathrm{G}}(b,\beta)\right\}\bigg|_{b=0} . \tag{2.61}$$

We then replace the partition function $\mathcal{Z}_{\mathrm{G}}(b,\beta)$ by its explicit expression (2.47) calculated in Section 2.5.1

$$\mathcal{Z}(\beta) = \mathcal{Z}_0(\beta) \exp\left[-\int_{-\beta/2}^{\beta/2} \mathrm{d}t\, V_{\mathrm{I}}\left(\frac{\delta}{\delta b(t)}\right)\right] \exp\left[\tfrac{1}{2}\int \mathrm{d}u\,\mathrm{d}v\, b(u)\Delta(u-v)b(v)\right]\bigg|_{b=0} . \tag{2.62}$$

If $V_{\mathrm{I}}(q)$ is a polynomial a more explicit form is obtained by expanding (2.62) in powers of $V_{\mathrm{I}}(q)$: this reduces perturbation theory to gaussian expectation values

$$\mathcal{Z}(\beta)/\mathcal{Z}_0(\beta) = \sum_{k=0} \frac{(-1)^k}{k!} \int \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_k \left\langle V_{\mathrm{I}}\big(q(t_1)\big) \ldots V_{\mathrm{I}}\big(q(t_k)\big)\right\rangle_0 ,$$

where $\langle \bullet \rangle_0$ means 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 (2.50). This is the basis of perturbation theory.

Example:

$$V_{\mathrm{I}}(q) = \lambda q^4 .$$

At second order in λ:

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

Then from Wick's theorem (2.50)

$$\left\langle q^4(t)\right\rangle_0 = 3\left(\left\langle q^2(t)\right\rangle\right)_0^2 = 3\Delta^2(0),$$

and

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

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

$$\begin{aligned}\mathcal{Z}(\beta)/\mathcal{Z}_0(\beta) &= 1 - 3\lambda\beta\Delta^2(0) + \frac{9}{2}\lambda^2\beta^2\Delta^4(0) + 36\beta\lambda^2\Delta^2(0)\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\Delta^2(t) \\ &\quad + 12\lambda^2\beta\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\Delta^4(t) + O\left(\lambda^3\right).\end{aligned} \tag{2.63}$$

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 (2.59). However, if we want the perturbative expansion to be associated with a formal expansion in powers of $\hbar$, then we infer from expression (2.19) that we have to expand the action, and thus the potential around a minimum. Calling q_0 a minimum of the potential we then write

$$V(q) = V(q_0) + \tfrac{1}{2}V''(q_0)(q-q_0)^2 + V_{\mathrm{I}}(q-q_0).$$

The expansion in powers of the coefficients of V_{I} can then be organized as an expansion in powers of $\hbar$, called loopwise expansion. Problems associated with a possible degeneracy of the classical minimum will be examined in Chapter 41.

Correlation functions and perturbation theory. We have shown 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 (2.43). The argument can immediately be generalized to the corresponding correlation functions.

2.6 Semi-Classical Expansion

In the formal limit $\hbar \to 0$ one expects that the quantum partition function approaches 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 (2.31). In particular, since $\hbar$ has a dimension, the expansion parameter must take the form of $\hbar$ divided by an action. The calculation will exhibit the true expansion parameter.

We then use this result to generate semi-classical WKB-like approximations for the spectrum.

2.6.1 Quantum partition function

For $\hbar \to 0$ the leading term in the action (2.32) is the kinetic term, and the dominant contributions come from constant paths. It is thus convenient to calculate first the matrix elements

$$\langle q_0|\,\mathrm{e}^{-\beta H}\,|q_0\rangle = \int_{q(0)=q(\beta)=q_0}[\mathrm{d}q(t)]\exp\left[-\mathcal{S}(q)/\hbar\right]. \tag{2.64}$$

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

$$\langle q_0 | \, e^{-\beta H} \, | q_0 \rangle = \int_{q(0)=q(\beta)=0} [\mathrm{d}q(t)] \exp\left[-\Sigma(q)\right] \tag{2.65}$$

with

$$\Sigma(q) = \int_0^\beta \mathrm{d}t \left[\tfrac{1}{2} m \dot{q}^2(t)/\hbar^2 + V\big(q_0 + q(t)\big)\right]. \tag{2.66}$$

With these new boundary conditions $q(t)$ is of order $\hbar$ and we can thus expand the potential in powers of q,

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

We, then, expand the integrand and calculate the successive terms:

$$\begin{aligned}\langle q_0 | \, e^{-\beta H} \, | q_0 \rangle = \mathcal{N}(\beta) \, e^{-\beta V(q_0)} \Bigg[& 1 - V'(q_0) \int_0^\beta \mathrm{d}t \, \langle q(t) \rangle + \tfrac{1}{2} (V'(q_0))^2 \\ & \times \int_0^\beta \mathrm{d}t \, \mathrm{d}u \, \langle q(t) q(u) \rangle - \tfrac{1}{2} V''(q_0) \int_0^\beta \mathrm{d}t \, \langle q^2(t) \rangle + O(\hbar^3) \Bigg],\end{aligned}$$

where $\langle \bullet \rangle$ means expectation value with the free action ($V = 0$). First $\langle q(t) \rangle = 0$. The two-point function of the free action with the proper boundary conditions is proportional to $\Delta(t, u) = \Delta(u, t)$ solution of

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

It follows

$$\frac{m}{\hbar^2} \langle q(t) q(u) \rangle = \Delta(t, u) = -\frac{1}{2}|t - u| + \frac{1}{2}(t + u - 2ut/\beta), \tag{2.67}$$

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

$$\mathcal{N}(\beta) = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}}. \tag{2.68}$$

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

$$\mathcal{Z}(\beta) = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}} \int \mathrm{d}q \exp\left[-\beta V(q) - \beta^2 \hbar^2 V''(q)/24m + O(\hbar^4)\right]. \tag{2.69}$$

Discussion.

(i) For $\hbar \to 0$ 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 a thermal wavelength

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

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

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

we see 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 temperatures 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 quantum mechanics examples (square well...).

(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. The corresponding classical function is instead given by an integral over the zero-mode, which is just one point, that is, of dimension zero.

2.6.2 WKB Spectrum

We show here how the spectrum in the WKB limit, that is, a limit in which $\hbar \to 0$ 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, as explained in Section 2.6.1.

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

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

The Laplace transform $G(E)$ 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 r.h.s. is obtained by analytic continuation from sufficiently negatives values of the energy variable E.

After continuation to arbitrary complex values of E we obtain

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

This gives us the eigenvalue distribution. For what follows, it is more convenient to consider the integrated distribution

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

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

$$\int_{-\infty}^{E_k} \mathrm{d}E' \, \frac{1}{2i\pi}\big(G(E'+i0) - G(E'-i0)\big) = k + 1/2\,, \tag{2.70}$$

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

We now approximate $\mathcal{Z}(\beta)$ by the leading term of the semi-classical result (2.69)

$$\mathcal{Z}(\beta) = \frac{1}{\hbar}\sqrt{\frac{m}{2\pi\beta}} \int \mathrm{d}q \; \mathrm{e}^{-\beta V(q)}\,.$$

From (2.70), we find

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

We take the discontinuity of $G(E)$ on the cut. We see that the discrete distribution is replaced by a continuous distribution, which 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

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

We thus recover the Bohr–Sommerfeld quantization condition

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

The l.h.s is finite for $\hbar \to 0$: we verify 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.

Bibliographical Notes

The formal analogy between time and inverse temperature has been first noticed in

F. Bloch, *Zeit. f. Phys.* 74 (1932) 295.

For the formulation of quantum mechanics in terms of path integrals, basic papers are

P.A.M. Dirac, *Physik. Z. Sowjetunion* 3 (1933) 64; R.P. Feynman, *Rev. Mod. Phys.* 367 (1948) 20; *Phys. Rev.* 80 (1950) 440.

An expanded version is found in

R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York 1965).

The approach through functional derivatives and the idea of reducing any hamiltonian problem to one in an external source is due to J. Schwinger. See, for example,

J. Schwinger, *Phys. Rev.* 82 (1951) 914; *Proc. Natl. Acad. Sci. USA* 37 (1951) 452.

This chapter has also benefited from

E.S. Abers and B.W. Lee, *Phys. Rep.* 9C (1973) 1.

For a discussion of the use of path integrals in statistical physics see, for example,

J. Ginibre in *Statistical Mechanics and Quantum Field Theory*, Les Houches 1970, C. De Witt and R. Stora eds. (Gordon and Breach, New York 1971).

For the Källen–Lehmann representation, see

G. Källen, *Quantum Electrodynamics* (Springer-Verlag, Berlin 1972).

Bibliographical Notes

The formal analogy between time and inverse temperature has been first noticed in
F. Bloch, *Zeit. f. Phys.* 74 (1932) 295.
For the formulation of quantum mechanics in terms of path integrals, basic papers are
P.A.M. Dirac, *Physik. Z. Sowjetunion* 3 (1933) 64; R.P. Feynman, *Rev. Mod. Phys.* 367 (1948) 20; *Phys. Rev.* 80 (1950) 440.
An expanded version is found in
R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York 1965).
The approach through functional derivatives and the idea of reducing any hamiltonian problem to one in an external source is due to J. Schwinger. See, for example,
J. Schwinger, *Phys. Rev.* 82 (1951) 914; *Proc. Natl. Acad. Sci. USA* 37 (1951) 452.
This chapter has also benefited from
E.S. Abers and B.W. Lee, *Phys. Rep.* 9C (1973) 1.
For a discussion of the use of path integrals in statistical physics see, for example,
J. Ginibre in *Statistical Mechanics and Quantum Field Theory*, Les Houches 1970, C. De Witt and R. Stora eds. (Gordon and Breach, New York 1971).
For the Källen–Lehmann representation, see
G. Källen, *Quantum Electrodynamics* (Springer-Verlag, Berlin 1972).

APPENDIX A2

A2.1 The Two-Point Function: Spectral Representation

The two-point correlation function plays a special role in quantum field theory and statistical physics. Here, we derive a few of its general properties. We assume that the hamiltonian H is hermitian, bounded from below, and, to simplify the notation, has a discrete spectrum. In the basis in which H is diagonal the two-point correlation function

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

can then be written as

$$Z^{(2)}(t) = \sum_{n\geq 0} |\langle 0|\, \hat{q}\, |n\rangle|^2 \, \mathrm{e}^{-(\varepsilon_n - \varepsilon_0)|t|} . \tag{A2.1}$$

The quantities $|n\rangle$ and ε_n are, respectively, the eigenfunctions and eigenvalues of H. For $|t|$ large $Z^{(2)}$ has a limit, $Z^{(2)}(t) \to (\langle 0|\, \hat{q}\, |0\rangle)^2$, which can always be removed by translating $\hat{q}$: $\hat{q} \mapsto \hat{q} - \langle 0|\, \hat{q}\, |0\rangle$.

Note that, as a consequence of the hermiticity of H, the eigenvalues are real, and the exponentials in the sum of the r.h.s. have positive coefficients. The expansion $(A2.1)$ leads to a representation of the Fourier transform $\widetilde{Z}^{(2)}(\omega)$ of the two-point function:

$$\widetilde{Z}^{(2)}(\omega) = \int \mathrm{d}t\, Z^{(2)}(t)\, \mathrm{e}^{i\omega t} = 2\pi\, |\langle 0|\, \hat{q}\, |0\rangle|^2\, \delta(\omega) + 2\sum_{n>0} \frac{(\varepsilon_n - \varepsilon_0)\, |\langle 0|\, \hat{q}\, |n\rangle|^2}{\left[\omega^2 + (\varepsilon_n - \varepsilon_0)^2\right]} . \tag{A2.2}$$

Two properties of the Fourier transform follow: except for a possible distributive part at $\omega = 0$, it is an analytic function of ω^2 with poles only on the real negative axis. Moreover, the pole residues are all positive; it follows that $\widetilde{Z}^{(2)}(\omega)$ cannot decrease faster than $1/\omega^2$ for ω^2 large. More precisely, let us calculate the limit when $t \to 0+$ of the derivative of $Z^{(2)}(t)$

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

Since the l.h.s. is real we can replace the operator in the r.h.s. by its hermitian part:

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

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

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

Combining this result with the representation $(A2.1)$, one finds

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

APPENDIX A2

A2.1 The Two-Point Function: Spectral Representation

The two-point correlation function plays a special role in quantum field theory and statistical physics. Here, we derive a few of its general properties. We assume that the hamiltonian H is hermitian, bounded from below, and, to simplify the notation, has a discrete spectrum. In the basis in which H is diagonal the two-point correlation function

$$Z^{(2)}(t) \equiv \langle q(0)q(t)\rangle = \langle 0|\,\hat{q}\;\mathrm{e}^{-|t|H}\,\hat{q}\;\mathrm{e}^{|t|H}\,|0\rangle\,,$$

can then be written as

$$Z^{(2)}(t) = \sum_{n\geq 0} |\langle 0|\,\hat{q}\,|n\rangle|^2\;\mathrm{e}^{-(\varepsilon_n-\varepsilon_0)|t|}\,. \tag{A2.1}$$

The quantities $|n\rangle$ and ε_n are, respectively, the eigenfunctions and eigenvalues of H. For $|t|$ large $Z^{(2)}$ has a limit, $Z^{(2)}(t) \to (\langle 0|\,\hat{q}\,|0\rangle)^2$, which can always be removed by translating $\hat{q}$: $\hat{q} \mapsto \hat{q} - \langle 0|\,\hat{q}\,|0\rangle$.

Note that, as a consequence of the hermiticity of H, the eigenvalues are real, and the exponentials in the sum of the r.h.s. have positive coefficients. The expansion ($A2.1$) leads to a representation of the Fourier transform $\widetilde{Z}^{(2)}(\omega)$ of the two-point function:

$$\widetilde{Z}^{(2)}(\omega) = \int \mathrm{d}t\, Z^{(2)}(t)\,\mathrm{e}^{i\omega t} = 2\pi\,|\langle 0|\,\hat{q}\,|0\rangle|^2\,\delta(\omega) + 2\sum_{n>0}\frac{(\varepsilon_n-\varepsilon_0)\,|\langle 0|\,\hat{q}\,|n\rangle|^2}{\left[\omega^2+(\varepsilon_n-\varepsilon_0)^2\right]}\,. \tag{A2.2}$$

Two properties of the Fourier transform follow: except for a possible distributive part at $\omega = 0$, it is an analytic function of ω^2 with poles only on the real negative axis. Moreover, the pole residues are all positive; it follows that $\widetilde{Z}^{(2)}(\omega)$ cannot decrease faster than $1/\omega^2$ for ω^2 large. More precisely, let us calculate the limit when $t \to 0+$ of the derivative of $Z^{(2)}(t)$

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

Since the l.h.s. is real we can replace the operator in the r.h.s. by its hermitian part:

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

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

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

Combining this result with the representation ($A2.1$), one finds

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

and, therefore, from ($A2.2$):

$$\widetilde{Z}^{(2)}(\omega) \underset{\omega\to\infty}{\sim} \frac{1}{m\omega^2}\,.$$

This interesting result is not surprising. The behaviour for ω large is related to short time evolution, and we have seen that the most singular part of the matrix elements of $\mathrm{e}^{-\beta H}$ 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 ($A2.2$) is replaced by an integral, the poles are replaced by a cut with a positive discontinuity but the other conclusions are the same. The relativistic generalization of representation ($A2.2$) is called the Källen–Lehmann representation.

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

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

One verifies that the result is the same:

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

This implies in particular

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

and, therefore, from ($A2.2$):

$$\widetilde{Z}^{(2)}(\omega) \underset{\omega\to\infty}{\sim} \frac{1}{m\omega^2}\,.$$

This interesting result is not surprising. The behaviour for ω large is related to short time evolution, and we have seen that the most singular part of the matrix elements of $\mathrm{e}^{-\beta H}$ 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 ($A2.2$) is replaced by an integral, the poles are replaced by a cut with a positive discontinuity but the other conclusions are the same. The relativistic generalization of representation ($A2.2$) is called the Källen–Lehmann representation.

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

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

One verifies that the result is the same:

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

This implies in particular

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

3 PATH INTEGRALS IN QUANTUM MECHANICS: GENERALIZATIONS

In Chapter 2, we constructed a path integral representation of the matrix elements of the statistical operator $e^{-\beta H}$ in the case of hamiltonians H of the separable form $p^2/2m + V(q)$. In this chapter, we extend the construction to hamiltonians which are general functions of phase space variables. This results in integrals over trajectories or paths in phase space. When the hamiltonian is quadratic in 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, which is somewhat ambiguous, reflecting the problem of order between quantum operators.

Hamiltonians which are general quadratic functions of momentum variables provide other important examples, and we, therefore, analyse them thoroughly. We find that the problem of ambiguities is even more severe.

Such hamiltonians arise in the quantization of the motion on Riemannian manifolds. We illustrate the analysis by the quantization of the free motion on the sphere S_{N-1}.

In the appendix, we also discuss path integral quantization of systems for which an action generating the classical equation of motion cannot be globally defined. This 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 now consider a general *local* quantum hamiltonian $\hat{H}(t)$ and again want to calculate the matrix elements of the operator U, solution of

$$\hbar \frac{\partial U}{\partial t}(t,t') = -\hat{H}(t)U(t,t'), \qquad U(t',t') = \mathbf{1}\,, \tag{3.1}$$

in the basis in which the position operator $\hat{q}$ is diagonal. We follow the strategy of Section 2.1 and start from the identity (2.7),

$$\langle q''|\, U(t'',t')\,|q'\rangle = \int \prod_{k=1}^{n-1} \mathrm{d}q_k \prod_{k=1}^{n} \langle q_k|\, U(t_k,t_{k-1})\,|q_{k-1}\rangle\,, \tag{3.2}$$

with the conventions

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

In the large n, small ε limit, the calculation of the matrix elements of $U(t'',t')$ reduces to the evaluation of the matrix elements $\langle q|\, U(t+\varepsilon,t)\,|q'\rangle$ in the small time interval limit. In this limit, as a consequence of the locality of the hamiltonian, the support of $\langle q|\, U(t+\varepsilon,t)\,|q'\rangle$ is concentrated around $q = q'$.

If the quantum hamiltonian is known explicitly, and if it is no more than quadratic in p, we can follow the strategy of Section 2.2 and solve equation (3.1) written in terms of

matrix elements. Actually, often, the problem has a different formulation: we are given a classical hamiltonian $H(p,q,t)$ and we want to quantize it while preserving its symmetries. We, therefore, proceed rather formally to establish a path integral representation and then discuss the properties of the resulting expression.

The phase space path integral. To calculate the operator $U(t,t-\varepsilon)$ we solve equation (3.1) formally:

$$U(t,t-\varepsilon)=1-\frac{\varepsilon}{\hbar}\hat{H}(\hat{p},\hat{q},t)+O\left(\varepsilon^2\right), \tag{3.3}$$

in which $\hat{H}$ is one quantum hamiltonian which has H as a classical limit. We then take the matrix elements of the equation. The matrix elements of the identity yield a Dirac δ-function for which we use the Fourier representation:

$$\delta(q-q')=\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{ip(q-q')/\hbar}\,.$$

Since the hamiltonian $\hat{H}$ is local, $\hat{q}$ can be replaced by some average value q_{av}, a function of q and q' which tends towards the value q in the limit $q=q'$. Then $\hat{H}$ becomes a function of only one operator $\hat{p}$ and no commutation problems remain. The problem of quantization is now reduced to the choice of q_{av}, which is determined by the classical hamiltonian H only in the limit $q'=q$, $\hbar=0$. Note that any symmetric choice

$$q_{\mathrm{av}}(q,q')=q_{\mathrm{av}}(q',q)\quad\Rightarrow\quad\langle q|\,\hat{H}\,|q'\rangle=\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{ip(q-q')/\hbar}\,H\left(p,q_{\mathrm{av}},t\right), \tag{3.4}$$

has the property that a real classical hamiltonian is associated with a hermitian quantum hamiltonian. The Wigner quantization rule $q_{\mathrm{av}}(q,q')=\frac{1}{2}(q+q')$ shares this property. Only in the case of hamiltonians of the form $p^2+V(q)$ is the final result independent of the choice of quantization, as we have shown in Section 2.2. We shall also find out later that the situation in field theory is more favourable, in a sense, since for local hamiltonians commutators of conjugate variables are divergent (equation (6.55)) and have to be removed by renormalization. Dimensional regularization will be especially useful in this respect (see Section 9.6).

At order ε the matrix elements of equation (3.3) can be written in exponential form:

$$\langle q|\,U(t,t-\varepsilon)\,|q'\rangle\underset{\varepsilon\to0}{\sim}\int\frac{\mathrm{d}p}{2\pi\hbar}\exp\left\{\frac{1}{\hbar}\left[ip(q-q')-\varepsilon H\left(p,q_{\mathrm{av}},t\right)\right]\right\}. \tag{3.5}$$

Inserting this expression into equation (2.7) we obtain

$$\langle q''|\,U(t'',t')\,|q'\rangle=\lim_{n\to\infty}\int\prod_{k=1}^{n}\frac{\mathrm{d}p_k\mathrm{d}q_k}{2\pi\hbar}\delta(q_n-q'')\exp\left[-\mathcal{S}_\varepsilon(p,q)/\hbar\right] \tag{3.6}$$

with ($q_0\equiv q'$)

$$\mathcal{S}_\varepsilon(p,q)=\varepsilon\sum_{k=1}^{n}\left[-ip_k\left(q_k-q_{k-1}\right)/\varepsilon+H\left(p_k,q_{\mathrm{av}}\left(q_k,q_{k-1}\right),t_k\right)\right]. \tag{3.7}$$

We now introduce a trajectory in phase space $\{p(t),q(t)\}$ such that

$$p(t_k)=p_k\,,\quad q(t_k)=q_k\,,$$

matrix elements. Actually, often, the problem has a different formulation: we are given a classical hamiltonian $H(p,q,t)$ and we want to quantize it while preserving its symmetries. We, therefore, proceed rather formally to establish a path integral representation and then discuss the properties of the resulting expression.

The phase space path integral. To calculate the operator $U(t,t-\varepsilon)$ we solve equation (3.1) formally:

$$U(t,t-\varepsilon) = 1 - \frac{\varepsilon}{\hbar}\hat{H}(\hat{p},\hat{q},t) + O\left(\varepsilon^2\right), \tag{3.3}$$

in which $\hat{H}$ is one quantum hamiltonian which has H as a classical limit. We then take the matrix elements of the equation. The matrix elements of the identity yield a Dirac δ-function for which we use the Fourier representation:

$$\delta(q-q') = \int \frac{\mathrm{d}p}{2\pi\hbar}\, \mathrm{e}^{ip(q-q')/\hbar}\,.$$

Since the hamiltonian $\hat{H}$ is local, $\hat{q}$ can be replaced by some average value q_{av}, a function of q and q' which tends towards the value q in the limit $q = q'$. Then $\hat{H}$ becomes a function of only one operator $\hat{p}$ and no commutation problems remain. The problem of quantization is now reduced to the choice of q_{av}, which is determined by the classical hamiltonian H only in the limit $q' = q$, $\hbar = 0$. Note that any symmetric choice

$$q_{\mathrm{av}}(q,q') = q_{\mathrm{av}}(q',q) \quad \Rightarrow \quad \langle q|\,\hat{H}\,|q'\rangle = \int \frac{\mathrm{d}p}{2\pi\hbar}\, \mathrm{e}^{ip(q-q')/\hbar}\, H\left(p, q_{\mathrm{av}}, t\right), \tag{3.4}$$

has the property that a real classical hamiltonian is associated with a hermitian quantum hamiltonian. The Wigner quantization rule $q_{\mathrm{av}}(q,q') = \frac{1}{2}(q+q')$ shares this property. Only in the case of hamiltonians of the form $p^2 + V(q)$ is the final result independent of the choice of quantization, as we have shown in Section 2.2. We shall also find out later that the situation in field theory is more favourable, in a sense, since for local hamiltonians commutators of conjugate variables are divergent (equation (6.55)) and have to be removed by renormalization. Dimensional regularization will be especially useful in this respect (see Section 9.6).

At order ε the matrix elements of equation (3.3) can be written in exponential form:

$$\langle q|\,U(t,t-\varepsilon)\,|q'\rangle \underset{\varepsilon\to 0}{\sim} \int \frac{\mathrm{d}p}{2\pi\hbar} \exp\left\{\frac{1}{\hbar}\left[ip(q-q') - \varepsilon H\left(p, q_{\mathrm{av}}, t\right)\right]\right\}. \tag{3.5}$$

Inserting this expression into equation (2.7) we obtain

$$\langle q''|\,U(t'',t')\,|q'\rangle = \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}p_k \mathrm{d}q_k}{2\pi\hbar}\, \delta(q_n - q'') \exp\left[-\mathcal{S}_\varepsilon(p,q)/\hbar\right] \tag{3.6}$$

with $(q_0 \equiv q')$

$$\mathcal{S}_\varepsilon(p,q) = \varepsilon \sum_{k=1}^{n} \left[-ip_k\left(q_k - q_{k-1}\right)/\varepsilon + H\left(p_k, q_{\mathrm{av}}\left(q_k, q_{k-1}\right), t_k\right)\right]. \tag{3.7}$$

We now introduce a trajectory in phase space $\{p(t), q(t)\}$ such that

$$p(t_k) = p_k\,, \quad q(t_k) = q_k\,,$$

matrix elements. Actually, often, the problem has a different formulation: we are given a classical hamiltonian $H(p,q,t)$ and we want to quantize it while preserving its symmetries. We, therefore, proceed rather formally to establish a path integral representation and then discuss the properties of the resulting expression.

The phase space path integral. To calculate the operator $U(t,t-\varepsilon)$ we solve equation (3.1) formally:

$$U(t,t-\varepsilon)=1-\frac{\varepsilon}{\hbar}\hat{H}(\hat{p},\hat{q},t)+O\left(\varepsilon^2\right), \tag{3.3}$$

in which $\hat{H}$ is one quantum hamiltonian which has H as a classical limit. We then take the matrix elements of the equation. The matrix elements of the identity yield a Dirac δ-function for which we use the Fourier representation:

$$\delta(q-q')=\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{ip(q-q')/\hbar}\,.$$

Since the hamiltonian $\hat{H}$ is local, $\hat{q}$ can be replaced by some average value q_{av}, a function of q and q' which tends towards the value q in the limit $q=q'$. Then $\hat{H}$ becomes a function of only one operator $\hat{p}$ and no commutation problems remain. The problem of quantization is now reduced to the choice of q_{av}, which is determined by the classical hamiltonian H only in the limit $q'=q$, $\hbar=0$. Note that any symmetric choice

$$q_{\mathrm{av}}(q,q')=q_{\mathrm{av}}(q',q)\quad\Rightarrow\quad\langle q|\,\hat{H}\,|q'\rangle=\int\frac{\mathrm{d}p}{2\pi\hbar}\,\mathrm{e}^{ip(q-q')/\hbar}\,H\left(p,q_{\mathrm{av}},t\right), \tag{3.4}$$

has the property that a real classical hamiltonian is associated with a hermitian quantum hamiltonian. The Wigner quantization rule $q_{\mathrm{av}}(q,q')=\frac{1}{2}(q+q')$ shares this property. Only in the case of hamiltonians of the form $p^2+V(q)$ is the final result independent of the choice of quantization, as we have shown in Section 2.2. We shall also find out later that the situation in field theory is more favourable, in a sense, since for local hamiltonians commutators of conjugate variables are divergent (equation (6.55)) and have to be removed by renormalization. Dimensional regularization will be especially useful in this respect (see Section 9.6).

At order ε the matrix elements of equation (3.3) can be written in exponential form:

$$\langle q|\,U(t,t-\varepsilon)\,|q'\rangle\underset{\varepsilon\to0}{\sim}\int\frac{\mathrm{d}p}{2\pi\hbar}\exp\left\{\frac{1}{\hbar}\left[ip(q-q')-\varepsilon H\left(p,q_{\mathrm{av}},t\right)\right]\right\}. \tag{3.5}$$

Inserting this expression into equation (2.7) we obtain

$$\langle q''|\,U(t'',t')\,|q'\rangle=\lim_{n\to\infty}\int\prod_{k=1}^{n}\frac{\mathrm{d}p_k\mathrm{d}q_k}{2\pi\hbar}\delta(q_n-q'')\exp\left[-\mathcal{S}_\varepsilon(p,q)/\hbar\right] \tag{3.6}$$

with $(q_0\equiv q')$

$$\mathcal{S}_\varepsilon(p,q)=\varepsilon\sum_{k=1}^{n}\left[-ip_k\left(q_k-q_{k-1}\right)/\varepsilon+H\left(p_k,q_{\mathrm{av}}(q_k,q_{k-1}),t_k\right)\right]. \tag{3.7}$$

We now introduce a trajectory in phase space $\{p(t),q(t)\}$ such that

$$p(t_k)=p_k\,,\quad q(t_k)=q_k\,,$$

and express $\mathcal{S}_\varepsilon(p,q)$ in terms of the trajectory. We can then take the formal continuum time limit $\varepsilon \to 0$, $n \to \infty$. The limit of $\mathcal{S}_\varepsilon(p,q)$ is the euclidean action $\mathcal{S}(p,q)$ in the hamiltonian formalism:

$$\mathcal{S}(p,q) = \int_{t'}^{t''} \mathrm{d}t \left[-ip(t)\dot{q}(t) + H\left(p(t), q(t), t\right)\right]. \tag{3.8}$$

The limit of expression (3.6) is a path integral in phase space:

$$\langle q''|\, U(t'', t')\, |q'\rangle = \int_{q(t')=q'}^{q(t'')=q''} [\mathrm{d}p(t)\mathrm{d}q(t)] \exp\left[-\mathcal{S}(p,q)\,/\hbar\right], \tag{3.9}$$

in which the measure in phase space is normalized with respect to $2\pi\hbar$. Expression (3.9) 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

$$(p,q) \mapsto (P(p,q), Q(p,q))$$

with

$$\{p,q\} = \{P,Q\}.$$

Again the extension to several degrees of freedom is straightforward: one substitutes in the path integral representation (3.9) the corresponding classical action (after continuation to imaginary time) and Liouville measure.

Remarks.

(i) Note that in the formal continuum limit no trace of the special choice of quantization (3.4) remains. One thus suspects that the phase space path integral is not well-defined in general. In other words the continuum limit is in general not unique, and the formal path integral should be supplemented with some information about how the limit is taken.

(ii) The canonical invariance can be true only for a very restricted class of transformations. We will show elsewhere (Appendix A39.1) that in the case of a one-dimensional, one degree of freedom hamiltonian H one can always find a canonical transformation which maps H onto a free hamiltonian:

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

One could then naively conclude that semi-classical approximations are always exact. It is very easy to produce counter-examples. The discrete form (3.7) shows one origin of the difficulty. A variable p_k is associated to a pair (q_k, q_{k-1}) and, therefore, the canonical invariance is not true for the discretized form.

The space of integration. Problems arise when one tries to characterize in general the space of trajectories in phase space which contribute to the path integral (3.9). The term which connects different time steps in expression (3.7) is now $ip_k(q_k - q_{k-1})$. It leads to oscillations in the integral (3.6) which suppress trajectories not regular enough. The typical scale of the difference $(q_k - q_{k-1})$ for contributing trajectories is given by the typical values of p_k. For example if, as in Section 2.2, the hamiltonian is quadratic in p, the relevant values of p_k in integral (3.6) are of order $1/\sqrt{\varepsilon}$ and one, therefore, finds that $(q_k - q_{k-1})$ is of order $\sqrt{\varepsilon}$, a result which is consistent with the analysis of Section 2.2.

In the same way, we can rearrange expression (3.7) to transform the term $ip_k(q_k - q_{k-1})$ into $iq_k(p_k - p_{k-1})$ ("integration by parts"). To find the regularity conditions imposed on the function $p(t)$ we have to know the relevant values of q_k in integral (3.6). Considering, again, the example of a hamiltonian $p^2 + V(q)$ we see that if, for example, $V(q)$ grows for q large like q^{2N} the relevant values of q_k are such that εq^{2N} is of order 1, and, therefore, the difference $(p_k - p_{k-1})$ should be of order $\varepsilon^{1/2N}$. For a general hamiltonian, the discussion obviously becomes rather involved.

The conclusion is that the path integral in phase space is more difficult to handle in practice than the path integral over positions defined previously and, thus, has found fewer applications up to now. For any non-standard example one has to return to the discrete form (3.6) and make a special analysis.

3.2 Hamiltonians Quadratic in Momentum Variables

We have explained the difficulties one encounters when one tries to define a path integral in phase space. To show that expression (3.8) has, nevertheless, at least some heuristic value, we now discuss the example of general hamiltonians quadratic in momenta.

The consistency. We first verify that in the case of hamiltonians

$$H = p^2/2m + V(q),$$

the path integral (2.19) is recovered after integration over $p(t)$ in expression (3.9).

The classical action is

$$\mathcal{S}(p,q) = \int_{t'}^{t''} \mathrm{d}t\bigl(-ip\dot{q} + p^2/2m + V(q)\bigr). \tag{3.10}$$

In the path integral (3.9) the integral over the momentum variables $p(t)$ is then gaussian. Following the strategy explained in Section 2.5.1, we change variables $p(t) \mapsto r(t)$,

$$p(t) = im\dot{q}(t) + r(t). \tag{3.11}$$

The action becomes

$$\mathcal{S} = \int_{t''}^{t'} \mathrm{d}t \left(\frac{1}{2m} r^2(t) + \frac{1}{2} m\dot{q}^2 + V(q)\right).$$

The path integral thus factorizes into an integral over $r(t)$:

$$\mathcal{N}(t',t'') = \int [\mathrm{d}r(t)] \exp\left(-\frac{1}{\hbar}\int_{t'}^{t''} \frac{r^2(t)}{2m}\mathrm{d}t\right),$$

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 $q(t)$. The factor $\mathcal{N}$ can be calculated from the discretized expression (3.6), and is exactly the factor (2.20) that has been incorporated into the measure $[\mathrm{d}q(t)]$. Thus,

$$\langle q''|\, U(t'',t')\, |q'\rangle = \int_{q(t')=q'}^{q(t'')=q''} [\mathrm{d}q(t)] \exp\left[-\frac{1}{\hbar}\int_{t'}^{t''} \mathrm{d}t\bigl(\tfrac{1}{2}m\dot{q}^2 + V(q)\bigr)\right]. \tag{3.12}$$

We have, therefore, verified explicitly the consistency between the representations (3.9) and (2.19).

In the same way, we can rearrange expression (3.7) to transform the term $ip_k(q_k - q_{k-1})$ into $iq_k(p_k - p_{k-1})$ ("integration by parts"). To find the regularity conditions imposed on the function $p(t)$ we have to know the relevant values of q_k in integral (3.6). Considering, again, the example of a hamiltonian $p^2 + V(q)$ we see that if, for example, $V(q)$ grows for q large like q^{2N} the relevant values of q_k are such that εq^{2N} is of order 1, and, therefore, the difference $(p_k - p_{k-1})$ should be of order $\varepsilon^{1/2N}$. For a general hamiltonian, the discussion obviously becomes rather involved.

The conclusion is that the path integral in phase space is more difficult to handle in practice than the path integral over positions defined previously and, thus, has found fewer applications up to now. For any non-standard example one has to return to the discrete form (3.6) and make a special analysis.

3.2 Hamiltonians Quadratic in Momentum Variables

We have explained the difficulties one encounters when one tries to define a path integral in phase space. To show that expression (3.8) has, nevertheless, at least some heuristic value, we now discuss the example of general hamiltonians quadratic in momenta.

The consistency. We first verify that in the case of hamiltonians

$$H = p^2/2m + V(q),$$

the path integral (2.19) is recovered after integration over $p(t)$ in expression (3.9).

The classical action is

$$\mathcal{S}(p,q) = \int_{t'}^{t''} \mathrm{d}t\left(-ip\dot{q} + p^2/2m + V(q)\right). \tag{3.10}$$

In the path integral (3.9) the integral over the momentum variables $p(t)$ is then gaussian. Following the strategy explained in Section 2.5.1, we change variables $p(t) \mapsto r(t)$,

$$p(t) = im\dot{q}(t) + r(t). \tag{3.11}$$

The action becomes

$$\mathcal{S} = \int_{t''}^{t'} \mathrm{d}t \left(\frac{1}{2m} r^2(t) + \frac{1}{2} m\dot{q}^2 + V(q)\right).$$

The path integral thus factorizes into an integral over $r(t)$:

$$\mathcal{N}(t', t'') = \int [\mathrm{d}r(t)] \exp\left(-\frac{1}{\hbar}\int_{t'}^{t''} \frac{r^2(t)}{2m}\mathrm{d}t\right),$$

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 $q(t)$. The factor $\mathcal{N}$ can be calculated from the discretized expression (3.6), and is exactly the factor (2.20) that has been incorporated into the measure $[\mathrm{d}q(t)]$. Thus,

$$\langle q''|\, U(t'', t')\, |q'\rangle = \int_{q(t')=q'}^{q(t'')=q''} [\mathrm{d}q(t)] \exp\left[-\frac{1}{\hbar}\int_{t'}^{t''} \mathrm{d}t\left(\tfrac{1}{2}m\dot{q}^2 + V(q)\right)\right]. \tag{3.12}$$

We have, therefore, verified explicitly the consistency between the representations (3.9) and (2.19).

3.2.1 Quantization in a magnetic field

We now consider the classical hamiltonian of a particle in a potential and a magnetic field:

$$H = \frac{1}{2m}\left[\mathbf{p} + e\mathbf{A}(\mathbf{q})\right]^2 + V(\mathbf{q}), \tag{3.13}$$

where $\mathbf{A}$ is the vector potential and e the electric charge. The quantization of this hamiltonian leads to a problem of order of quantum operators in the term $\mathbf{p} \cdot \mathbf{A}(\mathbf{q})$. However, the condition of hermiticity uniquely determines the quantum hamiltonian,

$$\hat{H} = \frac{1}{2m}\left[\hat{\mathbf{p}}^2 + e\mathbf{A}(\hat{\mathbf{q}}) \cdot \hat{\mathbf{p}} + e\hat{\mathbf{p}} \cdot \mathbf{A}(\hat{\mathbf{q}}) + e^2\mathbf{A}^2(\hat{\mathbf{q}})\right] + V(\hat{\mathbf{q}}), \tag{3.14}$$

since a change in the order of operators is equivalent to the addition of an imaginary potential proportional to $ie\hbar\nabla \cdot \mathbf{A}/2m$.

We recall that a characteristic property of the hamiltonian (3.14) is *gauge invariance.* In the presence of a magnetic field the phase of wave functions $\psi(\mathbf{q})$ can be changed at each point of space independently. The transformation

$$\psi(\mathbf{q}) \mapsto \psi(\mathbf{q})\, \mathrm{e}^{-ie\Lambda(\mathbf{q})/\hbar}, \tag{3.15}$$

can be cancelled by adding a gradient term to the vector potential $\mathbf{A}(\mathbf{q})$:

$$\mathbf{A}(\mathbf{q}) \mapsto \mathbf{A}(\mathbf{q}) + \nabla\Lambda(\mathbf{q}), \tag{3.16}$$

which does not affect the magnetic field.

A direct solution of equation (3.1) for small time intervals using hamiltonian (3.14), yields a discretized form of the additional magnetic term which can be written (see Appendix A18.3) as

$$ie\sum_k \int_0^1 \mathrm{d}s\, (\mathbf{q}_k - \mathbf{q}_{k-1}) \cdot \mathbf{A}\big((1-s)\mathbf{q}_{k-1} + s\mathbf{q}_k\big) = ie\int \mathrm{d}t\, \dot{\mathbf{q}}(t) \cdot \mathbf{A}\big(\mathbf{q}(t)\big), \tag{3.17}$$

where $\mathbf{q}(t)$ is the piece-wise linear trajectory interpolating the points $\mathbf{q}_k$, $\mathbf{q}(t_k) = \mathbf{q}_k$. Note that the argument of $\mathbf{A}$ is symmetric in $\mathbf{q}_k, \mathbf{q}_{k-1}$. A non-symmetric choice would differ by a term proportional $(\mathbf{q}_k - \mathbf{q}_{k-1})_i(\mathbf{q}_k - \mathbf{q}_{k-1})_j$ which is of order ε and, thus, would contribute in the continuum limit. This would result in a non-hermitian quantization of the classical hamiltonian.

If we use the general action (3.8), instead, we see that the integral over momentum is still gaussian. To eliminate the terms linear in $\mathbf{p}$ in the action we again change variables $\mathbf{p}(t) \mapsto \mathbf{r}(t)$,

$$\mathbf{p}(t) = im\dot{\mathbf{q}}(t) - e\mathbf{A}\big(\mathbf{q}(t)\big) + \mathbf{r}(t).$$

After integration over $\mathbf{r}(t)$, we obtain an integral over the path $\mathbf{q}(t)$ with the action

$$\mathcal{S}(\mathbf{q}) = \int_{t'}^{t''} \mathrm{d}t \left[\tfrac{1}{2}m\dot{\mathbf{q}}^2 + ie\mathbf{A}(\mathbf{q}) \cdot \dot{\mathbf{q}} + V(\mathbf{q})\right], \tag{3.18}$$

in which we recognize the euclidean classical action, with a magnetic term consistent with expression (3.17).

We notice 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 (3.16) adds to the lagrangian a total derivative and the variation of the action, $\delta\mathcal{S}$, is

$$\delta\mathcal{S} = ie\left[\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}')\right].$$

It follows that the matrix elements $\langle \mathbf{q}''|\, U(t'',t')\, |\mathbf{q}'\rangle$ are multiplied by the phase factor $\exp\left[-ie\left(\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}')\right)/\hbar\right]$, in agreement with the transformation law (3.15).

Hermiticity and the $\epsilon(0)$ *problem.* In the formal continuum limit the precise choice of quantization is lost. To understand where the problem emerges, we again expand the path integral corresponding to the action

$$\mathcal{S}(\mathbf{q}) = \int \mathrm{d}t \left[\tfrac{1}{2}\dot{\mathbf{q}}^2 + ie\mathbf{A}(\mathbf{q})\cdot\dot{\mathbf{q}} + \tfrac{1}{2}\omega^2\mathbf{q}^2\right],$$

in powers of the charge e. At first order, we find a contribution which, using Wick's theorem, can be written as

$$-\frac{ie}{\hbar}\int \mathrm{d}t\, \langle \dot{q}_i(t) q_j(t)\rangle \left\langle \frac{\partial A_i}{\partial q_j}\right\rangle.$$

From equations (2.46,2.49), we derive

$$\langle \dot{q}_i(t_1) q_j(t_2)\rangle = -\tfrac{1}{2}\hbar\,\delta_{ij}\,\epsilon(t_1 - t_2)\,\mathrm{e}^{-\omega|t_1 - t_2|},$$

where $\epsilon(t)$ is the sign function: $\epsilon(t) = 1$ for $t > 0$, $\epsilon(-t) = -\epsilon(t)$. Therefore, formally

$$-\frac{ie}{\hbar}\int \mathrm{d}t\, \langle \dot{q}_i(t) q_j(t)\rangle \left\langle \frac{\partial A_i}{\partial q_j}\right\rangle = \tfrac{1}{2}ie\epsilon(0)\int \mathrm{d}t\, \langle \nabla\cdot\mathbf{A}(\mathbf{q})\rangle .$$

Of course, the result is ambiguous because $\epsilon(0) = 0$ is undefined.

Note that the assignment $\epsilon(0) = 0$ preserves invariance under time reversal. Moreover, from (2.46)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle q_i(t) q_j(t)\rangle = 0 .$$

Therefore, only the choice $\epsilon(0) = 0$ is *consistent with the commutation between time derivative and averaging.*

Another choice would be equivalent to the addition of a term $-\frac{1}{2}ie\epsilon(0)\hbar\nabla\cdot\mathbf{A}(\mathbf{q})$ 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 complicated formalism.

Finally, the operator $\mathrm{e}^{-\beta H}$ 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 $\mathcal{S} \to \mathcal{S}^*$ (complex conjugation) and $t \to -t$ (transposition and thus exchange of boundary conditions). By preserving this symmetry throughout the calculation, and, therefore, setting $\epsilon(0) = 0$, one ensures consistency with the (hermitian) choice of quantization (3.14).

We shall meet this problem again, in Chapter 4, in Sections 5.2,5.3,39.6, and it will be implicit in field theories with derivative couplings, like gauge theories.

We notice 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 (3.16) adds to the lagrangian a total derivative and the variation of the action, $\delta\mathcal{S}$, is

$$\delta\mathcal{S} = ie\left[\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}')\right].$$

It follows that the matrix elements $\langle\mathbf{q}''|\, U(t'',t')\,|\mathbf{q}'\rangle$ are multiplied by the phase factor $\exp\left[-ie\big(\Lambda(\mathbf{q}'') - \Lambda(\mathbf{q}')\big)/\hbar\right]$, in agreement with the transformation law (3.15).

Hermiticity and the $\epsilon(0)$ problem. In the formal continuum limit the precise choice of quantization is lost. To understand where the problem emerges, we again expand the path integral corresponding to the action

$$\mathcal{S}(\mathbf{q}) = \int \mathrm{d}t\,\left[\tfrac{1}{2}\dot{\mathbf{q}}^2 + ie\mathbf{A}(\mathbf{q})\cdot\dot{\mathbf{q}} + \tfrac{1}{2}\omega^2\mathbf{q}^2\right],$$

in powers of the charge e. At first order, we find a contribution which, using Wick's theorem, can be written as

$$-\frac{ie}{\hbar}\int \mathrm{d}t\,\langle\dot{q}_i(t)q_j(t)\rangle\left\langle\frac{\partial A_i}{\partial q_j}\right\rangle.$$

From equations (2.46,2.49), we derive

$$\langle\dot{q}_i(t_1)q_j(t_2)\rangle = -\tfrac{1}{2}\hbar\,\delta_{ij}\,\epsilon(t_1-t_2)\,\mathrm{e}^{-\omega|t_1-t_2|},$$

where $\epsilon(t)$ is the sign function: $\epsilon(t)=1$ for $t>0$, $\epsilon(-t)=-\epsilon(t)$. Therefore, formally

$$-\frac{ie}{\hbar}\int \mathrm{d}t\,\langle\dot{q}_i(t)q_j(t)\rangle\left\langle\frac{\partial A_i}{\partial q_j}\right\rangle = \tfrac{1}{2}ie\epsilon(0)\int\mathrm{d}t\,\langle\nabla\cdot\mathbf{A}(\mathbf{q})\rangle.$$

Of course, the result is ambiguous because $\epsilon(0)=0$ is undefined.

Note that the assignment $\epsilon(0)=0$ preserves invariance under time reversal. Moreover, from (2.46)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle q_i(t)q_j(t)\rangle = 0.$$

Therefore, only the choice $\epsilon(0)=0$ is *consistent with the commutation between time derivative and averaging.*

Another choice would be equivalent to the addition of a term $-\frac{1}{2}ie\epsilon(0)\hbar\nabla\cdot\mathbf{A}(\mathbf{q})$ 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 complicated formalism.

Finally, the operator $\mathrm{e}^{-\beta H}$ 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 $\mathcal{S}\to\mathcal{S}^*$ (complex conjugation) and $t\to -t$ (transposition and thus exchange of boundary conditions). By preserving this symmetry throughout the calculation, and, therefore, setting $\epsilon(0)=0$, one ensures consistency with the (hermitian) choice of quantization (3.14).

We shall meet this problem again, in Chapter 4, in Sections 5.2,5.3,39.6, and it will be implicit in field theories with derivative couplings, like gauge theories.

3.2.2 General quadratic hamiltonian

A general hamiltonian quadratic in $\mathbf{p}$ can be derived from a general lagrangian quadratic in the velocities. Because in all examples we shall encounter, the quantization problem will initially be formulated in terms of a classical lagrangian, we now assume that we want to quantize the real time lagrangian $\mathcal{L}(\dot{q}, q)$:

$$\mathcal{L}(\dot{q}, q) = \tfrac{1}{2}\dot{q}^\alpha g_{\alpha\beta}(q)\dot{q}^\beta - \dot{q}^\alpha h_\alpha(q) - v(q), \tag{3.19}$$

where $g_{\alpha\beta}(q)$ is a positive matrix. The most interesting examples correspond to quantum mechanics in Riemannian manifolds, as we shall briefly discuss in Section 4.8 (for notation and details see Chapter 22). The tensor $g_{\alpha\beta}(q)$ then is the metric tensor. Field theoretical generalizations will be studied in Chapters 14,15.

The corresponding classical hamiltonian is obtained by a *Legendre transformation* (see also Section 1.8). The conjugate momenta are

$$p_\alpha = \frac{\partial \mathcal{L}(\dot{q}, q)}{\partial \dot{q}^\alpha} = g_{\alpha\beta}(q)\dot{q}^\beta - h_\alpha(q),$$

and, therefore, the hamiltonian is

$$\begin{aligned} H(p,q) &= p_\alpha \dot{q}^\alpha - \mathcal{L}(\dot{q}, q) \\ &= \tfrac{1}{2}\left(p_\alpha + h_\alpha(q)\right) g^{\alpha\beta}(q)\left(p_\beta + h_\beta(q)\right) + v(q), \end{aligned} \tag{3.20}$$

where the traditional notation $g^{\alpha\beta}$ for the matrix inverse of $g_{\alpha\beta}$:

$$g_{\alpha\gamma}(q) g^{\gamma\beta}(q) = \delta_\alpha^\beta$$

has been used.

Again the integration over $p(t)$ is gaussian and can be performed. However, a difficulty appears with the evaluation of the determinant resulting from the gaussian integration and we, therefore, perform the p integration on the discretized form (3.6) or actually (3.5). The expression (3.5) in the case of the hamiltonian (3.20) reads:

$$\begin{aligned} \langle \mathbf{q}| U(t, t-\varepsilon) |\mathbf{q}'\rangle = \int \frac{\mathrm{d}p}{2\pi\hbar} \exp\Bigg\{ \frac{\varepsilon}{\hbar}\Bigg[&\frac{i}{\varepsilon}\left(q - q'\right)^\alpha p_\alpha - \frac{1}{2}\left(p_\alpha + h_\alpha(\mathbf{q}_{\mathrm{av}})\right) \\ &\times g^{\alpha\beta}(\mathbf{q}_{\mathrm{av}})\left(p_\beta + h_\beta(\mathbf{q}_{\mathrm{av}})\right) - v(\mathbf{q}_{\mathrm{av}})\Bigg]\Bigg\}. \end{aligned} \tag{3.21}$$

The integration yields

$$\langle \mathbf{q}| U(t, t-\varepsilon) |\mathbf{q}'\rangle = \left[2\pi\hbar\varepsilon \det \mathbf{g}(\mathbf{q}_{\mathrm{av}})\right]^{1/2} \exp\left[-\mathcal{S}(\mathbf{q}, \mathbf{q}'; \varepsilon)/\hbar\right],$$

with

$$\begin{aligned} \mathcal{S}(\mathbf{q}, \mathbf{q}'; \varepsilon)/\varepsilon = {} & \tfrac{1}{2}\left[(q^\alpha - q'^\alpha)/\varepsilon\right] g_{\alpha\beta}(\mathbf{q}_{\mathrm{av}}) \left[(q^\beta - q'^\beta)/\varepsilon\right] \\ & + i\left[(q^\alpha - q'^\alpha)/\varepsilon\right] h_\alpha(\mathbf{q}_{\mathrm{av}}) + v(\mathbf{q}_{\mathrm{av}}), \end{aligned} \tag{3.22}$$

and $\mathbf{g}$ denotes the matrix with matrix elements $g_{\alpha\beta}$.

Returning to expression (3.6) and taking the formal continuum limit we verify that the exponential factor again yields the euclidean action $\mathcal{S}$, integral of the classical lagrangian (3.19) after continuation to imaginary time:

$$\mathcal{S}(\mathbf{q}) = \int_{t'}^{t''} \mathrm{d}t \left[\tfrac{1}{2}\dot{q}^\alpha g_{\alpha\beta}(\mathbf{q})\dot{q}^\beta + i\dot{q}^\alpha h_\alpha(\mathbf{q}) + v(\mathbf{q})\right]. \tag{3.23}$$

This is not surprising since the integration over p_α is equivalent to a Legendre transformation.

However, in contrast with the previous two examples, the integration generates a non-trivial $q(t)$ dependent normalization factor $\mathcal{N}(\mathbf{q})$:

$$\langle \mathbf{q}''| U(t''t') |\mathbf{q}'\rangle = \int [\mathrm{d}q(t)]\, \mathcal{N}(\mathbf{q}) \exp\left[-\mathcal{S}(\mathbf{q})/\hbar\right] \tag{3.24}$$

with

$$\mathcal{N}(\mathbf{q}) \underset{\varepsilon\to 0}{\sim} (2\pi\hbar\,\varepsilon)^{-n/2} \exp\left[\frac{1}{2}\sum_{k=1}^{n} \ln\det \mathbf{g}(\mathbf{q}_{\mathrm{av},k})\right], \tag{3.25}$$

and thus yields a divergent quantum correction to the classical action (it has no factor $1/\hbar$):

$$\mathcal{N}(\mathbf{q}) \sim \exp\left[\frac{1}{2\varepsilon}\int_{t'}^{t''} \ln\det \mathbf{g}\left[\mathbf{q}(t)\right] \mathrm{d}t\right], \tag{3.26}$$

or equivalently using the identity (1.101)

$$\ln\det \mathbf{g} = \operatorname{tr}\ln \mathbf{g},$$

$$\mathcal{N}(\mathbf{q}) \sim \exp\left[\frac{1}{2\varepsilon}\int_{t'}^{t''} \operatorname{tr}\ln \mathbf{g}\left[\mathbf{q}(t)\right] \mathrm{d}t\right]. \tag{3.27}$$

A formal calculation, starting from expression (3.9), yields a similar result with $1/\varepsilon$ replaced by $\delta(0)$ (δ being the Dirac δ-function). This difficulty is directly related to the problem of ordering operators. If one performs a small $\hbar$ (semi-classical) expansion of the path integral, one finds a divergent quantum correction (see Chapters 14,15). This divergence is cancelled by the leading contribution coming from (3.27). However, the remaining finite part is ill-defined in the formal continuum time limit. It is necessary to use the discretized form (3.6), which reflects a choice of quantization, to calculate it. Another direct way of understanding this ambiguity is to notice that, since in expression (3.26) the difference $|q - q'|$ is generically of order $\sqrt{\varepsilon}$, a replacement of $\mathbf{g}(\mathbf{q}_{\mathrm{av}})$ by any other symmetric function of q and q' which has the same $q = q'$ limit changes in general this quantity at order ε. 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 Appendix A39.2.

From the point of view of the 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.

We finally note that this problem becomes worse if the classical hamiltonian is a polynomial of higher degree in p.

Remark. When $g_{\alpha\beta}$ is the metric tensor in a Riemannian manifold, the factor (3.25) formally reconstructs the *covariant measure* in the manifold (see Section 22.5).

Returning to expression (3.6) and taking the formal continuum limit we verify that the exponential factor again yields the euclidean action $\mathcal{S}$, integral of the classical lagrangian (3.19) after continuation to imaginary time:

$$\mathcal{S}(\mathbf{q}) = \int_{t'}^{t''} \mathrm{d}t \left[\tfrac{1}{2}\dot{q}^\alpha g_{\alpha\beta}(\mathbf{q})\dot{q}^\beta + i\dot{q}^\alpha h_\alpha(\mathbf{q}) + v(\mathbf{q})\right]. \tag{3.23}$$

This is not surprising since the integration over p_α is equivalent to a Legendre transformation.

However, in contrast with the previous two examples, the integration generates a non-trivial $q(t)$ dependent normalization factor $\mathcal{N}(\mathbf{q})$:

$$\langle \mathbf{q}''| U(t''t') |\mathbf{q}'\rangle = \int [\mathrm{d}q(t)] \mathcal{N}(\mathbf{q}) \exp\left[-\mathcal{S}(\mathbf{q})/\hbar\right] \tag{3.24}$$

with

$$\mathcal{N}(\mathbf{q}) \underset{\varepsilon\to 0}{\sim} (2\pi\hbar\,\varepsilon)^{-n/2} \exp\left[\frac{1}{2}\sum_{k=1}^{n} \ln\det \mathbf{g}(\mathbf{q}_{\mathrm{av},k})\right], \tag{3.25}$$

and thus yields a divergent quantum correction to the classical action (it has no factor $1/\hbar$):

$$\mathcal{N}(\mathbf{q}) \sim \exp\left[\frac{1}{2\varepsilon}\int_{t'}^{t''} \ln\det \mathbf{g}\left[\mathbf{q}(t)\right] \mathrm{d}t\right], \tag{3.26}$$

or equivalently using the identity (1.101)

$$\ln\det \mathbf{g} = \operatorname{tr}\ln \mathbf{g},$$

$$\mathcal{N}(\mathbf{q}) \sim \exp\left[\frac{1}{2\varepsilon}\int_{t'}^{t''} \operatorname{tr}\ln \mathbf{g}\left[\mathbf{q}(t)\right] \mathrm{d}t\right]. \tag{3.27}$$

A formal calculation, starting from expression (3.9), yields a similar result with $1/\varepsilon$ replaced by $\delta(0)$ (δ being the Dirac δ-function). This difficulty is directly related to the problem of ordering operators. If one performs a small $\hbar$ (semi-classical) expansion of the path integral, one finds a divergent quantum correction (see Chapters 14,15). This divergence is cancelled by the leading contribution coming from (3.27). However, the remaining finite part is ill-defined in the formal continuum time limit. It is necessary to use the discretized form (3.6), which reflects a choice of quantization, to calculate it. Another direct way of understanding this ambiguity is to notice that, since in expression (3.26) the difference $|q - q'|$ is generically of order $\sqrt{\varepsilon}$, a replacement of $\mathbf{g}(\mathbf{q}_{\mathrm{av}})$ by any other symmetric function of q and q' which has the same $q = q'$ limit changes in general this quantity at order ε. 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 Appendix A39.2.

From the point of view of the 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.

We finally note that this problem becomes worse if the classical hamiltonian is a polynomial of higher degree in p.

Remark. When $g_{\alpha\beta}$ is the metric tensor in a Riemannian manifold, the factor (3.25) formally reconstructs the *covariant measure* in the manifold (see Section 22.5).

3.3 The Spectrum of the $O(2)$ Symmetric Rigid Rotator

To illustrate the discussion of hamiltonians quadratic in momentum variables, we now calculate the spectrum of the $O(N)$ rigid rotator from the path integral representation (this model is also the one-dimensional $O(N)$ non-linear σ model, see Chapter 14). 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.

For a general parametrization of the circle the hamiltonian is a quadratic function of momenta with a position-dependent coefficient, and quantization problems. However, if the circle is parametrized by an angle θ, the hamiltonian of the $O(2)$ rotator becomes

$$H = -\frac{1}{2}\frac{\partial^2}{(\partial\theta)^2}, \tag{3.28}$$

and would be a free hamiltonian if θ were not an angular variable. As a consequence, the spectrum, instead of being continuous, is discrete

$$E_l = \tfrac{1}{2}l^2, \tag{3.29}$$

where l is an integer, the angular momentum.

The path integral representation of the matrix elements of the operator $\mathrm{e}^{-\beta H}$ is similar to the free path integral:

$$\langle\theta''|\,\mathrm{e}^{-\beta H}\,|\theta'\rangle = \int_{\theta(0)=\theta'}^{\theta(\beta)=\theta''}[\mathrm{d}\theta(t)]\exp\left[-\frac{1}{2}\int_0^\beta\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\mathrm{d}t\right]. \tag{3.30}$$

However, we now show that the cyclic character of the variable also modifies the evaluation of the path integral. As usual, we first solve the classical equation of motion. Since θ' and θ'' are angles, we now find a family of trajectories which go from θ' to θ''

$$\theta_c(t) = \theta' + t\left(\theta'' - \theta' + 2\pi n\right)/\beta, \quad \text{with } n \in \mathbb{Z}\,. \tag{3.31}$$

These trajectories are topologically distinct, that is, they cannot be related by continuous deformation. Hence, they have all to be taken into account because fluctuations around one trajectory do not include any other one.

We then shift $\theta(t)$, $\theta(t) \mapsto u(t)$, by the solution

$$\theta(t) = \theta_c(t) + u(t)\,. \tag{3.32}$$

The path integral (3.30) then becomes a sum of contributions:

$$\langle\theta''|\,\mathrm{e}^{-\beta H}\,|\theta'\rangle = \sum_{n=-\infty}^{+\infty}\mathcal{N}(\beta)\exp\left[-\frac{1}{2\beta}\left(\theta'' - \theta' + 2\pi n\right)^2\right]. \tag{3.33}$$

The normalization $\mathcal{N}(\beta)$ is given by a path integral which is independent of θ', θ'' and n. 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,

$$\mathcal{N}(\beta) = \sqrt{2\pi/\beta}\,.$$

3.3 The Spectrum of the $O(2)$ Symmetric Rigid Rotator

To illustrate the discussion of hamiltonians quadratic in momentum variables, we now calculate the spectrum of the $O(N)$ rigid rotator from the path integral representation (this model is also the one-dimensional $O(N)$ non-linear σ model, see Chapter 14). 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.

For a general parametrization of the circle the hamiltonian is a quadratic function of momenta with a position-dependent coefficient, and quantization problems. However, if the circle is parametrized by an angle θ, the hamiltonian of the $O(2)$ rotator becomes

$$H = -\frac{1}{2}\frac{\partial^2}{(\partial\theta)^2}, \tag{3.28}$$

and would be a free hamiltonian if θ were not an angular variable. As a consequence, the spectrum, instead of being continuous, is discrete

$$E_l = \tfrac{1}{2}l^2, \tag{3.29}$$

where l is an integer, the angular momentum.

The path integral representation of the matrix elements of the operator $\mathrm{e}^{-\beta H}$ is similar to the free path integral:

$$\langle\theta''|\,\mathrm{e}^{-\beta H}\,|\theta'\rangle = \int_{\theta(0)=\theta'}^{\theta(\beta)=\theta''}[\mathrm{d}\theta(t)]\exp\left[-\frac{1}{2}\int_0^\beta\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\mathrm{d}t\right]. \tag{3.30}$$

However, we now show that the cyclic character of the variable also modifies the evaluation of the path integral. As usual, we first solve the classical equation of motion. Since θ' and θ'' are angles, we now find a family of trajectories which go from θ' to θ''

$$\theta_c(t) = \theta' + t\,(\theta'' - \theta' + 2\pi n)\,/\beta, \quad \text{with } n \in \mathbb{Z}\,. \tag{3.31}$$

These trajectories are topologically distinct, that is, they cannot be related by continuous deformation. Hence, they have all to be taken into account because fluctuations around one trajectory do not include any other one.

We then shift $\theta(t)$, $\theta(t) \mapsto u(t)$, by the solution

$$\theta(t) = \theta_c(t) + u(t)\,. \tag{3.32}$$

The path integral (3.30) then becomes a sum of contributions:

$$\langle\theta''|\,\mathrm{e}^{-\beta H}\,|\theta'\rangle = \sum_{n=-\infty}^{+\infty}\mathcal{N}(\beta)\exp\left[-\frac{1}{2\beta}\left(\theta'' - \theta' + 2\pi n\right)^2\right]. \tag{3.33}$$

The normalization $\mathcal{N}(\beta)$ is given by a path integral which is independent of θ', θ'' and n. 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,

$$\mathcal{N}(\beta) = \sqrt{2\pi/\beta}\,.$$

The expression (3.33) can be rewritten using Poisson's formula. Being a periodic function of $\theta'' - \theta'$, it has a Fourier series expansion:

$$\langle \theta'' | \mathrm{e}^{-\beta H} | \theta' \rangle = \sum_{l=-\infty}^{+\infty} \mathrm{e}^{il(\theta''-\theta')} \, \mathrm{e}^{-\beta E_l}, \tag{3.34}$$

whose coefficients are given by

$$\exp\left[-\beta E_l\right] = \frac{1}{\sqrt{2\pi\beta}} \int_0^{2\pi} \mathrm{d}\theta \, \mathrm{e}^{-il\theta} \sum_{n=-\infty}^{+\infty} \exp\left[-\frac{1}{2\beta}\left(\theta + 2\pi n\right)^2\right].$$

Inverting summation and integration, and changing variables, $\theta + 2n\pi \mapsto \theta$, one then finds

$$\exp\left(-\beta E_l\right) = \frac{1}{\sqrt{2\pi\beta}} \int_{-\infty}^{+\infty} \mathrm{d}\theta \, \exp\left[-\left(il\theta + \theta^2/2\beta\right)\right].$$

The gaussian integration finally yields the exact spectrum (3.29).

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 shall see now, because the group manifold has curvature.

3.4 The Spectrum of the $O(N)$ Symmetric Rigid Rotator

The hamiltonian of the $O(N)$ rigid rotator can be written as

$$H = \tfrac{1}{2}\mathbf{L}^2, \tag{3.35}$$

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:

$$H = \tfrac{1}{2} g^{ij}(\mathbf{q}) p_i p_j \,, \tag{3.36}$$

where $g^{ij}(\mathbf{q})$ is the inverse of the metric tensor on the sphere (see also Section 4.8). For example, if the sphere is parametrized locally by a vector $\mathbf{r}$ in $\mathbb{R}^N$ of unit length and components $\left(\mathbf{q}, (1-\mathbf{q}^2)^{1/2}\right)$ the inverse metric tensor reads

$$g^{ij}(\mathbf{q}) = \delta^{ij} - q^i q^j. \tag{3.37}$$

According to the discussion of Section 3.2.2, the corresponding path integral representation of $\mathrm{e}^{-\beta H}$ is then:

$$\langle \mathbf{q}'' | \mathrm{e}^{-\beta H} | \mathbf{q}' \rangle = \int \left[\sqrt{g\left(\mathbf{q}(t)\right)} \mathrm{d}\mathbf{q}(t)\right] \exp\left[-\tfrac{1}{2} \int_0^\beta \mathrm{d}t \, g_{ij}\left(\mathbf{q}(t)\right) \dot{q}^i \dot{q}^j\right], \tag{3.38}$$

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. One verifies that the path integral (3.38) can also be expressed in terms of a vector $\mathbf{r}$ of unit length:

$$\langle \mathbf{r}'' | \mathrm{e}^{-\beta H} | \mathbf{r}' \rangle = \int_{\mathbf{r}(0)=\mathbf{r}'}^{\mathbf{r}(\beta)=\mathbf{r}''} \left[\mathrm{d}\mathbf{r}(t) \delta\left(1 - \mathbf{r}^2(t)\right)\right] \exp\left[-\tfrac{1}{2} \int_0^\beta \mathrm{d}t \, \dot{\mathbf{r}}^2(t)\right]. \tag{3.39}$$

The expression (3.33) can be rewritten using Poisson's formula. Being a periodic function of $\theta'' - \theta'$, it has a Fourier series expansion:

$$\langle \theta'' | \, \mathrm{e}^{-\beta H} \, | \theta' \rangle = \sum_{l=-\infty}^{+\infty} \mathrm{e}^{il(\theta''-\theta')} \, \mathrm{e}^{-\beta E_l} , \tag{3.34}$$

whose coefficients are given by

$$\exp\left[-\beta E_l\right] = \frac{1}{\sqrt{2\pi\beta}} \int_0^{2\pi} \mathrm{d}\theta \, \mathrm{e}^{-il\theta} \sum_{n=-\infty}^{+\infty} \exp\left[-\frac{1}{2\beta}\left(\theta + 2\pi n\right)^2\right].$$

Inverting summation and integration, and changing variables, $\theta + 2n\pi \mapsto \theta$, one then finds

$$\exp\left(-\beta E_l\right) = \frac{1}{\sqrt{2\pi\beta}} \int_{-\infty}^{+\infty} \mathrm{d}\theta \, \exp\left[-\left(il\theta + \theta^2/2\beta\right)\right].$$

The gaussian integration finally yields the exact spectrum (3.29).

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 shall see now, because the group manifold has curvature.

3.4 The Spectrum of the $O(N)$ Symmetric Rigid Rotator

The hamiltonian of the $O(N)$ rigid rotator can be written as

$$H = \tfrac{1}{2}\mathbf{L}^2, \tag{3.35}$$

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:

$$H = \tfrac{1}{2} g^{ij}(\mathbf{q}) p_i p_j , \tag{3.36}$$

where $g^{ij}(\mathbf{q})$ is the inverse of the metric tensor on the sphere (see also Section 4.8). For example, if the sphere is parametrized locally by a vector $\mathbf{r}$ in $\mathbb{R}^N$ of unit length and components $\left(\mathbf{q}, (1-\mathbf{q}^2)^{1/2}\right)$ the inverse metric tensor reads

$$g^{ij}(\mathbf{q}) = \delta^{ij} - q^i q^j . \tag{3.37}$$

According to the discussion of Section 3.2.2, the corresponding path integral representation of $\mathrm{e}^{-\beta H}$ is then:

$$\langle \mathbf{q}'' | \, \mathrm{e}^{-\beta H} \, | \mathbf{q}' \rangle = \int \left[\sqrt{g\left(\mathbf{q}(t)\right)} \mathrm{d}\mathbf{q}(t)\right] \exp\left[-\tfrac{1}{2}\int_0^\beta \mathrm{d}t \, g_{ij}\left(\mathbf{q}(t)\right) \dot{q}^i \dot{q}^j\right], \tag{3.38}$$

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. One verifies that the path integral (3.38) can also be expressed in terms of a vector $\mathbf{r}$ of unit length:

$$\langle \mathbf{r}'' | \, \mathrm{e}^{-\beta H} \, | \mathbf{r}' \rangle = \int_{\mathbf{r}(0)=\mathbf{r}'}^{\mathbf{r}(\beta)=\mathbf{r}''} \left[\mathrm{d}\mathbf{r}(t) \delta\left(1 - \mathbf{r}^2(t)\right)\right] \exp\left[-\tfrac{1}{2}\int_0^\beta \mathrm{d}t \, \dot{\mathbf{r}}^2(t)\right]. \tag{3.39}$$

Because the sphere has curvature for $N > 2$ 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 $\mathbf{r}''$:

$$\cos\theta = \mathbf{r}' \cdot \mathbf{r}'', \quad 0 \le \theta \le \pi .$$

We now 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 $(\mathbf{r}', \mathbf{r}'')$ space 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, $\mathbf{r}(t) \mapsto \boldsymbol{\rho}(t)$,

$$\mathbf{r}(t) = \mathbf{R}(t)\boldsymbol{\rho}(t).$$

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 $\boldsymbol{\rho}_{\mathrm{T}}$ the component in the orthogonal subspace. With this notation we find

$$\langle \mathbf{r}''|\, \mathrm{e}^{-\beta H}\, |\mathbf{r}'\rangle = \int_{\boldsymbol{\rho}(0)=\mathbf{r}'}^{\boldsymbol{\rho}(\beta)=\mathbf{r}'} \left[\mathrm{d}\boldsymbol{\rho}(t)\delta\left(1 - \boldsymbol{\rho}^2(t)\right)\right] \exp\left[-\mathcal{S}(\boldsymbol{\rho})\right] \tag{3.40}$$

with

$$\mathcal{S}(\boldsymbol{\rho}) = \frac{1}{2}\int_0^\beta \mathrm{d}t \left(\dot{\boldsymbol{\rho}}_{\mathrm{T}}^2 + \dot{u}^2 + \dot{v}^2 + \frac{\theta^2}{\beta^2}(u^2 + v^2) + 2\frac{\theta}{\beta}(\dot{v}u - \dot{u}v)\right), \tag{3.41}$$

and the constraint

$$u^2 + v^2 + \boldsymbol{\rho}_{\mathrm{T}}^2 = 1 . \tag{3.42}$$

In contrast with the abelian case where an exact calculation is possible, we can here perform only a small β (large temperature) expansion, corresponding to the WKB or semi-classical limit and valid for large quantum numbers. We take into account only fluctuations around the classical solution $u = 1$, $\boldsymbol{\rho}_{\mathrm{T}} = 0$, $v = 0$, neglecting contributions exponentially small in β^{-1}.

We thus eliminate the variable u from the action (3.41) by solving the constraint (3.42):

$$u = \left(1 - v^2 - \boldsymbol{\rho}_{\mathrm{T}}^2\right)^{1/2},$$

and expand the action in powers of $\boldsymbol{\rho}_{\mathrm{T}}$ and v. The leading order is

$$\mathcal{S}(\boldsymbol{\rho}) = \theta^2/2\beta ,$$

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 $\boldsymbol{\rho}_{\mathrm{T}}$ and v

$$\frac{1}{2}\int_0^\beta \mathrm{d}t \left[\dot{\boldsymbol{\rho}}_{\mathrm{T}}^2 - (\theta^2/\beta^2)\boldsymbol{\rho}_{\mathrm{T}}^2 + \dot{v}^2\right].$$

The integral over $v(t)$ is independent of θ and can be absorbed into the normalization. The integrals over the components of $\boldsymbol{\rho}_{\mathrm{T}}$ factorize and give identical results: the integral over $\boldsymbol{\rho}_{\mathrm{T}}$ 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 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

$$\langle \mathbf{r}''|\, \mathrm{e}^{-\beta H}\, |\mathbf{r}'\rangle \sim K(\beta)\, \mathrm{e}^{-\theta^2/2\beta} \left[\prod_{n>0} \left(1 - \frac{\theta^2}{n^2\pi^2}\right)\right]^{-(N-2)/2} .$$

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,

$$\langle \mathbf{r}''|\, \mathrm{e}^{-\beta H}\, |\mathbf{r}'\rangle \sim K(\beta) \left(\frac{\theta}{\sin\theta}\right)^{(N-2)/2} \mathrm{e}^{-\theta^2/2\beta} . \tag{3.43}$$

To extract the eigenvalues of H, we project this expression over the orthogonal polynomials $P_l^N(\cos\theta)$ associated with the $O(N)$ group:

$$\int_0^\pi \mathrm{d}\theta\, (\sin\theta)^{N-2} P_l^N(\cos\theta) P_{l'}^N(\cos\theta) = \delta_{ll'}\,. \tag{3.44}$$

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:

$$P_l^N(\cos\theta) = P_l^N(1) \left(1 - \frac{l(l+N-2)}{2(N-1)}\theta^2 + O\left(\theta^4\right)\right).$$

If we assume that to each value of l corresponds only one eigenvalue E_l of H then we can write

$$\begin{aligned} \mathrm{e}^{-\beta E_l} &\propto K(\beta) \int_0^\pi \mathrm{d}\theta\, P_l^N(\cos\theta)\, (\theta\sin\theta)^{(N-2)/2}\, \mathrm{e}^{-\theta^2/(2\beta)} \\ &= \mathrm{e}^{-\beta E_0} \left(1 - \tfrac{1}{2} l(l+N-2)\beta + O\left(\beta^2\right)\right), \end{aligned} \tag{3.45}$$

and, therefore,

$$E_l = E_0 + \tfrac{1}{2} l(l+N-2) + O(\beta). \tag{3.46}$$

Since E_l is independent of β we can infer from this calculation the exact result, up to an additive constant E_0.

The integral over $v(t)$ is independent of θ and can be absorbed into the normalization. The integrals over the components of $\boldsymbol{\rho}_{\mathrm{T}}$ factorize and give identical results: the integral over $\boldsymbol{\rho}_{\mathrm{T}}$ 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 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

$$\langle \mathbf{r}''|\, \mathrm{e}^{-\beta H}\, |\mathbf{r}'\rangle \sim K(\beta)\, \mathrm{e}^{-\theta^2/2\beta} \left[\prod_{n>0} \left(1 - \frac{\theta^2}{n^2\pi^2}\right)\right]^{-(N-2)/2} .$$

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,

$$\langle \mathbf{r}''|\, \mathrm{e}^{-\beta H}\, |\mathbf{r}'\rangle \sim K(\beta) \left(\frac{\theta}{\sin\theta}\right)^{(N-2)/2} \mathrm{e}^{-\theta^2/2\beta} . \tag{3.43}$$

To extract the eigenvalues of H, we project this expression over the orthogonal polynomials $P_l^N(\cos\theta)$ associated with the $O(N)$ group:

$$\int_0^\pi \mathrm{d}\theta\, (\sin\theta)^{N-2} P_l^N(\cos\theta) P_{l'}^N(\cos\theta) = \delta_{ll'}\,. \tag{3.44}$$

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:

$$P_l^N(\cos\theta) = P_l^N(1) \left(1 - \frac{l(l+N-2)}{2(N-1)}\theta^2 + O\left(\theta^4\right)\right).$$

If we assume that to each value of l corresponds only one eigenvalue E_l of H then we can write

$$\begin{aligned}
\mathrm{e}^{-\beta E_l} &\propto K(\beta) \int_0^\pi \mathrm{d}\theta\, P_l^N(\cos\theta)\, (\theta\sin\theta)^{(N-2)/2}\, \mathrm{e}^{-\theta^2/(2\beta)} \\
&= \mathrm{e}^{-\beta E_0} \left(1 - \tfrac{1}{2} l(l+N-2)\beta + O\left(\beta^2\right)\right),
\end{aligned} \tag{3.45}$$

and, therefore,

$$E_l = E_0 + \tfrac{1}{2} l(l+N-2) + O(\beta). \tag{3.46}$$

Since E_l is independent of β we can infer from this calculation the exact result, up to an additive constant E_0.

Concerning this calculation a comment is now in order: we have explained in Section 3.2 that the path integrals (3.38,3.39) are ill-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 discretizations of the path integral a subclass which corresponds to an $O(N)$ symmetric quantized hamiltonian. We shall see later, in the discussion of the non-linear σ model, that such a hamiltonian is fully determined up to an additive constant. The ambiguities of the quantization are here entirely contained in E_0.

Bibliographical Notes

Most of the references of Chapter 2 are still relevant here. The construction of path integrals in phase space for general hamiltonian systems can be found in

R.P. Feynman, *Phys. Rev.* 84 (1951) 108 (Appendix B); C. Garrod, *Rev. Mod. Phys.* 38 (1966) 483.

Several sections have been inspired by

L.D. Faddeev, in *Methods in Field Theory*, Les Houches School 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

The generalization to constrained systems is given in

L.D. Faddeev, *Theor. Math. Phys.* 1 (1969) 3.

An early discussion of the problem of quantization of velocity dependent potentials in the operator formalism can be found in

T.D. Lee and C.N. Yang, *Phys. Rev.* 128 (1962) 885.

APPENDIX A3
QUANTIZATION OF SPIN DEGREES OF FREEDOM, TOPOLOGICAL ACTIONS

In Section 3.3 we have evaluated the path integral in an example where space has non-trivial topological properties. We now want to discuss two other examples where topology plays an essential role, in the sense that an action generating the classical equations of motion cannot be globally defined. We first quantize in the path integral formalism angular momentum operators 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 speaks about topological action. The property that a topological action cannot be globally defined leads to the quantization of its amplitude, a property specific to quantum mechanics. Indeed, in classical mechanics, a multiplication of the action by a constant does not modify the equations of motion.

Note that in this appendix, to simplify the notation, we use a *real-time* formalism.

A3.1 Symplectic Form and Quantization: General Remarks

We first indicate how non-trivial topological properties of phase space, irrelevant from the point of view of classical mechanics, affect quantization. As an example, in the next section, we discuss the example of spin quantization.

In the hamiltonian formulation of classical mechanics the action has the form (3.8):

$$\mathcal{A}(p,q)=\int_{t'}^{t''}\mathrm{d}t\,\left[p(t)\dot q(t)-H\bigl(p(t),q(t),t\bigr)\right]. \tag{A3.1}$$

We note that the term $\int p\dot q\,\mathrm{d}t$ represents the area in phase space between the classical trajectory C and the axis $p=0$. It can be written as

$$\int_{\partial D}p(t)\dot q(t)\,\mathrm{d}t=\int_D \mathrm{d}p\wedge\mathrm{d}q\,, \tag{A3.2}$$

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 r.h.s. of the equation becomes

$$\int_D \mathrm{d}p\wedge\mathrm{d}q=\int_D \omega_{\alpha\beta}(u)\,\mathrm{d}u_\alpha\wedge\mathrm{d}u_\beta\,.$$

In the language of forms $\omega=\omega_{\alpha\beta}\,\mathrm{d}u_\alpha\wedge\mathrm{d}u_\beta$ is a two-form (see Section 1.4) which by construction is obtained by differentiating a one-form; here,

$$\omega=\mathrm{d}\omega', \tag{A3.3}$$

$$\omega'=p(u)\frac{\mathrm{d}q}{u_\alpha}\mathrm{d}u_\alpha=\omega_\alpha\mathrm{d}u_\alpha\,, \tag{A3.4}$$

and is called the *symplectic form*.

APPENDIX A3
QUANTIZATION OF SPIN DEGREES OF FREEDOM, TOPOLOGICAL ACTIONS

In Section 3.3 we have evaluated the path integral in an example where space has non-trivial topological properties. We now want to discuss two other examples where topology plays an essential role, in the sense that an action generating the classical equations of motion cannot be globally defined. We first quantize in the path integral formalism angular momentum operators 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 speaks about topological action. The property that a topological action cannot be globally defined leads to the quantization of its amplitude, a property specific to quantum mechanics. Indeed, in classical mechanics, a multiplication of the action by a constant does not modify the equations of motion.

Note that in this appendix, to simplify the notation, we use a *real-time* formalism.

A3.1 Symplectic Form and Quantization: General Remarks

We first indicate how non-trivial topological properties of phase space, irrelevant from the point of view of classical mechanics, affect quantization. As an example, in the next section, we discuss the example of spin quantization.

In the hamiltonian formulation of classical mechanics the action has the form (3.8):

$$\mathcal{A}(p,q)=\int_{t'}^{t''}\mathrm{d}t\,\left[p(t)\dot{q}(t)-H\bigl(p(t),q(t),t\bigr)\right]. \tag{A3.1}$$

We note that the term $\int p\dot{q}\,\mathrm{d}t$ represents the area in phase space between the classical trajectory C and the axis $p=0$. It can be written as

$$\int_{\partial D} p(t)\dot{q}(t)\,\mathrm{d}t=\int_D \mathrm{d}p\wedge\mathrm{d}q\,, \tag{A3.2}$$

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 r.h.s. of the equation becomes

$$\int_D \mathrm{d}p\wedge\mathrm{d}q=\int_D \omega_{\alpha\beta}(u)\,\mathrm{d}u_\alpha\wedge\mathrm{d}u_\beta\,.$$

In the language of forms $\omega=\omega_{\alpha\beta}\,\mathrm{d}u_\alpha\wedge\mathrm{d}u_\beta$ is a two-form (see Section 1.4) which by construction is obtained by differentiating a one-form; here,

$$\omega=\mathrm{d}\omega'\,, \tag{A3.3}$$

$$\omega'=p(u)\frac{\mathrm{d}q}{u_\alpha}\mathrm{d}u_\alpha=\omega_\alpha\mathrm{d}u_\alpha\,, \tag{A3.4}$$

and is called the *symplectic form*.

Since the operator d acting on differential forms is nilpotent, the form ω is closed, that is, it satisfies

$$\mathrm{d}\omega = 0 . \tag{A3.5}$$

Example. In Section 5.1, we will discuss the holomorphic formalism. We will introduce a complex parametrization of phase space (equation (5.36)):

$$p - iq = -i\sqrt{2}z , \qquad p + iq = i\sqrt{2}\bar{z} .$$

In terms of $z, \bar{z}$, the symplectic form becomes

$$\mathrm{d}p \wedge \mathrm{d}q = \frac{1}{i}\mathrm{d}z \wedge \mathrm{d}\bar{z} .$$

Previous 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

$$\mathcal{A}(u) = \int_D \omega - \int_{\partial D} \mathrm{d}t\, H\bigl(u(t), t\bigr), \tag{A3.6}$$

where ω, a symplectic form, is a *closed* two-form:

$$\omega = \omega_{\alpha\beta}\mathrm{d}u_\alpha \wedge \mathrm{d}u_\beta , \qquad \mathrm{d}\omega = 0 .$$

This latter condition is a sufficient one for the equations of motion to depend only on the boundary ∂D of the domain D but not on the interior, as was obvious for the initial action (A3.1). They then take the form

$$\omega_{\alpha\beta}\dot{u}_\beta = \frac{\partial \mathcal{H}}{\partial u_\alpha} . \tag{A3.7}$$

Locally, equation (A3.5) can be integrated in the same way that $\mathrm{d}p\wedge\mathrm{d}q$ can be integrated into $p\,\mathrm{d}q$. 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 quantum mechanics since the path integral involves the action explicitly in the form $\mathrm{e}^{i\mathcal{A}/\hbar}$. For the path integral to make sense, this phase factor must be independent of the choice of the action $\mathcal{A}$. Let us examine this problem in the example of the quantization of spin degrees of freedom.

A3.2 Spin Dynamics and Quantization

A3.2.1 Classical spin dynamics

We consider a vector $\mathbf{S}$ in three dimensions of fixed length s:

$$\mathbf{S}^2 = s^2 .$$

The simplest dynamics we can write for $\mathbf{S}$ is

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}t} = \mathbf{H} \times \mathbf{S} , \tag{A3.8}$$

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. Phase space is, therefore, isomorphic to the sphere S_2. The hamiltonian is simply

$$\mathcal{H} = \mathbf{H} \cdot \mathbf{S} . \tag{A3.9}$$

More precisely, it can be verified that if we parametrize $\mathbf{S}$ as

$$\mathbf{S} = s(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta),$$

then the action $\mathcal{A}$,

$$\mathcal{A}(\theta,\varphi) = \int [s\cos\theta\dot{\varphi} - \mathcal{H}]\,\mathrm{d}t\,, \tag{A3.10}$$

generates equation ($A3.8$). We see in this expression that $\cos\theta$ and φ play the role of conjugate variables. However, two remarks are in order: first we have integrated the symplectic two-form but, for this purpose, we have been forced to use 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 two-form cannot be integrated globally into a one-form because the integral $\int \mathrm{d}\varphi\,\mathrm{d}\cos\theta$, which is the area on the sphere, is defined only $(\mathrm{mod}\, 4\pi)$.

In classical mechanics, of course, these properties are irrelevant since only equations of motion are physical.

Note that the symplectic form has other useful representations:

$$\mathrm{d}\cos\theta \wedge \mathrm{d}\varphi = \tfrac{1}{2} s^{-3} \varepsilon_{ijk} S_i \mathrm{d}S_j \wedge \mathrm{d}S_k = 2i\mathrm{d}z_i \wedge \mathrm{d}\bar{z}_i\,,$$

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 Chapter 15):

$$\mathbf{S} = s\bar{z}_i \boldsymbol{\sigma} z_i\,, \qquad \bar{z}_i z_i = 1\,,$$

where $\boldsymbol{\sigma}$ is the set of Pauli matrices. In a special *gauge* the vector $\mathbf{z}$ can also be written as

$$\begin{aligned} z_1 &= \mathrm{e}^{i\varphi/2}\cos(\theta/2), \\ z_2 &= \mathrm{e}^{-i\varphi/2}\sin(\theta/2). \end{aligned}$$

A3.2.2 Quantization of spin degrees of freedom

We consider the path integral representation of the corresponding evolution operator in quantum mechanics. The action itself now appears explicitly and the problem discussed in the preceding section becomes relevant. The path integral exists only if the integrand $\mathrm{e}^{i\mathcal{A}/\hbar}$ is defined. Since the total area is defined only $(\mathrm{mod}\, 4\pi)$ 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 quantum mechanics: 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 provides another example of such a situation.

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. Phase space is, therefore, isomorphic to the sphere S_2. The hamiltonian is simply

$$\mathcal{H} = \mathbf{H} \cdot \mathbf{S} \,. \tag{A3.9}$$

More precisely, it can be verified that if we parametrize $\mathbf{S}$ as

$$\mathbf{S} = s(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta),$$

then the action $\mathcal{A}$,

$$\mathcal{A}(\theta,\varphi) = \int [s\cos\theta\dot{\varphi} - \mathcal{H}]\,\mathrm{d}t\,, \tag{A3.10}$$

generates equation ($A3.8$). We see in this expression that $\cos\theta$ and φ play the role of conjugate variables. However, two remarks are in order: first we have integrated the symplectic two-form but, for this purpose, we have been forced to use 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 two-form cannot be integrated globally into a one-form because the integral $\int \mathrm{d}\varphi\,\mathrm{d}\cos\theta$, which is the area on the sphere, is defined only $(\mathrm{mod}\, 4\pi)$.

In classical mechanics, of course, these properties are irrelevant since only equations of motion are physical.

Note that the symplectic form has other useful representations:

$$\mathrm{d}\cos\theta \wedge \mathrm{d}\varphi = \tfrac{1}{2}s^{-3}\varepsilon_{ijk}S_i\mathrm{d}S_j \wedge \mathrm{d}S_k = 2i\mathrm{d}z_i \wedge \mathrm{d}\bar{z}_i\,,$$

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 Chapter 15):

$$\mathbf{S} = s\bar{z}_i\boldsymbol{\sigma}z_i\,, \qquad \bar{z}_i z_i = 1\,,$$

where $\boldsymbol{\sigma}$ is the set of Pauli matrices. In a special *gauge* the vector $\mathbf{z}$ can also be written as

$$\begin{aligned} z_1 &= \mathrm{e}^{i\varphi/2}\cos(\theta/2), \\ z_2 &= \mathrm{e}^{-i\varphi/2}\sin(\theta/2). \end{aligned}$$

A3.2.2 Quantization of spin degrees of freedom

We consider the path integral representation of the corresponding evolution operator in quantum mechanics. The action itself now appears explicitly and the problem discussed in the preceding section becomes relevant. The path integral exists only if the integrand $\mathrm{e}^{i\mathcal{A}/\hbar}$ is defined. Since the total area is defined only $(\mathrm{mod}\,4\pi)$ 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 quantum mechanics: 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 provides another example of such a situation.

In the parametrization ($A3.10$) the action can be written as (we now set $\hbar = 1$)

$$\mathcal{A} = \int [(\gamma + \cos\theta)\dot{\varphi} - \mathcal{H}]\, \mathrm{d}t\,, \qquad (A3.11)$$

an expression that differs from ($A3.10$) only by a total derivative: by choosing $\gamma = 1/2, 0$ for s half-integer, integer, respectively, one renders $\mathrm{e}^{i\mathcal{A}}$ regular near $\theta = 0, \pi$.

To relate this action to the usual operator formulation of the angular momentum relations we first note that, classically, we have

$$S_{\pm} = \mathrm{e}^{\pm i\varphi}\left(s^2 - S_z^2\right)^{1/2}.$$

After quantization S_z becomes p_φ the conjugate momentum of the angular variables φ:

$$S_z \equiv p_\varphi = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}\varphi}.$$

It can be verified that the quantum operator $S_{\pm}$ can be written as

$$S_{\pm} = \mathrm{e}^{\pm i\varphi/2}\left(s^2 - p_\varphi^2\right)^{1/2} \mathrm{e}^{\pm i\varphi/2}.$$

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:

$$s = j + 1/2\,. \qquad (A3.12)$$

In particular, since s is quantized, we recover a property of quantum mechanics: quantization of spin.

If we call m the eigenvalues of S_z we observe that in the φ configuration space the corresponding eigenvectors have the form $\mathrm{e}^{im\varphi}$ and the projector K on the basis is

$$K(\varphi'', \varphi') = \sum_{m=-j}^{m=j} \mathrm{e}^{im(\varphi'-\varphi'')} = \frac{\sin\left[(j+1/2)\left(\varphi' - \varphi''\right)\right]}{\sin\left[\left(\varphi' - \varphi''\right)/2\right]}. \qquad (A3.13)$$

We can compare this expression with the short-time path integral representation which leads to

$$\begin{aligned} K(\varphi'', \varphi') &\propto \sum_n \int_{-s}^{s} \mathrm{d}p_\varphi\, \mathrm{e}^{ip_\varphi(\varphi'-\varphi''+2n\pi)}, \\ &\propto \sin\left[s\left(\varphi' - \varphi''\right)\right] \sum_n \frac{2(-1)^n}{(\varphi' - \varphi'' + 2n\pi)}. \end{aligned}$$

The sum over n has to be regularized but the factor $\sin s(\varphi' - \varphi'')$ is consistent with ($A3.13$) and the identification ($A3.12$).

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 $s \to \infty$), ambiguities due to operator ordering lead to difficulties in explicit calculations.

A3.3 The Magnetic Monopole

Electromagnetism provides another example of the situation encountered in Section A3.2.1. 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, as above, generally written as the integral of a two-form (involving the magnetic field) since this form is closed:

$$\mathcal{A}_{\text{mag.}} = e \int F_{ij}\, \mathrm{d}x_i \wedge \mathrm{d}x_j \,, \quad F_{ij} = \varepsilon_{ijk} B_k \,.$$

Because the two-form F is closed and one can integrate it locally introducing a vector potential, which is a one-form (see (3.18)),

$$F_{ij} = \partial_i A_j - \partial_j A_i \quad \Rightarrow \mathcal{A}_{\text{mag.}} = e \int \mathbf{A}(\mathbf{x}) \cdot \mathrm{d}\mathbf{x}.$$

However, if the two-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 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:

$$\int F_{ij} \mathrm{d}x_i \wedge \mathrm{d}x_j = g \int \frac{r}{r^3} \times r^2 \mathrm{d}\cos\theta \, \mathrm{d}\varphi = 4\pi g \,.$$

If the vector potential could be globally defined the integral would obviously vanish. More directly, if we try to calculate the corresponding vector potential we find in a family of gauges

$$A_i(x) = g \epsilon_{ijk} n_k x_j \frac{\mathbf{n} \cdot \mathbf{x}}{r\left(r^2 - (\mathbf{n} \cdot \mathbf{x})^2\right)} \,,$$

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 unphysical and can be displaced but not removed.

Again this property has no classical consequences. However, in quantum mechanics since the classical action can be defined only $(\bmod\, 4\pi eg)$ the weight factor $\mathrm{e}^{i\mathcal{A}}$ is only defined if

$$4\pi eg = 0 \quad (\bmod\ 2\pi) \ \Rightarrow 2eg = \text{ integer} \,,$$

which we recognize as Dirac's quantization condition.

Note that when this condition is fulfilled, parallel transport (see Section 18.3.1) is globally defined in $\mathbb{R}^3$.

4 STOCHASTIC DIFFERENTIAL EQUATIONS: LANGEVIN, FOKKER–PLANCK EQUATIONS

In Chapters 2,3 we have shown how the concept of path integral naturally emerges in the calculation of the matrix elements of the quantum statistical operator $e^{-\beta H}$ when the hamiltonian H is local. We have noted that, when hamiltonians are even, quadratic, functions of momenta, the integrand in the path integral defines a positive measure and, therefore, has a probabilistic interpretation. In this chapter, we discuss Langevin equations, that is, stochastic differential equations related to diffusion processes, brownian motion or random walk. From the Langevin equation we derive the Fokker–Planck (FP) equation for the probability distribution of the stochastic variables. The FP equation has a form analogous to the equation for the statistical operator in a magnetic field we have studied in Section 3.2 (but the corresponding hamiltonian, in general, is non-hermitian). We then show that averaged observables can also be calculated from path integrals, whose integrands define automatically positive measures. 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. Time discretization provides a solution to the problem.

This chapter is also meant to serve as an introduction to Chapters 17 and 36 in which *stochastic quantization* and *critical dynamics* are discussed.

4.1 The Langevin Equation

We call Langevin equation a first order in time stochastic differential equation of the form ($\dot{q} \equiv \mathrm{d}q/\mathrm{d}t$)

$$\dot{q}_i(t) = -\tfrac{1}{2} f_i\big(\mathbf{q}(t)\big) + \nu_i(t), \tag{4.1}$$

in which $\mathbf{q}(t)$ is a trajectory in $\mathbb{R}^d$, $f_i(\mathbf{q})$ a differentiable function of $\mathbf{q}$ and $\boldsymbol{\nu}(t)$ a set of stochastic functions called hereafter the "noise". The noise has a functional probability distribution $[d\rho(\nu)]$. In what follows, we specialize to a gaussian noise with a probability distribution of the form

$$[\mathrm{d}\rho(\nu)] = [\mathrm{d}\nu] \exp\left[-\frac{1}{2\Omega}\int \mathrm{d}t \sum_i \nu_i^2(t)\right]. \tag{4.2}$$

This particular form of the gaussian noise, called gaussian white noise, is related to Markov's processes (see Appendix A4). The constant Ω characterizes the width of the noise distribution. Note that by a rescaling of time $t \mapsto \Omega t$ (and a redefinition of the noise variable) it can be transferred in front of the driving term $f \mapsto f/\Omega$.

Alternatively the gaussian noise can be characterized by its one- and two-point correlation functions:

$$\langle \nu_i(t)\rangle = 0\,, \qquad \langle \nu_i(t)\nu_j(t')\rangle = \Omega\,\delta_{ij}\,\delta(t-t'), \tag{4.3}$$

as one verifies immediately using the results of Chapter 2. Equation (4.1) is not the most general stochastic first order differential equation. First, we have explicitly assumed *time translation invariance*. But even with this restriction more general equations can be considered, which we discuss in Section 4.8, because additional problems then arise.

Given the value of $\mathbf{q}(t)$ at initial time t_0, $\mathbf{q}(t_0) = \mathbf{q}_0$, the Langevin equation generates a time-dependent probability distribution $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ for the stochastic vector $\mathbf{q}(t)$ which can be formally written as

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = \left\langle \textstyle\prod_{i=1}^{d} \delta\left[q_i(t) - q_i\right] \right\rangle_\nu, \quad t \ge t_0 . \tag{4.4}$$

In equation (4.4), as in equation (4.3), brackets mean average over the noise. The vector $\mathbf{q}$ is the argument of $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ and has no relation with the function $\mathbf{q}(t)$. If $\mathcal{O}(\mathbf{q})$ is an arbitrary function, the definition (4.4) then implies

$$\int P(\mathbf{q}, t; \mathbf{q}_0, t_0)\mathcal{O}(\mathbf{q})\mathrm{d}q = \left\langle \mathcal{O}\big(\mathbf{q}(t)\big)\right\rangle_\nu . \tag{4.5}$$

Because the Langevin equation is local in time, and the noise at different times is uncorrelated, the distribution P satisfies a Markov property, analogous to the semi-group property (2.4) of the solution of equation (2.5):

$$P(\mathbf{q}_3, t_3; \mathbf{q}_1, t_1) = \int \mathrm{d}^d q_2 \, P(\mathbf{q}_3, t_3; \mathbf{q}_2, t_2) P(\mathbf{q}_2, t_2; \mathbf{q}_1, t_1), \quad t_1 \le t_2 \le t_3 . \tag{4.6}$$

Therefore, the distribution P is completely determined from its knowledge for small time intervals. It is convenient to to introduce the quantum notation of bras and kets (not to be confused with the notation $\langle \bullet \rangle$ meaning average over the noise ν):

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) \equiv \langle \mathbf{q} | \, \mathbf{P}(t, t_0) \, | \mathbf{q}_0 \rangle . \tag{4.7}$$

Time translation invariance and the relation (4.6) then imply the existence of an operator H, analogous to a quantum hamiltonian, such that the operator $\mathbf{P}(t, t_0)$ can be written as

$$\mathbf{P}(t, t_0) = \mathrm{e}^{-(t-t_0)H} , \tag{4.8}$$

which we call the Fokker–Planck hamiltonian.

Notation. In what follows we will often omit the dependence on the initial data and use instead of $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ the simplified notation $P(\mathbf{q}, t)$.

4.2 A Simple Example: The Linear Langevin Equation

A linear Langevin equation with gaussian white noise can be written as

$$\dot{q} = -\omega q + \nu(t), \tag{4.9}$$

in which $\nu(t)$ is the gaussian noise of equation (4.3), with $\Omega = 1$. In this case the Langevin equation (4.9) can be solved explicitly. With the boundary condition $q(0) = q_0$ the solution is

$$q(t) = q_0 \, \mathrm{e}^{-\omega t} + \int_0^t \mathrm{e}^{-\omega(t-t')} \, \nu(t') \mathrm{d}t' . \tag{4.10}$$

Random walk. In the special case $\omega = 0$ the equation describes a simple random walk in the continuum. One finds $\langle q(t) \rangle_\nu = q_0$ and $\left\langle (q(t) - q_0)^2 \right\rangle_\nu = t$.

The example exhibits an important property, also relevant for the general case. From

$$q(t+\varepsilon)-q(t)=\int_t^{t+\varepsilon}\nu(\tau)\mathrm{d}\tau\,,$$

we obtain, after averaging over the noise,

$$\left\langle [q(t+\varepsilon)-q(t)]^2\right\rangle_\nu=\varepsilon\,.$$

For $\varepsilon\to 0$ this result implies that typical paths are not differentiable, but only continuous or more precisely,

$$|q(t+\varepsilon)-q(t)|\underset{\varepsilon\to 0}{=}O(\sqrt{\varepsilon}),\tag{4.11}$$

a result which generalizes to any Langevin equation of the form (4.1). We note that this is also the class of paths which 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 case. The average of equation (4.10) over the noise yields

$$\langle q(t)\rangle_\nu=q_0\,\mathrm{e}^{-\omega t}\,.\tag{4.12}$$

Using (4.10) the calculation of the second moment of the distribution is also simple:

$$\left\langle [q(t)-\langle q(t)\rangle]^2\right\rangle_\nu=\frac{1}{2\omega}\left(1-\mathrm{e}^{-2\omega t}\right).\tag{4.13}$$

The expressions (4.12,4.13) immediately show that for $\omega<0$ the distribution of $q(t)$ has no limit. Moreover, for $q_0\neq 0$, $\langle q(t)\rangle_\nu$ grows exponentially with time. For $\omega>0$, instead, both moments have a finite large time limit.

Since $q(t)$ is linearly related to $\nu(t)$, it has a gaussian distribution $P(q,t)$ which is characterized by its two first moments $\langle q(t)\rangle$ and $\langle q^2(t)\rangle$. We thus find

$$P(q,t)=\left[\frac{\pi}{\omega}\left(1-\mathrm{e}^{-2\omega t}\right)\right]^{-1/2}\exp\left[-\frac{\omega}{(1-\mathrm{e}^{-2\omega t})}\left(q-q_0\,\mathrm{e}^{-\omega t}\right)^2\right].\tag{4.14}$$

This expression is analogous to expression (2.30). For $\omega>0$ the asymptotic distribution $P(q,t)$ for $t\to\infty$, called the ***equilibrium distribution***, is thus

$$P(q,t)\underset{t\to+\infty}{\to}\sqrt{\omega/\pi}\,\mathrm{e}^{-\omega q^2}\,.\tag{4.15}$$

Remark. In Section 4.3, we derive a general equation, the FP equation, satisfied by the probability distribution $P(q,t)$. To illustrate the difficulty one may encounter in the derivation of the equation we calculate the $q(t)$ two-point function for $q_0=0$:

$$\langle q(t_1)q(t_2)\rangle_\nu=\frac{1}{2\omega}\left(\mathrm{e}^{-\omega|t_2-t_1|}-\mathrm{e}^{-\omega(t_1+t_2)}\right).\tag{4.16}$$

Differentiating with respect to t_1, we find

$$\langle \dot q(t_1)q(t_2)\rangle_\nu=-\tfrac{1}{2}\left(\epsilon(t_1-t_2)\,\mathrm{e}^{-\omega|t_2-t_1|}-\mathrm{e}^{-\omega(t_1+t_2)}\right),\tag{4.17}$$

where $\epsilon(t)$ is the sign function. The limit $t_1 \to t_2$ is clearly ill-defined (a similar problem has been encountered in Section 3.2) because $\epsilon(0)$ appears in the r.h.s. If, instead, we take this limit in expression (4.16) first and then differentiate we find

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle q^2(t)\right\rangle_\nu = \frac{1}{2}\,\mathrm{e}^{-2\omega t}.$$

The difficulty is related to a property of the Langevin equation (4.1): $\mathrm{d}\left\langle \mathbf{q}^2(t)\right\rangle/\mathrm{d}t$ is well-defined but $2\left\langle \mathbf{q}(t)\cdot\dot{\mathbf{q}}(t)\right\rangle$ is not. Due to the singular nature of the noise two-point correlation function, time differentiation and averaging are two operations which do not commute in general. Note, however, that the choice $\epsilon(0)=0$ in equation (4.17), that is, taking the half sum of the derivative from above and below, ensures compatibility between the different results and thus the commutation of averaging and time differentiation.

4.3 The Fokker–Planck Equation

We now show that the equations (4.1,4.2) imply a differential equation for $P(\mathbf{q},t)$. To avoid the formal problems associated with averages of the form $\langle \dot{q}(t)q(t)\rangle$ (equation (4.17)) we first integrate the Langevin equation (4.1) in an infinitesimal time interval between time t and $t+\varepsilon$:

$$q_i(t+\varepsilon) = q_i(t) - \tfrac{1}{2}\varepsilon f_i(\mathbf{q}(t)) + \int_t^{t+\varepsilon} \nu_i(\tau)\mathrm{d}\tau + O(\varepsilon^{3/2}), \tag{4.18}$$

where the evaluation of the error follows from the estimate (4.11). By averaging over the noise between the times t and $t+\varepsilon$, one obtains the probability distribution of $\mathbf{q}=\mathbf{q}(t+\varepsilon)$, at fixed $\mathbf{q}'=\mathbf{q}(t)$, that is, $P(\mathbf{q},t+\varepsilon;\mathbf{q}',t)$

$$P(\mathbf{q},t+\varepsilon;\mathbf{q}',t) = \left\langle \delta\big(\mathbf{q}-\mathbf{q}(t+\varepsilon)\big)\right\rangle_\nu .$$

Discretized Langevin equation. At order ε the continuum Langevin equation is equivalent to the discretized Langevin equation:

$$q_i(t+\varepsilon) = q_i(t) - \tfrac{1}{2}\varepsilon f_i(\mathbf{q}(t)) + \sqrt{\varepsilon}\,\bar{\nu}_i(t). \tag{4.19}$$

The noise $\bar{\nu}(t)$ is linearly related to $\nu(t)$

$$\int_t^{t+\varepsilon} \nu_i(\tau)\mathrm{d}\tau = \sqrt{\varepsilon}\,\bar{\nu}_i(t), \tag{4.20}$$

and, therefore, also has a gaussian distribution defined for instance by $((t-t')/\varepsilon\in\mathbb{Z})$

$$\langle\bar{\nu}_i(t)\rangle = 0\,,\qquad \langle\bar{\nu}_i(t)\bar{\nu}_j(t')\rangle = \Omega\,\delta_{ij}\,\delta_{tt'},$$

where $\delta_{tt'}$ is a Kronecker δ.

Fokker–Planck equation. For $\varepsilon\to 0$ the Fourier transform $\tilde{P}$ of P with respect to $\mathbf{q}$ is given by

$$\begin{aligned}\tilde{P}(\mathbf{p},t+\varepsilon;\mathbf{q}',t) &= \int \mathrm{d}^d q\; \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}}\, P(\mathbf{q},t+\varepsilon;\mathbf{q}',t) = \left\langle \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}(t+\varepsilon)}\right\rangle_\nu \\ &= \exp\left[-i\mathbf{p}\cdot\left(\mathbf{q}'-\varepsilon\mathbf{f}(\mathbf{q}')/2\right)\right]\left\langle \exp\left[-i\mathbf{p}\cdot\int_t^{t+\varepsilon}\boldsymbol{\nu}(\tau)\mathrm{d}\tau\right]\right\rangle_\nu .\end{aligned}$$

The integral over $\nu(\tau)$ is then gaussian and can be performed explicitly:

$$\tilde{P}(\mathbf{p}, t+\varepsilon; \mathbf{q}', t) = \mathrm{e}^{-\varepsilon\Omega\mathbf{p}^2/2 + i\varepsilon\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')/2}\,\mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}'}\,. \tag{4.21}$$

According to the discussion of Section 4.1, the coefficient of ε in the expansion of the expression for $\varepsilon \to 0$ is a matrix element of the FP hamiltonian (equation (4.8)) in the mixed $\mathbf{p}, \mathbf{q}$ representation,

$$\begin{aligned}\tilde{P}(\mathbf{p}, t+\varepsilon; \mathbf{q}', t) &= \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}'}\left[1 - \tfrac{1}{2}\varepsilon\big(\Omega\mathbf{p}^2 - i\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')\big)\right] + O(\varepsilon^2) \\ &= \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}'} - \varepsilon H(\mathbf{p}, \mathbf{q}') + O(\varepsilon^2).\end{aligned}$$

Inverting the Fourier transformation we find that H is a second-order differential operator:

$$\begin{aligned}\langle\mathbf{q}|\,H\,|\mathbf{q}'\rangle &= \frac{1}{2}\int \frac{\mathrm{d}^d p}{(2\pi)^d}\,\mathrm{e}^{i\mathbf{p}\cdot(\mathbf{q}-\mathbf{q}')}\big(\Omega\mathbf{p}^2 - i\mathbf{p}\cdot\mathbf{f}(\mathbf{q}')\big) \\ &= \tfrac{1}{2}\big(-\Omega\nabla_{\mathbf{q}}^2 - \nabla_{\mathbf{q}}\cdot\mathbf{f}(\mathbf{q})\big)\delta^{(d)}(\mathbf{q}-\mathbf{q}'),\end{aligned}$$

or using the operator notation of quantum mechanics (Section 2.2, with $\hbar = 1$),

$$H = \tfrac{1}{2}\left(\Omega\hat{\mathbf{p}}^2 - i\hat{\mathbf{p}}\cdot\mathbf{f}(\hat{\mathbf{q}})\right). \tag{4.22}$$

Since, from equation (4.8),

$$\dot{\mathbf{P}}(t) = -H\mathbf{P}(t)\,,$$

the probability distribution $P(\mathbf{q}, t)$ (equation (4.4)) at finite time $t > t_0$, satisfies the evolution equation,

$$\dot{P}(\mathbf{q}, t) = \tfrac{1}{2}\partial_i\big(\Omega\partial_i P(\mathbf{q}, t) + f_i(\mathbf{q})P(\mathbf{q}, t)\big). \tag{4.23}$$

This is the FP equation associated with the Langevin equation (4.1) and noise (4.2). Integrating over space we verify that the form of the FP equation automatically ensures that probability is conserved:

$$\frac{\partial}{\partial t}\int \mathrm{d}^d q\, P(\mathbf{q}, t) = 0\,.$$

We have, therefore, established a formal relation between stochastic differential equations and euclidean quantum mechanics. All observables which can be calculated from the Langevin equation by averaging over the noise, can be recovered by methods of quantum mechanics using the FP hamiltonian H.

Finally, inverting the Fourier transformation we infer from equation (4.21)

$$P(\mathbf{q}, t+\varepsilon; \mathbf{q}', t) \sim (2\pi\Omega)^{-d/2}\exp\left[-\frac{1}{2\varepsilon\Omega}\big(\mathbf{q} - \mathbf{q}' + \tfrac{1}{2}\varepsilon\mathbf{f}(\mathbf{q}')\big)^2\right]. \tag{4.24}$$

The form (4.24) of the distribution, valid for $\varepsilon \to 0$, will be used to derive a path integral representation.

4.4 Equilibrium Distribution. Correlation Functions

The methods of quantum mechanics can now be used to study several interesting questions, as, for example, the existence of an equilibrium distribution. We have to be careful that the FP hamiltonian is not hermitian and, therefore, its left and right eigenvectors do not coincide. Note that if the system is not ergodic the arguments that follow can be generalized to each connected component and the configuration space is decomposed into disconnected components, a situation we will encounter in Quantum Field Theory.

The equilibrium distribution $P_0(\mathbf{q})$ is the large time limit, if it exists, of $P(\mathbf{q}, t)$:

$$P_0(\mathbf{q}) = \lim_{t \to +\infty} P(\mathbf{q}, t). \tag{4.25}$$

This implies that $P_0(\mathbf{q})$ is a time-independent solution of the FP equation (4.23), that is, a right eigenvector of H with eigenvalue 0 and that it is positive and normalizable:

$$\int P_0(\mathbf{q}) \mathrm{d}\mathbf{q} < \infty \,. \tag{4.26}$$

We now note that conservation of probability has imposed the special form of the FP hamiltonian (4.22), that is, the factorization on the left of the differential operators $\partial / \partial q_i$. Therefore, a constant is a candidate for being a left eigenvector with eigenvalue zero of the hamiltonian H. Introducing the bra and ket notation we can write

$$\langle 0 | H = 0 \quad \text{with} \quad \langle 0 | \mathbf{q} \rangle = 1 \,.$$

With the same notation we can write

$$H \, |0\rangle = 0 \quad \text{with} \quad \langle \mathbf{q} | 0 \rangle = P_0(\mathbf{q}) \,.$$

The condition (4.26) is thus the condition that the vector $|0\rangle$ has a finite norm, a necessary condition for the eigenvalue zero to really belong to the spectrum

$$\langle 0 | 0 \rangle = \int P_0(\mathbf{q}) \mathrm{d}\mathbf{q} < \infty \,.$$

Since one can show that the eigenvalues of H have a non-negative real part, the vector $|0\rangle$ is then the ground state of the hamiltonian.

If no such solution exists, all eigenvalues of H have strictly positive real parts and the Langevin equation (4.1) has runaway solutions in the sense that the probability of finding $\mathbf{q}(t)$ inside a ball of arbitrary but finite radius goes to zero at large time:

(i) algebraically, if the spectrum of H has a continuous part extending down to the origin (this is the case with the brownian motion, for example),

(ii) exponentially, with a rate which is the inverse of the real part of the ground state eigenvalue, otherwise. This rate is called the relaxation time. (For more details see Appendix A4).

Finally, we note that the sole knowledge of the equilibrium distribution,

$$P_0(q) = \mathrm{e}^{-E(q)}, \tag{4.27}$$

4.4 Equilibrium Distribution. Correlation Functions

The methods of quantum mechanics can now be used to study several interesting questions, as, for example, the existence of an equilibrium distribution. We have to be careful that the FP hamiltonian is not hermitian and, therefore, its left and right eigenvectors do not coincide. Note that if the system is not ergodic the arguments that follow can be generalized to each connected component and the configuration space is decomposed into disconnected components, a situation we will encounter in Quantum Field Theory.

The equilibrium distribution $P_0(\mathbf{q})$ is the large time limit, if it exists, of $P(\mathbf{q},t)$:

$$P_0(\mathbf{q}) = \lim_{t\to+\infty} P(\mathbf{q},t). \tag{4.25}$$

This implies that $P_0(\mathbf{q})$ is a time-independent solution of the FP equation (4.23), that is, a right eigenvector of H with eigenvalue 0 and that it is positive and normalizable:

$$\int P_0(\mathbf{q})\mathrm{d}\mathbf{q} < \infty\,. \tag{4.26}$$

We now note that conservation of probability has imposed the special form of the FP hamiltonian (4.22), that is, the factorization on the left of the differential operators $\partial/\partial q_i$. Therefore, a constant is a candidate for being a left eigenvector with eigenvalue zero of the hamiltonian H. Introducing the bra and ket notation we can write

$$\langle 0|\,H = 0 \quad \text{with} \quad \langle 0|\mathbf{q}\rangle = 1\,.$$

With the same notation we can write

$$H\,|0\rangle = 0 \quad \text{with} \quad \langle \mathbf{q}|0\rangle = P_0(\mathbf{q})\,.$$

The condition (4.26) is thus the condition that the vector $|0\rangle$ has a finite norm, a necessary condition for the eigenvalue zero to really belong to the spectrum

$$\langle 0|0\rangle = \int P_0(\mathbf{q})\mathrm{d}\mathbf{q} < \infty\,.$$

Since one can show that the eigenvalues of H have a non-negative real part, the vector $|0\rangle$ is then the ground state of the hamiltonian.

If no such solution exists, all eigenvalues of H have strictly positive real parts and the Langevin equation (4.1) has runaway solutions in the sense that the probability of finding $\mathbf{q}(t)$ inside a ball of arbitrary but finite radius goes to zero at large time:

(i) algebraically, if the spectrum of H has a continuous part extending down to the origin (this is the case with the brownian motion, for example),

(ii) exponentially, with a rate which is the inverse of the real part of the ground state eigenvalue, otherwise. This rate is called the relaxation time. (For more details see Appendix A4).

Finally, we note that the sole knowledge of the equilibrium distribution,

$$P_0(q) = \mathrm{e}^{-E(q)}, \tag{4.27}$$

does not determine the functions $f_i(q)$ in equation (4.23) uniquely. Indeed, by demanding that $P_0(q)$ is a time-independent solution of the FP equation, we only obtain the condition

$$(\partial_i - \partial_i E)\,(\Omega \partial_i E - f_i) = 0\,, \tag{4.28}$$

which, setting

$$f_i(q) = \Omega \partial_i E(q) + V_i(q)\,\mathrm{e}^{E(q)}, \tag{4.29}$$

can be rewritten as a current conservation equation:

$$\partial_i V_i(q) = 0\,. \tag{4.30}$$

Correlation functions. Up to now 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 $Z^{(n)}(t_1, t_2, \ldots, t_n)$:

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \langle q(t_1) q(t_2) \ldots q(t_n) \rangle_\nu\,. \tag{4.31}$$

Assuming that the boundary conditions are $q(t_0) = q_0$ and ordering times $t_0 \le t_1 \le t_2 \le \cdots \le t_n$, we can use the definition of the FP hamiltonian to formally rewrite $Z^{(n)}$. 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

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \int \mathrm{d}q_n\, q_n \, \langle P(q_n, t_n; q(t_{n-1}), t_{n-1}) q(t_1) q(t_2) \ldots q(t_{n-1}) \rangle_\nu\,.$$

We then average over the noise for times $t_{n-2} < t \le t_{n-1}$, where now $q(t_{n-2})$ yields the boundary condition. We find

$$\begin{aligned} Z^{(n)}(t_1, t_2, \ldots, t_n) = \int & \mathrm{d}q_n \, \mathrm{d}q_{n-1}\, q_n P(q_n, t_n; q_{n-1}, t_{n-1}) q_{n-1} \\ & \times \langle P(q_{n-1}, t_{n-1}; q(t_{n-2}), t_{n-2}) q(t_1) q(t_2) \ldots q(t_{n-2}) \rangle_\nu\,. \end{aligned}$$

The procedure can be iterated until the averaging over the noise is completed. Finally, using the representation (4.8), and the arguments of Section 2.4.2, we can express $Z^{(n)}(t_1, \ldots, t_n)$ in terms of an operator expectation value:

$$\begin{aligned} Z^{(n)}(t_1, t_2, \ldots, t_n) &= \langle 0 |\, \hat{q}\, \mathrm{e}^{-(t_n - t_{n-1})H}\, \hat{q} \ldots \hat{q}\, \mathrm{e}^{-(t_2 - t_1)H}\, \hat{q}\, \mathrm{e}^{-(t_1 - t_0)H} \, | q_0 \rangle & (4.32) \\ &= \langle 0 |\, \hat{Q}(t_n) \ldots \hat{Q}(t_2) \hat{Q}(t_1)\, \mathrm{e}^{H t_0} \, | q_0 \rangle\,, & (4.33) \end{aligned}$$

where, to simplify notations, we have introduced $\hat{Q}(t)$, the operator $\hat{q}$ in the "Heisenberg" representation:

$$\hat{Q}(t) = \mathrm{e}^{Ht}\, \hat{q}\, \mathrm{e}^{-Ht}\,.$$

When an equilibrium distribution exists, in the limit $t_0 \to -\infty$ (boundary conditions in the far past) the correlation functions converge towards *equilibrium correlation functions*:

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \langle 0 |\, \hat{Q}(t_n) \ldots \hat{Q}(t_2) \hat{Q}(t_1) \, | 0 \rangle\,. \tag{4.34}$$

We recognize the analogue of the representation (2.40) of correlation functions as vacuum or ground state expectation values of product of operators in quantum mechanics, as discussed in Section 2.4. As shown in Section A6.1 these expressions can be symmetrized in time by introducing the time ordering operation.

Time evolution of observables. Going from the FP equation to expressions in terms of time-dependent operators, we have changed from a Schrödinger picture to a Heisenberg picture. In the Heisenberg picture, one directly writes equations for the evolution of time-dependent operators. Let us give an application here. In terms of Heisenberg operators the average of an observable $\mathcal{O}(q)$ at time t can be written as

$$\langle \mathcal{O}(q(t))\rangle_\nu = \langle 0|\, \mathcal{O}\big[\hat{Q}(t)\big]\, \mathrm{e}^{Ht_0}\, |q_0\rangle\,.$$

We can write an evolution equation for the operator $\mathcal{O}\big[\hat{Q}(t)\big]\,\mathrm{e}^{Ht_0}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{O}[\hat{Q}(t)]\,\mathrm{e}^{Ht_0} = [\,H\,,\; \mathcal{O}[\hat{Q}(t)]\,\mathrm{e}^{Ht_0}\,].$$

For the matrix elements $\mathcal{O}(q,t)$ of the operator

$$\mathcal{O}(q,t) \equiv \langle 0|\mathcal{O}[\hat{Q}(t)]\,\mathrm{e}^{Ht_0}\,|q\rangle,$$

this translates into the partial differential equation

$$\dot{\mathcal{O}}(q,t) = \frac{1}{2}\left[\Omega\frac{\partial}{\partial q} - f(q)\right]\frac{\partial}{\partial q}\mathcal{O}(q,t). \tag{4.35}$$

Then the averages $\langle \mathcal{O}(q(t))\rangle_\nu$ are simply obtained by integrating $\mathcal{O}(q,t)$ with the initial distribution at time t_0, here

$$\langle \mathcal{O}(q(t))\rangle_\nu = \int \mathrm{d}q\, \mathcal{O}(q,t)\delta(q-q_0). \tag{4.36}$$

4.5 A Special Case: The Dissipative Langevin Equation

In one special case the hamiltonian (4.22) is equivalent to a hermitian hamiltonian, when $f_i(\mathbf{q})$ is a gradient:

$$f_i(\mathbf{q}) = \Omega\,\partial_i E(\mathbf{q}). \tag{4.37}$$

This corresponds to the purely dissipative Langevin equation:

$$\dot{q}_i = -\tfrac{1}{2}\Omega\partial_i E(\mathbf{q}) + \nu_i(t). \tag{4.38}$$

The linear Langevin equation (4.9) provides the simplest example of such an equation

$$\omega q = \frac{\partial}{\partial q}\left(\frac{1}{2}\omega q^2\right).$$

Then, the transformation

$$P(\mathbf{q},t) = \mathrm{e}^{-E(\mathbf{q})/2}\,\langle \mathbf{q}\,|U(t,t_0)|\,\mathbf{q}_0\rangle\,\mathrm{e}^{E(\mathbf{q}_0)/2}, \tag{4.39}$$

introduced into equation (4.23), leads to

$$\partial_t \langle \mathbf{q} | U(t,t_0) | \mathbf{q}_0 \rangle = -\tfrac{1}{2}\Omega \left(-\partial_i + \tfrac{1}{2}\partial_i E\right) \left(\partial_i + \tfrac{1}{2}\partial_i E\right) \langle \mathbf{q} | U(t,t_0) | \mathbf{q}_0 \rangle . \tag{4.40}$$

In terms of the hamiltonian $\widetilde{H}$,

$$\widetilde{H} = \tfrac{1}{2}\Omega A_i^\dagger A_i \equiv \frac{\Omega}{2} \left[\hat{\mathbf{p}}^2 + \tfrac{1}{4}\left(\nabla E(\hat{\mathbf{q}})\right)^2 - \tfrac{1}{2}\Delta E(\hat{\mathbf{q}})\right], \tag{4.41}$$

where we have defined

$$\mathbf{A}(q, \partial/\partial q) \equiv \nabla_q + \tfrac{1}{2}\nabla E \equiv i\mathbf{p} + \tfrac{1}{2}\nabla E , \tag{4.42}$$

the operator $U(t,t_0)$ takes the form of a statistical operator $U(t,t_0) = \mathrm{e}^{-(t-t_0)\widetilde{H}}$.

We see that the hamiltonian $\widetilde{H}$ is positive. Moreover, if the wave function $\mathrm{e}^{-E(\mathbf{q})/2}$ is normalizable

$$\langle \mathbf{q}|0\rangle = \mathrm{e}^{-E(\mathbf{q})/2}, \qquad \langle 0|0\rangle = \int \mathrm{d}\mathbf{q}\, \mathrm{e}^{-E(\mathbf{q})} < \infty ,$$

it corresponds to the eigenvalue zero:

$$\mathbf{A}|0\rangle = 0 \quad \Rightarrow \quad \widetilde{H}|0\rangle = 0 ,$$

and thus is the ground state of $\widetilde{H}$. At large times the operator $\mathrm{e}^{-(t-t_0)\widetilde{H}}$ projects onto its ground state. The interpretation of this result in terms of the probability distribution $P(\mathbf{q},t)$ then is

$$\lim_{t\to\infty} P(\mathbf{q},t) = \mathrm{e}^{-E(\mathbf{q})/2}\, \mathrm{e}^{E(\mathbf{q}_0)/2} \langle \mathbf{q}|0\rangle\langle 0|\mathbf{q}_0\rangle = \mathrm{e}^{-E(\mathbf{q})} .$$

The distribution $P(\mathbf{q},t)$ converges at large time towards the equilibrium distribution $\mathrm{e}^{-E(\mathbf{q})}$.

If the wave function $\mathrm{e}^{-E(\mathbf{q})/2}$ is not normalizable, instead, there exists no equilibrium distribution; the Langevin equation (4.1) has only runaway solutions.

Remark. The case of the purely dissipative Langevin equation (4.38) corresponds to detailed balance for discrete processes (see Appendix A4.2). The driving term f_i is then called conservative. In the absence of noise, the Langevin equation reduces to a gradient flow:

$$\dot{q}_i(t) = -\tfrac{1}{2}\Omega\, \partial_i E\big(\mathbf{q}(t)\big) . \tag{4.43}$$

Taking the scalar product with the vector $\dot{\mathbf{q}}$ and integrating one obtains

$$\int_{t_0}^{t} \dot{\mathbf{q}}^2(t')\mathrm{d}t' = -\tfrac{1}{2}\Omega \left[E\big(\mathbf{q}(t)\big) - E\big(\mathbf{q}(t_0)\big)\right] . \tag{4.44}$$

Therefore, in the absence of noise, $E(\mathbf{q}(t))$ is a monotonically decreasing function of time.

The linear Langevin equation. After the transformation (4.39), the FP hamiltonian associated with equation (4.9) takes the form

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 q^2 - \tfrac{1}{2}\omega . \tag{4.45}$$

The eigenvalues ϵ_n of H are

$$\epsilon_n = \left(n + \tfrac{1}{2}\right)|\omega| - \tfrac{1}{2}\omega\,, \quad n \geq 0\,. \tag{4.46}$$

We see immediately that if ω is positive, ϵ_0 vanishes. Correspondingly, $\mathrm{e}^{-\omega q^2}$ is normalizable and is thus the equilibrium distribution.

If instead ω is negative,

$$\omega < 0\,, \qquad \epsilon_0 = -\omega\,,$$

the lowest eigenvalue is positive, $\mathrm{e}^{-\omega q^2}$ is not normalizable and for large time the distribution $P(q,t)$ becomes (taking $t_0 = 0$ for convenience)

$$P(q,t) \sim \mathrm{e}^{-|\omega|t}\,\mathrm{e}^{-|\omega|q_0^2} + O\left(\mathrm{e}^{-2|\omega|t}\right). \tag{4.47}$$

This 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 the simple 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 $1/\sqrt{t}$.

4.6 Path Integral Representation

Applying the method described in Chapter 2 to equation (4.23), we can derive path integral representations for the probability distribution $P(\mathbf{q}, t; \mathbf{q}_0, t_0)$ (equation (4.4)) and averaged observables. Here, it is more convenient to start directly from expression (4.24) and combine it with the semi-group property (4.6). We obtain

$$P(\mathbf{q}, t; \mathbf{q}_0, t_0) = \int_{\mathbf{q}(t_0)=\mathbf{q}_0}^{\mathbf{q}(t)=\mathbf{q}} [\mathrm{d}\mathbf{q}(\tau)] \exp\left[-\mathcal{S}(\mathbf{q})/\Omega\right] \tag{4.48}$$

with

$$\mathcal{S}(\mathbf{q}) = \lim_{\varepsilon \to 0} \sum_{k=1} \frac{1}{2\varepsilon}\left(\mathbf{q}_k - \mathbf{q}_{k-1} + \tfrac{1}{2}\varepsilon \mathbf{f}(\mathbf{q}_{k-1})\right)^2.$$

To study the small time step limit we expand in powers of ε up to order ε. Following the convention that we have adopted in Section 3.1 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,

$$\mathbf{q} - \mathbf{q}' + \tfrac{1}{2}\varepsilon \mathbf{f}(\mathbf{q}') = \mathbf{q} - \mathbf{q}' + \tfrac{1}{2}\varepsilon \mathbf{f}\left((\mathbf{q} + \mathbf{q}')/2\right) - \tfrac{1}{4}\varepsilon(q_i - q'_i)\frac{\partial \mathbf{f}}{\partial q_i} + O\left(\varepsilon^2\right).$$

To get rid of the last term, we can change variables $\mathbf{q} \mapsto \tilde{\mathbf{q}}$, setting

$$\mathbf{q} - \mathbf{q}' - \tfrac{1}{4}\varepsilon(q_i - q'_i)\frac{\partial \mathbf{f}}{\partial q_i} \mapsto \tilde{\mathbf{q}} - \mathbf{q}'\,,$$

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

$$\det\left[\frac{\partial}{\partial q_i}\left(q_j - q'_j - \tfrac{1}{4}\varepsilon(q_l - q'_l)\frac{\partial f_j}{\partial q_l}\right)\right]^{-1} = \exp\left(\tfrac{1}{4}\varepsilon \frac{\partial f_i}{\partial q_i}\right),$$

where the identity (1.101), $\ln\det = \mathrm{tr}\ln$, has been used. Collecting all terms of order ε we finally obtain the dynamic action:

$$\mathcal{S}(\mathbf{q}) = \int_{t_0}^{t} \tfrac{1}{2}\left[\left(\dot{\mathbf{q}} + \tfrac{1}{2}\mathbf{f}(\mathbf{q})\right)^2 - \tfrac{1}{2}\Omega\partial_i f_i(\mathbf{q})\right] \mathrm{d}\tau\,. \tag{4.49}$$

Remarks.

(i) Since the action contains a term linear in the time derivative, perturbative calculations will involve the ill-defined quantity $\epsilon(0)$. Consistency with the choice of symmetrizing the argument of $\mathbf{f}$ requires $\epsilon(0) = 0$.

(ii) In the case of the dissipative Langevin equation $f_i(\mathbf{q}) = \Omega\partial_i E(\mathbf{q})$, and, therefore,

$$\frac{1}{2}\int_{t_0}^{t} \mathrm{d}\tau\left(\dot{\mathbf{q}} + \tfrac{1}{2}\mathbf{f}(q)\right)^2 = \frac{1}{2}\int_{t_0}^{t} \mathrm{d}\tau\left[\dot{\mathbf{q}}^2 + \tfrac{1}{4}\Omega^2\left(\nabla E(\mathbf{q})\right)^2\right] + \tfrac{1}{2}\Omega\left[E(\mathbf{q}(t)) - E(\mathbf{q}(t_0))\right].$$

This result is consistent with the equations (4.39,4.41). It again requires $\epsilon(0) = 0$, in order that the derivative and the average commute.

(iii) Combining the representation (4.32) with the results of Section 2.4.2, we obtain a path integral representation of correlation functions:

$$Z^{(n)}_{i_1 i_2 \ldots i_n}(t_1, t_2, \ldots, t_n) = \int [\mathrm{d}q(\tau)] q_{i_1}(t_1) q_{i_2}(t_2) \ldots q_{i_n}(t_n)\, \mathrm{e}^{-\mathcal{S}(\mathbf{q})/\Omega} \tag{4.50}$$

with

$$\mathcal{S}(\mathbf{q}) = \int_{t_0}^{\infty} \tfrac{1}{2}\left[\left(\dot{\mathbf{q}} + \tfrac{1}{2}\mathbf{f}(\mathbf{q})\right)^2 - \tfrac{1}{2}\Omega\partial_i f_i(\mathbf{q})\right] \mathrm{d}\tau\,,$$

and the boundary condition is $\mathbf{q}(t_0) = \mathbf{q}_0$.

Perturbative expansion. The parameter Ω here plays the role of $\hbar$ in quantum mechanics. It orders, naturally, perturbation theory. We then note that, at leading order, the classical action reduces to

$$\mathcal{S}(\mathbf{q}) = \int \tfrac{1}{2}\left(\dot{\mathbf{q}} + \tfrac{1}{2}\mathbf{f}(\mathbf{q})\right)^2 \mathrm{d}t\,. \tag{4.51}$$

This means that in the particular example of the Langevin equation (4.38), perturbation theory has to be expanded around one of the extrema of $E(q)$. However, not all extrema are equivalent once one takes into account the first quantum correction due to the additional term $\Omega\partial_i f_i(q)$, as the example of Section 4.2 reveals (see equation (4.46)). Expression (4.46) indeed shows that the second term $\Delta E(q)$ lifts the degeneracy between minima and maxima and that only minima are suitable starting points for a perturbative expansion.

4.7 General Discretized Langevin Equation

Before dealing with a more general continuum Langevin equation we first discuss a discretized, markovian, time translation invariant Langevin equation in the small step size limit and in the simple case of a gaussian noise. This is a useful exercise in the sense that it clarifies a few subtle points that appear in the discussion of the continuum equation. We call ε the time step, which we will assume to be small.

Since the main application is the brownian motion on Riemannian manifolds, it is convenient to adopt the notation and conventions of Chapter 22, to which we refer for details, with, in particular, summation over repeated lower and upper indices for tensors.

A more general form of the discretized Langevin equation is

$$q^i(t+\varepsilon) = q^i(t) - \tfrac{1}{2}\varepsilon f^i\big(\mathbf{q}(t)\big) + e^i_a\big(\mathbf{q}(t)\big)\nu_a(t) + \tfrac{1}{2}d^i_{ab}\big(\mathbf{q}(t)\big)\nu_a(t)\nu_b(t) \tag{4.52}$$

with the gaussian noise still defined by

$$\langle \nu_a(t)\rangle = 0\,, \qquad \langle \nu_a(t)\nu_b(t')\rangle = \varepsilon\,\Omega\,\delta_{tt'}\delta_{ab}\,. \tag{4.53}$$

Note that we have expanded the r.h.s. of the Langevin equation only up to second order in $\boldsymbol{\nu}(t)$. The reasons are the following:

(i) the noise two-point function must be proportional to ε for dimensional reasons; the consistency of this choice will be checked in the derivation of the FP equation;

(ii) since $\nu_a(t)$ is of order $\sqrt{\varepsilon}$, it is sufficient to consider terms at most quadratic in the noise in the Langevin equation; higher order terms give contributions of order at least $\varepsilon^{3/2}$ and are, therefore, negligible in the continuum, small ε, limit.

From the Langevin equation (4.52) we can derive an equation for the probability distribution $P(\mathbf{q},t+\varepsilon;\mathbf{q}',t)$:

$$P(\mathbf{q},t+\varepsilon;\mathbf{q}',t) = \langle \mathbf{q}-\mathbf{q}(t+\varepsilon)\rangle_{\nu(t)}\,.$$

Following the method of Section 4.3, we first calculate $\tilde P$, the Fourier transform of P:

$$\tilde P(\mathbf{p},t+\varepsilon;\mathbf{q}',t) = \int \mathrm{d}q\; \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}}\, P(\mathbf{q},t+\varepsilon;\mathbf{q}',t) = \left\langle \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}(t+\varepsilon)}\right\rangle_{\nu(t)}\,.$$

Using equation (4.52) and the noise distribution we find

$$\begin{aligned}\tilde P(\mathbf{p},t+\varepsilon;\mathbf{q}',t) \propto \int &\mathrm{d}\nu\, \mathrm{e}^{-(\nu_a)^2/2\varepsilon\Omega} \\ &\times \exp\left[-ip_i\left(q'^i - \frac{\varepsilon}{2}f^i(\mathbf{q}') + e^i_a(\mathbf{q}')\nu_a + \tfrac{1}{2}d^i_{ab}(\mathbf{q}')\nu_a\nu_b\right)\right].\end{aligned}$$

We now perform the gaussian integral over the variables ν_a, keeping terms up to order ε. We obtain

$$\begin{aligned}\tilde P(\mathbf{p},t+\varepsilon;\mathbf{q}',t) \propto \exp\big[&-ip_i\left(q'^i - \tfrac{1}{2}\varepsilon f^i(\mathbf{q}')\right) \\ &-\tfrac{1}{2}i\varepsilon\Omega p_i d^i_{aa}(\mathbf{q}') - \tfrac{1}{2}\varepsilon\Omega e^i_a(\mathbf{q}')e^j_a(\mathbf{q}')p_ip_j\big]\,.\end{aligned} \tag{4.54}$$

We introduce the positive symmetric matrix $g^{ij}(\mathbf{q})$ and the inverse determinant g:

$$g^{ij}(\mathbf{q}) = e^i_a(\mathbf{q})e^j_a(\mathbf{q}), \tag{4.55a}$$

$$g = 1/\det g^{ij}\,, \tag{4.55b}$$

and set

$$h^i(\mathbf{q}) = f^i(\mathbf{q}) - \Omega d^i_{aa}(\mathbf{q}). \tag{4.56}$$

We use the standard notation g_{ij} for the inverse of the matrix g^{ij}:

$$g^{ij} g_{jk} = \delta^i_k \,.$$

We then invert the Fourier transformation

$$P(\mathbf{q}, t+\varepsilon; \mathbf{q}', t) = \frac{1}{(2\pi\varepsilon\Omega)^{d/2}} \sqrt{g(\mathbf{q}')} \exp\left[-\frac{1}{2\varepsilon\Omega} g_{ij}(\mathbf{q}') d^i d^j\right], \tag{4.57}$$

in which we have set

$$\mathbf{d} = \mathbf{q} - \mathbf{q}' + \tfrac{1}{2}\varepsilon \mathbf{h}(\mathbf{q}'). \tag{4.58}$$

Again the distribution P can be written in terms of the matrix elements of an operator of the form $\mathrm{e}^{-\varepsilon H}$, where H is a second order differential operator. Following the method of Section 4.3, we expand equation (4.54) to first order in ε, using the definitions (4.55a) and(4.56). Collecting the terms of order ε, we then find the matrix elements of the corresponding hamiltonian in the mixed $\mathbf{p}, \mathbf{q}'$ representation:

$$\langle \mathbf{p} | H | \mathbf{q}' \rangle = \tfrac{1}{2} \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{q}'} \left[\Omega\, g^{ij}(\mathbf{q}') p_i p_j - i p_i h^i(\mathbf{q}')\right]. \tag{4.59}$$

Inverting the Fourier transformation, we obtain in operator notation

$$H = \tfrac{1}{2}\hat{p}_i \left[\Omega\, \hat{p}_j g^{ij}(\hat{\mathbf{q}}) - i h^i(\hat{\mathbf{q}})\right]. \tag{4.60}$$

The FP equation follows:

$$\frac{\partial}{\partial t} P(\mathbf{q}, t) = \frac{1}{2}\frac{\partial}{\partial q^i}\left[\Omega \frac{\partial}{\partial q^j}\left(g^{ij} P\right) + h^i P\right]. \tag{4.61}$$

This result calls for a simple observation: at leading order in ε the term quadratic in the noise in the Langevin equation (4.52) is equivalent to its average over the noise and only leads to a shift of the function $\mathbf{f}$. In the continuum time limit only equations linear in the noise thus need to be considered.

Finally, from the Langevin equation, which we now write without loss of generality in the small ε limit as

$$q^i(t+\varepsilon) = q^i(t) - \tfrac{1}{2}\varepsilon f^i\big(\mathbf{q}(t)\big) + e^i_a\big(\mathbf{q}(t)\big)\nu_a(t), \tag{4.62}$$

we can directly derive a path integral representation in discretized form for $P(\mathbf{q}, t)$.

4.8 Brownian Motion on Riemannian Manifolds

We now discuss the general Langevin equation directly in the continuum limit, still with the notation and conventions of Chapter 22.

We now consider the general markovian Langevin equation:

$$\dot{q}^i(t) = -\tfrac{1}{2} f^i\big(q(t)\big) + e^i_a\big(q(t)\big)\nu_a(t), \tag{4.63}$$

in which $\nu_a(t)$ is again a gaussian white noise:

$$\langle \nu_a(t)\rangle = 0\,, \qquad \langle \nu_a(t)\nu_b(t')\rangle = \Omega\delta(t-t')\delta_{ab}\,. \tag{4.64}$$

Equations (4.63,4.64) are the continuum analogues of equations (4.53,4.62). The linearity in the noise is a consequence of the markovian character of the noise distribution which implies that the noise correlation functions are proportional to δ-functions in time (see Section 4.7).

Since we have already discussed the discretized Langevin equation, it is sufficient here to integrate the Langevin equation between times t and $t+\varepsilon$ and use the results of Section 4.7. At order ε we find (with $q(t) = q'$)

$$\begin{aligned} q^i(t+\varepsilon) &= q'^i - \tfrac{1}{2}\varepsilon f^i(q') + e^i_a(q') \int_t^{t+\varepsilon} \mathrm{d}\tau\, \nu_a(\tau) \\ &\quad + \partial_j e^i_a(q') e^j_b(q') \int_t^{t+\varepsilon} \mathrm{d}\tau \int_t^{\tau} \mathrm{d}\tau'\, \nu_a(\tau)\nu_b(\tau') + O(\varepsilon^{3/2}). \end{aligned}$$

A term quadratic in the noise has appeared which can be replaced by its average (see the remark at the end of Section 4.7):

$$\begin{aligned} \int_t^{t+\varepsilon} \mathrm{d}\tau \int_t^{\tau} \mathrm{d}\tau'\, \langle \nu_a(\tau)\nu_b(\tau')\rangle &= \delta_{ab}\Omega \int_t^{t+\varepsilon} \mathrm{d}\tau \int_t^{\tau} \mathrm{d}\tau'\, \delta(\tau-\tau') \\ &= \delta_{ab}\Omega\varepsilon\theta(0), \end{aligned}$$

where $\theta(t) = \frac{1}{2}(1+\epsilon(t))$ is the step function. We obtain a term proportional to $\theta(0)$, which is ambiguous. The choice $\epsilon(0) = 0$ and, thus, $\theta(0) = 1/2$ leads to

$$q^i(t+\varepsilon) = q'^i - \tfrac{1}{2}\varepsilon f^i(q') + e^i_a(q') \int_t^{t+\varepsilon} \mathrm{d}\tau\, \nu_a(\tau) + \tfrac{1}{2}\varepsilon\Omega\partial_j e^i_a(q') e^j_a(q'). \tag{4.65}$$

Introducing now the average noise (4.20) we obtain a discretized Langevin equation valid for small ε:

$$q^i(t+\varepsilon) = q^i(t) - \tfrac{1}{2}\varepsilon f^i\big(q(t)\big) + \tfrac{1}{2}\varepsilon\Omega\partial_j e^i_a\big(q(t)\big) e^j_a\big(q(t)\big) + e^i_a\big(q(t)\big)\bar{\nu}_a(t). \tag{4.66}$$

Discussion. In contrast with the Langevin equation (4.1), equation (4.63) is somewhat ambiguous, the difficulty being of exactly the same nature as the problem of ordering operators arising in the quantization of a hamiltonian of the form (3.20).

The problem can, as usual, be best understood by discretizing time (see Section 4.7 for details). Since the variation $q^i(t+\varepsilon) - q^i(t)$ is of order $\sqrt{\varepsilon}$, if in the discretized equation (4.62) we replace, for example, $e^i_a(q(t))$ by $e^i_a\left[\frac{1}{2}(q(t) + q(t+\varepsilon))\right]$, quantities which are formally indistinguishable in the continuum limit, we change the equation at order ε. Instead, if we make the same substitution in $f(q)$, the equation is modified by terms of order $\varepsilon^{3/2}$, which are negligible.

More precisely, if we expand e^i_a for ε small we get

$$e^i_a\left\{\tfrac{1}{2}\left[\mathbf{q}(t)+\mathbf{q}(t+\varepsilon)\right]\right\}=e^i a\left[\mathbf{q}(t)\right]+\tfrac{1}{2}e^j_b\left[\mathbf{q}(t)\right]\nu_b(t)\frac{\partial}{\partial q^j}e^i_a\left[\mathbf{q}(t)\right]+O(\varepsilon). \tag{4.67}$$

This corresponds to an equation (4.52) (Itô calculus) with a specific form of the term quadratic in the noise. The substitution in equation (4.52) of the function d^i_{ab} by

$$d^i_{ab}(\mathbf{q})=e^j_b(\mathbf{q})\frac{\partial}{\partial q^j}e^i_a(\mathbf{q}), \tag{4.68}$$

leads to the equation (4.66).

The choice $\epsilon(0)=0$ thus corresponds to the symmetric form $\frac{1}{2}\left[\mathbf{q}(t)+\mathbf{q}(t+\varepsilon)\right]$ (Stratanovich convention). In what follows the symmetric Stratanovich convention will be adopted, although it is unnatural from the practical point of view, because it has simpler transformation properties under a change of variables and, therefore, of coordinates on a manifold.

4.8.1 The Fokker–Planck equation

Using the results of Section 4.7 we obtain the corresponding FP hamiltonian:

$$H=\tfrac{1}{2}\hat{p}_i\left[\Omega\hat{p}_j g^{ij}(\hat{q})-if^i(\hat{q})+i\Omega\partial_j e^i_a(\hat{q})e^j_a(\hat{q})\right]. \tag{4.69}$$

The FP equation for the time-dependent distribution $P(\mathbf{q},t)$ follows:

$$\dot{P}(q,t)=\tfrac{1}{2}\partial_i\left[\Omega e^i_a(q)\partial_j\left(e^j_a(q)P\right)+f^i(q)P\right]. \tag{4.70}$$

Brownian motion on a Riemannian manifold. Covariance under reparametrization of the manifold implies that the FP equation for the free brownian motion on a manifold is expected to have the form (for more details see Chapter 22, in particular Section 22.5)

$$\dot{D}(q,t)=\frac{\Omega}{2\sqrt{g}}\partial_i\left[\sqrt{g}\,g^{ij}\partial_j D\right], \tag{4.71}$$

in which $g_{ij}(q)$ is the metric tensor on the manifold, g the determinant of $g_{ij}(q)$ (equation (4.55*b*)), and $D(q,t)$ a scalar density, that is, a density normalized with the covariant measure $\sqrt{g}\,\mathrm{d}q$. The density D has a path integral representation, which, following the discussion of Section 3.2, has the covariant form

$$D=\int\left[\mathrm{d}q(\tau)g^{1/2}(q(\tau))\right]\exp\left[-\mathcal{S}(\mathbf{q})/\Omega\right], \tag{4.72}$$

with

$$\mathcal{S}(q)=\tfrac{1}{2}\int\mathrm{d}\tau\,\dot{q}^i g_{ij}(q)\dot{q}^j\,. \tag{4.73}$$

We have already discussed the ambiguity of this continuum expression, even when a hermitian regularization is chosen.

In general the functions $P(q,t)$ and $D(q,t)$ are related by

$$\sqrt{g}D(q,t) = P(q,t). \tag{4.74}$$

This also implies that if the matrix e^i_a is a square matrix, it is the inverse of the vielbein (Section 22.6). If the manifold is compact, a constant scalar density D is obviously a static solution to equation (4.71) and, therefore, corresponds to the equilibrium distribution.

Comparing equations (4.70) and (4.71) we observe that the free brownian motion does not correspond in general to $\mathbf{f} = 0$ since, as a short calculation shows, equations (4.70) and (4.71) are identical only if

$$f^i(q) = -e^i_a(q)\,\nabla_j\, e^j_a(q), \tag{4.75}$$

in which ∇ is the covariant derivative on the manifold. The equation shows that $\mathbf{f}$ does not depend only on the geometry of the manifold but also on the choice of the vielbein.

We finally note that equation (4.70) is the most general second order stochastic differential equation: the operator $\partial/\partial q^i$ is implied by total probability conservation, 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 will become clear once we construct the path integral representation.

4.8.2 Path integral representation

Applying the method of Section 4.6 to equation (4.57) we can derive a path integral representation for the probability distribution, and correlation functions. We give here a more direct derivation based on the Langevin equation (4.63) and the generalization of a simple identity to an infinite number of variables. This method will be systematically used in Chapters 16 and 17.

Integral representation of constraints. We consider a mapping $\mathbf{x} \mapsto \mathbf{y}$,

$$y_i = f_i(\mathbf{x}), \tag{4.76}$$

where the functions $f_i(\mathbf{x})$ are differentiable, and assume that it can be inverted for $|\mathbf{y}|$ small enough, defining a unique set of functions $x_i(\mathbf{y})$. We call $J(\mathbf{x})$ the jacobian of the change of variables $\mathbf{y} \mapsto \mathbf{x}$:

$$J(\mathbf{x}) = \left|\det \frac{\partial f_i}{\partial x_j}\right| . \tag{4.77}$$

We now want to calculate a function $\sigma(\mathbf{x})$, for $\mathbf{x}$ solution of the equation $\mathbf{f}(\mathbf{x}) = 0$, without solving the equation explicitly. We can then use a simple identity, which follows from the definition of Dirac's δ-function:

$$\sigma(\mathbf{x})|_{\mathbf{f}(\mathbf{x})=0} = \int \left\{\prod_{i=1}^{n} \mathrm{d}x_i\, \delta\left[f_i(\mathbf{x})\right]\right\} J(\mathbf{x})\sigma(\mathbf{x}). \tag{4.78}$$

This identity, written here for a finite number of variables, can easily be generalized to an infinite number of variables. In most applications the δ-function is replaced by its Fourier representation, leading to the alternative form

$$\sigma(\mathbf{x})|_{\mathbf{f}(\mathbf{x})=0} = \frac{1}{(2\pi)^n} \int \left\{\prod_{i=1}^{n} \mathrm{d}x_i\, \mathrm{d}\lambda_i\right\} J(\mathbf{x})\, \mathrm{e}^{i\boldsymbol{\lambda}\cdot\mathbf{x}}\, \sigma(\mathbf{x}), \tag{4.79}$$

In general the functions $P(q,t)$ and $D(q,t)$ are related by

$$\sqrt{g}D(q,t) = P(q,t). \tag{4.74}$$

This also implies that if the matrix e_a^i is a square matrix, it is the inverse of the vielbein (Section 22.6). If the manifold is compact, a constant scalar density D is obviously a static solution to equation (4.71) and, therefore, corresponds to the equilibrium distribution.

Comparing equations (4.70) and (4.71) we observe that the free brownian motion does not correspond in general to $\mathbf{f} = 0$ since, as a short calculation shows, equations (4.70) and (4.71) are identical only if

$$f^i(q) = -e_a^i(q)\, \nabla_j\, e_a^j(q), \tag{4.75}$$

in which ∇ is the covariant derivative on the manifold. The equation shows that $\mathbf{f}$ does not depend only on the geometry of the manifold but also on the choice of the vielbein.

We finally note that equation (4.70) is the most general second order stochastic differential equation: the operator $\partial/\partial q^i$ is implied by total probability conservation, 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 will become clear once we construct the path integral representation.

4.8.2 Path integral representation

Applying the method of Section 4.6 to equation (4.57) we can derive a path integral representation for the probability distribution, and correlation functions. We give here a more direct derivation based on the Langevin equation (4.63) and the generalization of a simple identity to an infinite number of variables. This method will be systematically used in Chapters 16 and 17.

Integral representation of constraints. We consider a mapping $\mathbf{x} \mapsto \mathbf{y}$,

$$y_i = f_i(\mathbf{x}), \tag{4.76}$$

where the functions $f_i(\mathbf{x})$ are differentiable, and assume that it can be inverted for $|\mathbf{y}|$ small enough, defining a unique set of functions $x_i(\mathbf{y})$. We call $J(\mathbf{x})$ the jacobian of the change of variables $\mathbf{y} \mapsto \mathbf{x}$:

$$J(\mathbf{x}) = \left|\det \frac{\partial f_i}{\partial x_j}\right|. \tag{4.77}$$

We now want to calculate a function $\sigma(\mathbf{x})$, for $\mathbf{x}$ solution of the equation $\mathbf{f}(\mathbf{x}) = 0$, without solving the equation explicitly. We can then use a simple identity, which follows from the definition of Dirac's δ-function:

$$\sigma(\mathbf{x})|_{\mathbf{f}(\mathbf{x})=0} = \int \left\{\prod_{i=1}^{n} \mathrm{d}x_i\, \delta\left[f_i(\mathbf{x})\right]\right\} J(\mathbf{x})\sigma(\mathbf{x}). \tag{4.78}$$

This identity, written here for a finite number of variables, can easily be generalized to an infinite number of variables. In most applications the δ-function is replaced by its Fourier representation, leading to the alternative form

$$\sigma(\mathbf{x})|_{\mathbf{f}(\mathbf{x})=0} = \frac{1}{(2\pi)^n} \int \left\{\prod_{i=1}^{n} \mathrm{d}x_i\, \mathrm{d}\lambda_i\right\} J(\mathbf{x})\, \mathrm{e}^{i\boldsymbol{\lambda}\cdot\mathbf{x}}\, \sigma(\mathbf{x}), \tag{4.79}$$

where one integrates over imaginary values of λ_i.

Path integral representation. We now want to evaluate functionals of $q^i(t)$ when the function $q^i(t)$ is a solution of the Langevin equation (4.63). We can then apply the identity (4.79).

The jacobian. The jacobian J is the determinant of a differential operator $\widetilde{\mathbf{M}}$, of matrix elements $\widetilde{M}_i^j(t,\tau)$, obtained by differentiating the Langevin equation with respect to $q^i(\tau)$:

$$\widetilde{M}_i^j(t,\tau) = \left[\delta_i^j \frac{\mathrm{d}}{\mathrm{d}t} + \tfrac{1}{2}\partial_i f^j(q) - \partial_i e_a^j(q)\nu_a\right] \delta(t-\tau). \tag{4.80}$$

To calculate the determinant, we first factorize the q-independent operator $\mathrm{d}/\mathrm{d}t$:

$$\widetilde{\mathbf{M}} = (\mathrm{d}/\mathrm{d}t)\,\mathbf{M}\,,$$

where $\mathbf{M}$ corresponds to a kernel which can be written as

$$M_i^j(t,\tau) = \delta_i^j\delta(t-\tau) + \theta(t-\tau)v_i^j(\tau), \tag{4.81}$$

$$v_i^j(\tau) = \tfrac{1}{2}\partial_i f^j\big(q(\tau)\big) - \partial_i e_a^j\big(q(\tau)\big)\nu_a(\tau), \tag{4.82}$$

in which $\theta(t)$ again is the step function.

We then use the identity (1.101), det = exp tr ln, to expand $\det \mathbf{M}$:

$$\begin{aligned}\ln\det\mathbf{M} = \theta(0)\int \operatorname{tr} v(t)\mathrm{d}t - \tfrac{1}{2}\int \theta(t_1-t_2)\theta(t_2-t_1)\operatorname{tr}\left[v(t_1)v(t_2)\right]\mathrm{d}t_1\mathrm{d}t_2 + \cdots \\ + \frac{(-1)^{n+1}}{n}\int \theta(t_1-t_2)\ldots\theta(t_n-t_1)\operatorname{tr}\left[v(t_1)\ldots v(t_n)\right]\mathrm{d}t_1\ldots\mathrm{d}t_n + \cdots\,.\end{aligned} \tag{4.83}$$

Due to the product of θ-functions all terms vanish but the first one:

$$\det\mathbf{M} = \exp\left[\theta(0)\int \operatorname{tr} v(t)\,\mathrm{d}t\right]. \tag{4.84}$$

The result involves the ill-defined quantity $\theta(0)$. Consistent with our general convention, we set $\theta(0) = 1/2$.

Replacing $v(t)$ by the explicit expression (4.82), we obtain

$$J \propto \det\mathbf{M} = \exp\left[\frac{1}{2}\int \mathrm{d}t\,\left(\tfrac{1}{2}\partial_i f^i - \partial_i e_a^i\nu_a\right)\right]. \tag{4.85}$$

The noise average. We then impose the Langevin equation by its Fourier representation, as in the identity (4.79),

$$\delta(\text{Langevin eq.}) = \int [\mathrm{d}\lambda_i(t)] \exp\left\{\int \mathrm{d}t\,\lambda_i(t)\left[\dot{q}^i(t) + \tfrac{1}{2}f^i(q) - e_a^i(q)\nu_a\right]\right\},$$

but with the change $\lambda \mapsto i\lambda$.

where one integrates over imaginary values of λ_i.

Path integral representation. We now want to evaluate functionals of $q^i(t)$ when the function $q^i(t)$ is a solution of the Langevin equation (4.63). We can then apply the identity (4.79).

The jacobian. The jacobian J is the determinant of a differential operator $\widetilde{\mathbf{M}}$, of matrix elements $\widetilde{M}_i^j(t,\tau)$, obtained by differentiating the Langevin equation with respect to $q^i(\tau)$:

$$\widetilde{M}_i^j(t,\tau) = \left[\delta_i^j \frac{\mathrm{d}}{\mathrm{d}t} + \tfrac{1}{2}\partial_i f^j(q) - \partial_i e_a^j(q)\nu_a\right]\delta(t-\tau). \tag{4.80}$$

To calculate the determinant, we first factorize the q-independent operator $\mathrm{d}/\mathrm{d}t$:

$$\widetilde{\mathbf{M}} = (\mathrm{d}/\mathrm{d}t)\,\mathbf{M}\,,$$

where $\mathbf{M}$ corresponds to a kernel which can be written as

$$M_i^j(t,\tau) = \delta_i^j\delta(t-\tau) + \theta(t-\tau)v_i^j(\tau), \tag{4.81}$$

$$v_i^j(\tau) = \tfrac{1}{2}\partial_i f^j\big(q(\tau)\big) - \partial_i e_a^j\big(q(\tau)\big)\nu_a(\tau), \tag{4.82}$$

in which $\theta(t)$ again is the step function.

We then use the identity (1.101), $\det = \exp\operatorname{tr}\ln$, to expand $\det\mathbf{M}$:

$$\begin{aligned}\ln\det\mathbf{M} = {}& \theta(0)\int \operatorname{tr} v(t)\mathrm{d}t - \tfrac{1}{2}\int \theta(t_1-t_2)\theta(t_2-t_1)\operatorname{tr}\left[v(t_1)v(t_2)\right]\mathrm{d}t_1\mathrm{d}t_2 + \cdots \\ & + \frac{(-1)^{n+1}}{n}\int \theta(t_1-t_2)\ldots\theta(t_n-t_1)\operatorname{tr}\left[v(t_1)\ldots v(t_n)\right]\mathrm{d}t_1\ldots\mathrm{d}t_n + \cdots\,.\end{aligned} \tag{4.83}$$

Due to the product of θ-functions all terms vanish but the first one:

$$\det\mathbf{M} = \exp\left[\theta(0)\int \operatorname{tr} v(t)\,\mathrm{d}t\right]. \tag{4.84}$$

The result involves the ill-defined quantity $\theta(0)$. Consistent with our general convention, we set $\theta(0) = 1/2$.

Replacing $v(t)$ by the explicit expression (4.82), we obtain

$$J \propto \det\mathbf{M} = \exp\left[\frac{1}{2}\int \mathrm{d}t\,\left(\tfrac{1}{2}\partial_i f^i - \partial_i e_a^i\nu_a\right)\right]. \tag{4.85}$$

The noise average. We then impose the Langevin equation by its Fourier representation, as in the identity (4.79),

$$\delta(\text{Langevin eq. }) = \int [\mathrm{d}\lambda_i(t)]\exp\left\{\int \mathrm{d}t\,\lambda_i(t)\left[\dot{q}^i(t) + \tfrac{1}{2}f^i(q) - e_a^i(q)\nu_a\right]\right\},$$

but with the change $\lambda \mapsto i\lambda$.

Collecting the noise-dependent factors coming both from the Langevin equation and the jacobian, we see that the noise appears only linearly in the exponential. The average over the noise reduces to the calculation of a simple gaussian integral:

$$\int [\mathrm{d}\nu_a(t)] \exp\left\{-\int \mathrm{d}t \left[\tfrac{1}{2}\nu_a^2/\Omega + \nu_a\left(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\right)\right]\right\}$$
$$\propto \exp\left[\tfrac{1}{2}\Omega\left(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\right)^2\right].$$

Introducing the metric tensor (4.55a), we can expand

$$\left(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\right)^2 = \lambda_i g^{ij}\lambda_j + \lambda_i e_a^i(q)\partial_j e_a^j(q) + \tfrac{1}{4}\left(\partial_i e_a^i(q)\right)^2.$$

The remaining λ integral is also gaussian and can be performed:

$$\int [\mathrm{d}\lambda_i(t)] \exp\left\{\int \mathrm{d}t \left[\tfrac{1}{2}\Omega\lambda_i g^{ij}\lambda_j + \lambda_i\left(\dot q^i + \tfrac{1}{2}f^i(q) + \tfrac{1}{2}\Omega e_a^i(q)\partial_j e_a^j(q)\right)\right]\right\}$$
$$\propto \sqrt{g}\exp\left[-\mathcal{S}_1(q)/\Omega\right]$$

with

$$\mathcal{S}_1(q) = \tfrac{1}{2}\left(\dot q^i + \tfrac{1}{2}f^i\right)g_{ij}\left(\dot q^j + \tfrac{1}{2}f^j\right) + \tfrac{1}{2}\Omega\left(\dot q^i + \tfrac{1}{2}f^i\right)e_{ai}\partial_j e_a^j + \tfrac{1}{8}\Omega^2 e_a^i\partial_j e_a^j e_{bi}\partial_k e_b^k\,.$$

If we now assume that e_a^i is a square invertible matrix (and, therefore, plays the role of a vielbein), the complete expression simplifies because the two terms which depend only on e cancel. Introducing covariant derivatives ∇_i and using the identity (equation (22.67))

$$\nabla_i V^i = \partial_i V^i + (\partial_i \ln e)V^i\,, \quad e = \det e_{ia} = \sqrt{g}\,,$$

we obtain after some algebra, and up to boundary terms,

$$P(q,t;q_0,t_0) = \int_{q(t_0)=q_0}^{q(t)=q} \left[\mathrm{d}q(\tau)\sqrt{g(q(\tau))}\right]\exp\left[-\mathcal{S}(q)/\Omega\right], \tag{4.86}$$

with the definition

$$\mathcal{S}(q) = \frac{1}{2}\int_{t_0}^{t}\mathrm{d}\tau\Big[\left(\dot q^i + \tfrac{1}{2}f^i(q)\right)g_{ij}(q)\left(\dot q^j + \tfrac{1}{2}f^j(q)\right) - \tfrac{1}{2}\Omega\nabla_i f^i + \Omega e_{ai}\nabla_j e_a^j\left(\dot q^i + \tfrac{1}{2}f^i(q)\right)\Big]. \tag{4.87}$$

We note that the functional $[\mathrm{d}q(t)\sqrt{g(q)}]$ measure is formally the covariant measure on the manifold (see Section 22.5), in such a way that the expression (4.86) is completely covariant.

This path integral representation suffers from the ambiguities we have already discussed in Section 3.2. In particular, the covariant functional integration measure is divergent. In both cases the appearance of this divergent factor is related to the question of the order of quantum operators in the quantum mechanics or FP hamiltonian, as already discussed in Section 3.2. To define the path integral more precisely one has to return to the time discretized form (4.57).

Collecting the noise-dependent factors coming both from the Langevin equation and the jacobian, we see that the noise appears only linearly in the exponential. The average over the noise reduces to the calculation of a simple gaussian integral:

$$\int [\mathrm{d}\nu_a(t)] \exp\left\{-\int \mathrm{d}t\, \left[\tfrac{1}{2}\nu_a^2/\Omega + \nu_a\big(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\big)\right]\right\}$$
$$\propto \exp\left[\tfrac{1}{2}\Omega\big(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\big)^2\right].$$

Introducing the metric tensor (4.55a), we can expand

$$\big(\lambda_i e_a^i(q) + \tfrac{1}{2}\partial_i e_a^i(q)\big)^2 = \lambda_i g^{ij}\lambda_j + \lambda_i e_a^i(q)\partial_j e_a^j(q) + \tfrac{1}{4}\big(\partial_i e_a^i(q)\big)^2.$$

The remaining λ integral is also gaussian and can be performed:

$$\int [\mathrm{d}\lambda_i(t)] \exp\left\{\int \mathrm{d}t\, \left[\tfrac{1}{2}\Omega\lambda_i g^{ij}\lambda_j + \lambda_i\big(\dot q^i + \tfrac{1}{2}f^i(q) + \tfrac{1}{2}\Omega e_a^i(q)\partial_j e_a^j(q)\big)\right]\right\}$$
$$\propto \sqrt{g}\exp\left[-\mathcal{S}_1(q)/\Omega\right]$$

with

$$\mathcal{S}_1(q) = \tfrac{1}{2}\big(\dot q^i + \tfrac{1}{2}f^i\big)g_{ij}\big(\dot q^j + \tfrac{1}{2}f^j\big) + \tfrac{1}{2}\Omega\big(\dot q^i + \tfrac{1}{2}f^i\big)e_{ai}\partial_j e_a^j + \tfrac{1}{8}\Omega^2 e_a^i\partial_j e_a^j e_{bi}\partial_k e_b^k\,.$$

If we now assume that e_a^i is a square invertible matrix (and, therefore, plays the role of a vielbein), the complete expression simplifies because the two terms which depend only on e cancel. Introducing covariant derivatives ∇_i and using the identity (equation (22.67))

$$\nabla_i V^i = \partial_i V^i + (\partial_i \ln e)V^i\,, \quad e = \det e_{ia} = \sqrt{g}\,,$$

we obtain after some algebra, and up to boundary terms,

$$P(q,t;q_0,t_0) = \int_{q(t_0)=q_0}^{q(t)=q} \left[\mathrm{d}q(\tau)\sqrt{g(q(\tau))}\right]\exp\left[-\mathcal{S}(q)/\Omega\right], \tag{4.86}$$

with the definition

$$\mathcal{S}(q) = \frac{1}{2}\int_{t_0}^{t}\mathrm{d}\tau\Big[\big(\dot q^i + \tfrac{1}{2}f^i(q)\big)\,g_{ij}(q)\,\big(\dot q^j + \tfrac{1}{2}f^j(q)\big) - \tfrac{1}{2}\Omega\nabla_i f^i + \Omega e_{ai}\nabla_j e_a^j\,\big(\dot q^i + \tfrac{1}{2}f^i(q)\big)\Big]. \tag{4.87}$$

We note that the functional $[\mathrm{d}q(t)\sqrt{g(q)}]$ measure is formally the covariant measure on the manifold (see Section 22.5), in such a way that the expression (4.86) is completely covariant.

This path integral representation suffers from the ambiguities we have already discussed in Section 3.2. In particular, the covariant functional integration measure is divergent. In both cases the appearance of this divergent factor is related to the question of the order of quantum operators in the quantum mechanics or FP hamiltonian, as already discussed in Section 3.2. To define the path integral more precisely one has to return to the time discretized form (4.57).

Bibliographical Notes

A historical reference is

P. Langevin, *Comptes Rendus Acad. Sci. Paris* 146 (1908) 530.

For an introduction to the subject of stochastic processes in physics see:

N.G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam 1981).

A selection of early papers can be found in

Selected Papers on Noise and Stochastic Processes, N. Wax ed. (Dover, New York 1954).

Mathematical contributions are contained in

New Stochastic Methods in Physics, C. De Witt-Morette and K.D. Elworthy eds., *Phys. Rep.* 77 (1981) 121.

Examples of mathematical textbooks are

I.I. Gihman and A.V. Skorohod, *Stochastic Differential Equations* (Springer Verlag, Berlin 1972); N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, Amsterdam 1981);

K.D. Elworthy, *Stochastic Differential Equations on Manifolds, London Math. Soc. Lectures Notes 70* (Cambridge University Press, Cambridge 1982).

For an early reference about the application of path integrals to brownian motion and statistical mechanics see for example

S.G. Brush, *Rev. Mod. Phys.* 33 (1961) 79.

Additional comments on numerical simulations and stochastic processes as discussed in Appendix A4 can be found in

S.K. Ma, *Statistical Mechanics*, Ch. 22 (World Scientific, Singapore 1985); G. Parisi, *Statistical Field Theory*, Ch. 19,20 (Addison-Wesley, New York 1988) .

APPENDIX A4
DISCRETE MARKOV STOCHASTIC PROCESSES: A FEW REMARKS

We recall in the appendix a few simple properties of Markov's processes because it is useful to have them in mind when discussing Langevin or FP equations.

We consider here a finite number N of discrete states labelled below by indices like $a, b...$ The stochastic process is defined in terms of a matrix Π_{ab} ($\Pi_{ab} \geq 0, \forall\, a, b$) which gives the probability of transition at any given time from a state b to a state a. The conservation of probabilities implies

$$\sum_{a=1}^{N} \Pi_{ab} = 1\,. \tag{A4.1}$$

The probability $P_n(a)$ to be at time n in a state a satisfies then the evolution (or master) equation:

$$P_{n+1}(a) = \sum_b \Pi_{ab} P_n(b)\,. \tag{A4.2}$$

Summing equation ($A4.2$) over the index a and using equation ($A4.1$), we find that the total probability is conserved:

$$\sum_a P_{n+1}(a) = \sum_b P_n(b).$$

A4.1 The Spectrum of the Transition Matrix

Equation ($A4.1$) also implies that the vector U_a

$$U_a = 1\,, \qquad \forall a\,,$$

is a left eigenvector of the matrix Π_{ab} with eigenvalue one:

$$\sum_a U_a \Pi_{ab} = \sum_a \Pi_{ab} = 1 = U_b\,.$$

The corresponding right eigenvector,

$$V(a) = \sum_b \Pi_{ab} V(b), \tag{A4.3}$$

is a stationary solution of equation ($A4.2$).

More generally let V_a be any eigenvector of Π_{ab} and v the corresponding eigenvalue:

$$\sum_b \Pi_{ab} V_b = v V_a\,. \tag{A4.4}$$

Summing over a and using equation ($A4.1$), we obtain

$$\sum_b V_b = v \sum_a V_a\,,$$

APPENDIX A4
DISCRETE MARKOV STOCHASTIC PROCESSES: A FEW REMARKS

We recall in the appendix a few simple properties of Markov's processes because it is useful to have them in mind when discussing Langevin or FP equations.

We consider here a finite number N of discrete states labelled below by indices like a, b... The stochastic process is defined in terms of a matrix Π_{ab} ($\Pi_{ab} \geq 0, \forall\, a, b$) which gives the probability of transition at any given time from a state b to a state a. The conservation of probabilities implies

$$\sum_{a=1}^{N} \Pi_{ab} = 1 . \tag{A4.1}$$

The probability $P_n(a)$ to be at time n in a state a satisfies then the evolution (or master) equation:

$$P_{n+1}(a) = \sum_b \Pi_{ab} P_n(b) . \tag{A4.2}$$

Summing equation ($A4.2$) over the index a and using equation ($A4.1$), we find that the total probability is conserved:

$$\sum_a P_{n+1}(a) = \sum_b P_n(b).$$

A4.1 The Spectrum of the Transition Matrix

Equation ($A4.1$) also implies that the vector U_a

$$U_a = 1 , \qquad \forall a ,$$

is a left eigenvector of the matrix Π_{ab} with eigenvalue one:

$$\sum_a U_a \Pi_{ab} = \sum_a \Pi_{ab} = 1 = U_b .$$

The corresponding right eigenvector,

$$V(a) = \sum_b \Pi_{ab} V(b), \tag{A4.3}$$

is a stationary solution of equation ($A4.2$).

More generally let V_a be any eigenvector of Π_{ab} and v the corresponding eigenvalue:

$$\sum_b \Pi_{ab} V_b = v V_a . \tag{A4.4}$$

Summing over a and using equation ($A4.1$), we obtain

$$\sum_b V_b = v \sum_a V_a ,$$

which implies either $v = 1$, or $\sum V_a = 0$.

Comparing the modulus of both sides of equation ($A4.4$) we also find

$$\sum_b \Pi_{ab} \left|V_b\right| \geq |v| \left|V_a\right|, \tag{A4.5}$$

and thus using equation ($A4.1$) again:

$$|v| \leq 1 \,. \tag{A4.6}$$

All eigenvalues of the matrix Π_{ab} have a modulus smaller than or equal to one. Equality is possible only if the inequality ($A4.5$) is an equality for all values of a:

$$\sum_b \Pi_{ab} \left|V_b\right| = \left|V_a\right|, \tag{A4.7}$$

which implies that $|V_a|$ is a stationary probability distribution.

Let us assume that for some subset I of $\{1, \ldots, N\}$, V_a vanishes. Then,

$$\forall a \in I \ : \quad \sum_{b \notin I} \Pi_{ab} \left|V_b\right| = 0 \,. \tag{A4.8}$$

This 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 defined below 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 $c_1, \ldots, c_r$ such that the product,

$$\forall a, b \qquad \exists \{c_1, \ldots, c_r\} : \quad \Pi_{bc_r} \Pi_{c_r c_{r-1}} \ldots \Pi_{c_1 a} > 0 \,.$$

We then call the space of states *connected.*

It follows that, as shown above, the stationary distribution has no vanishing component. Moreover, the eigenspace corresponding to the eigenvalue one has dimension one (if the dimension were larger than one, 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

$$P(a) = \mathrm{e}^{-E(a)} / \mathcal{Z} \quad \text{with} \quad E(a) > 0 \quad \text{and} \quad \mathcal{Z} = \sum_a \mathrm{e}^{-E(a)} \,. \tag{A4.9}$$

If all other eigenvalues of the transition matrix have a modulus smaller than one, then any distribution converges at infinite time towards the stationary distribution $P(a)$ which is thus the *equilibrium distribution.*

Other eigenvalues with modulus one. If there exists another eigenvalue of modulus one

$$v = \mathrm{e}^{i\theta}, \quad \theta \neq 0 \pmod{2\pi},$$

the corresponding eigenspace also has dimension one and the eigenvector V_a satisfies $|V_a| = P(a)$, where $P(a)$ is the stationary distribution ($A4.9$).

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 ($A4.5$) 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 increase to infinity an important new phenomenon may occur: the components of the normalized equilibrium distribution,

$$\sum_a P(a) = 1\,,$$

may all go to zero.

In this case even with the previous 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 23–36. 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.

A4.2 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 which 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 on the matrix Π_{ab} the condition

$$\Pi_{ab} P(b) = \Pi_{ba} P(a) \quad \text{for all pairs } (a, b)\,. \tag{A4.10}$$

This condition is called *detailed balance*. It is a local condition involving only the states a and b.

The condition implies that $P(a)$ is a stationary distribution. Indeed

$$\sum_b \Pi_{ab} P(b) = \sum_b \Pi_{ba} P(a) = P(a). \tag{A4.11}$$

We then set

$$P_n(a) = \sqrt{P(a)}\, \tilde{P}_n(a). \tag{A4.12}$$

The equation ($A4.2$) becomes

$$\tilde{P}_{n+1}(a) = \sum_b \tilde{\Pi}_{ab}\, \tilde{P}_n(b), \tag{A4.13}$$

where the matrix $\tilde{\Pi}_{ab}$,

$$\tilde{\Pi}_{ab} = \Pi_{ab} \sqrt{P(b)} / \sqrt{P(a)}, \tag{A4.14}$$

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 ($A4.5$) 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 $\mathrm{e}^{-i\theta}$ 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 $\mathrm{e}^{-i(k-1)\theta}$ 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 $\mathrm{e}^{in\theta} = 1$, 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 increase to infinity an important new phenomenon may occur: the components of the normalized equilibrium distribution,

$$\sum_a P(a) = 1 \, ,$$

may all go to zero.

In this case even with the previous 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 23–36. 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.

A4.2 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 which 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 on the matrix Π_{ab} the condition

$$\Pi_{ab} P(b) = \Pi_{ba} P(a) \quad \text{for all pairs } \ (a, b) \, . \tag{A4.10}$$

This condition is called *detailed balance.* It is a local condition involving only the states a and b.

The condition implies that $P(a)$ is a stationary distribution. Indeed

$$\sum_b \Pi_{ab} P(b) = \sum_b \Pi_{ba} P(a) = P(a). \tag{A4.11}$$

We then set

$$P_n(a) = \sqrt{P(a)}\, \tilde{P}_n(a). \tag{A4.12}$$

The equation ($A4.2$) becomes

$$\tilde{P}_{n+1}(a) = \sum_b \tilde{\Pi}_{ab}\, \tilde{P}_n(b), \tag{A4.13}$$

where the matrix $\tilde{\Pi}_{ab}$,

$$\tilde{\Pi}_{ab} = \Pi_{ab} \sqrt{P(b)} / \sqrt{P(a)}, \tag{A4.14}$$

is symmetric as a consequence of equation ($A4.10$). Its spectrum is thus real. The distribution $P(a)$ is an equilibrium distribution except if Π_{ab} has -1 as eigenvalue. This means, according to previous discussion, that 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.

A4.3 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

$$\begin{aligned}\Pi_{ab} &= \frac{1}{r}\,, &&\text{if}\quad P(a)\geq P(b)\quad\text{and}\quad a\neq b\,,\\ \Pi_{ab} &= \frac{1}{r}\frac{P(a)}{P(b)}\,, &&\text{if}\quad P(a)<P(b).\end{aligned}\qquad (A4.15)$$

These conditions imply detailed balance. In addition, as required

$$\Pi_{bb} = 1-\sum_{a\neq b}\Pi_{ab} > 0\,. \qquad (A4.16)$$

Another idea is to take

$$\Pi_{ab} = pP(a)\theta_{ab}\,,\quad a\neq b \qquad (A4.17)$$

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, it is easy in the discrete case 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}.

5 PATH AND FUNCTIONAL INTEGRALS IN QUANTUM STATISTICAL PHYSICS

Hamiltonians in quantum mechanics can be expressed in terms of creation and annihilation operators, instead of the more usual position and momentum operators, a method well adapted to the study of perturbed harmonic oscillators. In the holomorphic formalism these operators act by multiplication and differentiation on a vector space of analytic functions. Alternatively, they can also be represented by kernels, functions of complex variables $z, \bar{z}$ which in the classical limit correspond to a complex parametrization of phase space.

To this formalism corresponds a path integral representation of the statistical operator (the density matrix at thermal equilibrium) where paths belong to complex spaces instead of the more usual position–momentum phase space. Its construction provides a useful introduction to the construction of the path integral for fermion degrees of freedom. Both formalisms can then be generalized to a quantum gas of Bose or Fermi particles in the grand canonical formulation. A functional integral representation of quantum partition functions can be derived, a topic we discuss in the second part of the chapter. Finally, these formalisms allow the construction of the functional integral representation of the scattering S-matrix in quantum field theory as we show in Chapters 6 and 8.

5.1 Quantum Mechanics: Holomorphic Formalism

We first briefly recall the main ideas and properties of the holomorphic formalism. To motivate the construction, we consider the harmonic oscillator

$$H_0 = \tfrac{1}{2}\hat{p}^2 + \tfrac{1}{2}\omega^2\hat{q}^2 . \tag{5.1}$$

We introduce the usual annihilation and creation operators $a, a^\dagger$:

$$\hat{p} - i\omega\hat{q} = -i\sqrt{2\hbar\omega}\, a\,, \qquad \hat{p} + i\omega\hat{q} = i\sqrt{2\hbar\omega}\, a^\dagger \quad \Rightarrow \ [a, a^\dagger] = 1\,. \tag{5.2}$$

In terms of $a, a^\dagger$ the hamiltonian takes the standard form $\hbar\omega(a^\dagger a + 1/2)$ with $\omega > 0$. In what follows, we omit the constant energy shift $\hbar\omega/2$ and consider

$$H_0 = \hbar\omega a^\dagger a\,. \tag{5.3}$$

We then introduce a complex variable z and represent $a^\dagger, a$ by operators having the same commutation relations, z and $\partial/\partial z$, respectively,

$$a^\dagger \mapsto z\,, \quad a \mapsto \partial/\partial z\,, \tag{5.4}$$

acting by multiplication and differentiation on a complex vector space of analytic functions $f(z)$. The hamiltonian H_0 is then represented by

$$H_0 = \hbar\omega z \frac{\partial}{\partial z}\,. \tag{5.5}$$

PATH AND FUNCTIONAL INTEGRALS IN QUANTUM STATISTICAL PHYSICS

Hamiltonians in quantum mechanics can be expressed in terms of creation and annihilation operators, instead of the more usual position and momentum operators, a method well adapted to the study of perturbed harmonic oscillators. In the holomorphic formalism these operators act by multiplication and differentiation on a vector space of analytic functions. Alternatively, they can also be represented by kernels, functions of complex variables $z, \bar{z}$ which in the classical limit correspond to a complex parametrization of phase space.

To this formalism corresponds a path integral representation of the statistical operator (the density matrix at thermal equilibrium) where paths belong to complex spaces instead of the more usual position–momentum phase space. Its construction provides a useful introduction to the construction of the path integral for fermion degrees of freedom. Both formalisms can then be generalized to a quantum gas of Bose or Fermi particles in the grand canonical formulation. A functional integral representation of quantum partition functions can be derived, a topic we discuss in the second part of the chapter. Finally, these formalisms allow the construction of the functional integral representation of the scattering S-matrix in quantum field theory as we show in Chapters 6 and 8.

5.1 Quantum Mechanics: Holomorphic Formalism

We first briefly recall the main ideas and properties of the holomorphic formalism. To motivate the construction, we consider the harmonic oscillator

$$H_0 = \tfrac{1}{2}\hat{p}^2 + \tfrac{1}{2}\omega^2\hat{q}^2 . \tag{5.1}$$

We introduce the usual annihilation and creation operators $a, a^\dagger$:

$$\hat{p} - i\omega\hat{q} = -i\sqrt{2\hbar\omega}\, a , \qquad \hat{p} + i\omega\hat{q} = i\sqrt{2\hbar\omega}\, a^\dagger \quad \Rightarrow \ [a, a^\dagger] = 1 . \tag{5.2}$$

In terms of $a, a^\dagger$ the hamiltonian takes the standard form $\hbar\omega(a^\dagger a + 1/2)$ with $\omega > 0$. In what follows, we omit the constant energy shift $\hbar\omega/2$ and consider

$$H_0 = \hbar\omega a^\dagger a . \tag{5.3}$$

We then introduce a complex variable z and represent $a^\dagger, a$ by operators having the same commutation relations, z and $\partial/\partial z$, respectively,

$$a^\dagger \mapsto z , \quad a \mapsto \partial/\partial z , \tag{5.4}$$

acting by multiplication and differentiation on a complex vector space of analytic functions $f(z)$. The hamiltonian H_0 is then represented by

$$H_0 = \hbar\omega z \frac{\partial}{\partial z} . \tag{5.5}$$

In this representation, the eigenfunctions of H_0 are the monomials z^n:

$$\hbar\omega z \frac{\partial}{\partial z} z^n = n\hbar\omega\, z^n .$$

The action of the operator $U_0(t) = \mathrm{e}^{-H_0 t/\hbar}$ (again we first keep the $\hbar$ normalization of real-time evolution) on analytic functions is

$$U_0(t) f(z) = \mathrm{e}^{-H_0 t/\hbar} f(z) = f(\mathrm{e}^{-\omega t} z). \tag{5.6}$$

Hilbert space of analytic functions. We now give a structure of Hilbert space to the vector space of analytic functions. We define the scalar product of two functions g and f by

$$(g, f) = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi} \mathrm{e}^{-z\bar{z}}\, \overline{g(z)}\, f(z), \tag{5.7}$$

where complex integrals are defined and discussed in Section 1.3.

With this scalar product the eigenfunctions of the harmonic oscillator form an orthogonal basis:

$$\int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi} \mathrm{e}^{-z\bar{z}}\, \bar{z}^n z^m = n!\, \delta_{mn}\,, \tag{5.8}$$

a result that follows, for instance, from Wick's theorem (1.30). In particular, the norm of a function f is finite if

$$f(z) = \sum_n f_n z^n \ \Rightarrow\ (f, f) = \sum_n |f_n|^2 n! < \infty .$$

Therefore, square integrable functions belong to a subset of entire analytic functions.

A simple consequence of the orthogonality relations (5.8) is the representation of the "δ-function",

$$f(0) = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi} \mathrm{e}^{-z\bar{z}}\, f(z). \tag{5.9}$$

Operator kernels. Since the functions $z^n/\sqrt{n!}$ form an orthonormal basis the identity operator can be represented by the kernel

$$\mathrm{e}^{z\bar{z}} = \sum_{n=0} \frac{\bar{z}^n}{\sqrt{n!}} \frac{z^n}{\sqrt{n!}} .$$

This property is expressed by the identity

$$\int \frac{\mathrm{d}\bar{z}'\mathrm{d}z'}{2i\pi} \mathrm{e}^{\bar{z}' z}\, \mathrm{e}^{-z'\bar{z}'} f(z') = f(z), \tag{5.10}$$

which is a direct consequence of the identity (5.9).

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

$$a^{\dagger m} a^n \mapsto z^m \left(\partial/\partial z\right)^n ,$$

in the representation (5.4). 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 $\langle z|\, \mathcal{O}\, |\bar{z}\rangle$, obtained by acting on the identity,

$$\mathcal{O}(z, \bar{z}) \equiv \langle z|\, \mathcal{O}\, |\bar{z}\rangle = O(z, \partial/\partial z)\, \mathrm{e}^{z\bar{z}} = O(z, \bar{z})\, \mathrm{e}^{z\bar{z}} . \tag{5.11}$$

Acting with $O(z, \partial/\partial z)$ on both sides of equation (5.10) we find

$$(Of)(z) = \int \frac{\mathrm{d}\bar{z}' \mathrm{d}z'}{2i\pi} \mathcal{O}(z, \bar{z}')\, \mathrm{e}^{-z'\bar{z}'}\, f(z').$$

The kernel associated with the product of the two operators is then given by

$$\int \frac{\mathrm{d}\bar{z}' \mathrm{d}z'}{2i\pi} \langle z|\, \mathcal{O}_2\, |\bar{z}'\rangle\, \mathrm{e}^{-z'\bar{z}'} \langle z'|\, \mathcal{O}_1\, |\bar{z}\rangle = \langle z|\, \mathcal{O}_2 \mathcal{O}_1\, |\bar{z}\rangle . \tag{5.12}$$

From the form (5.7) of the scalar product, one also infers the expression of the trace of an operator,

$$\operatorname{tr} \mathcal{O} = \int \frac{\mathrm{d}\bar{z}\, \mathrm{d}z}{2i\pi}\, \mathrm{e}^{-z\bar{z}}\, \mathcal{O}(z, \bar{z}), \tag{5.13}$$

a form consistent with the cyclic property of the trace, as one can verify by taking the trace of equation (5.12).

With the definition (5.11), the matrix elements of the hamiltonian (5.5) and the operator $U_0(t)$ are, respectively,

$$\langle z|\, H_0\, |\bar{z}\rangle = \hbar\omega z\bar{z}\, \mathrm{e}^{z\bar{z}} , \qquad \langle z|\, U_0(t)\, |\bar{z}\rangle = \mathrm{e}^{z\bar{z}\, \mathrm{e}^{-\omega t}} . \tag{5.14}$$

The partition function $\mathcal{Z}_0(\beta)$ corresponding to H_0 is the trace of $U_0(\hbar\beta)$. Using equations (5.13,5.14), one finds the expected result,

$$\mathcal{Z}_0(\beta) = \operatorname{tr} U_0(\hbar\beta) = \int \frac{\mathrm{d}\bar{z}\, \mathrm{d}z}{2i\pi}\, \mathrm{e}^{-z\bar{z}}\, \mathrm{e}^{z\bar{z}\, \mathrm{e}^{-\omega\hbar\beta}} = \frac{1}{1 - \mathrm{e}^{-\hbar\omega\beta}} . \tag{5.15}$$

Remarks:

(i) The hermitian conjugation of operators becomes the complex conjugation of kernels

$$O \mapsto O^\dagger \quad \Rightarrow \quad \mathcal{O}(z, \bar{z}) \mapsto \overline{\mathcal{O}}(z, \bar{z}).$$

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 a mixed position–momentum representation $\langle q|\, \mathcal{O}\, |p\rangle$ which is obtained by a Fourier transformation on one argument:

$$\langle q|\, \mathcal{O}\, |p\rangle = \int \mathrm{d}q'\; \mathrm{e}^{ipq'/\hbar}\, \langle q|\, \mathcal{O}\, |q'\rangle .$$

in the representation (5.4). 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 $\langle z| \mathcal{O} |\bar{z}\rangle$, obtained by acting on the identity,

$$\mathcal{O}(z, \bar{z}) \equiv \langle z| \mathcal{O} |\bar{z}\rangle = O(z, \partial/\partial z)\, \mathrm{e}^{z\bar{z}} = O(z, \bar{z})\, \mathrm{e}^{z\bar{z}} . \tag{5.11}$$

Acting with $O(z, \partial/\partial z)$ on both sides of equation (5.10) we find

$$(Of)(z) = \int \frac{\mathrm{d}\bar{z}'\mathrm{d}z'}{2i\pi} \mathcal{O}(z, \bar{z}')\, \mathrm{e}^{-z'\bar{z}'}\, f(z').$$

The kernel associated with the product of the two operators is then given by

$$\int \frac{\mathrm{d}\bar{z}'\mathrm{d}z'}{2i\pi} \langle z| \mathcal{O}_2 |\bar{z}'\rangle\, \mathrm{e}^{-z'\bar{z}'}\, \langle z'| \mathcal{O}_1 |\bar{z}\rangle = \langle z| \mathcal{O}_2\mathcal{O}_1 |\bar{z}\rangle . \tag{5.12}$$

From the form (5.7) of the scalar product, one also infers the expression of the trace of an operator,

$$\operatorname{tr} \mathcal{O} = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\, \mathrm{e}^{-z\bar{z}}\, \mathcal{O}(z, \bar{z}), \tag{5.13}$$

a form consistent with the cyclic property of the trace, as one can verify by taking the trace of equation (5.12).

With the definition (5.11), the matrix elements of the hamiltonian (5.5) and the operator $U_0(t)$ are, respectively,

$$\langle z| H_0 |\bar{z}\rangle = \hbar\omega z\bar{z}\, \mathrm{e}^{z\bar{z}} , \qquad \langle z| U_0(t) |\bar{z}\rangle = \mathrm{e}^{z\bar{z}\,\mathrm{e}^{-\omega t}} . \tag{5.14}$$

The partition function $\mathcal{Z}_0(\beta)$ corresponding to H_0 is the trace of $U_0(\hbar\beta)$. Using equations (5.13,5.14), one finds the expected result,

$$\mathcal{Z}_0(\beta) = \operatorname{tr} U_0(\hbar\beta) = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\, \mathrm{e}^{-z\bar{z}}\, \mathrm{e}^{z\bar{z}\,\mathrm{e}^{-\omega\hbar\beta}} = \frac{1}{1 - \mathrm{e}^{-\hbar\omega\beta}} . \tag{5.15}$$

Remarks:

(i) The hermitian conjugation of operators becomes the complex conjugation of kernels

$$O \mapsto O^\dagger \quad \Rightarrow \quad \mathcal{O}(z, \bar{z}) \mapsto \overline{\mathcal{O}}(z, \bar{z}).$$

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 a mixed position–momentum representation $\langle q| \mathcal{O} |p\rangle$ which is obtained by a Fourier transformation on one argument:

$$\langle q| \mathcal{O} |p\rangle = \int \mathrm{d}q'\; \mathrm{e}^{ipq'/\hbar}\, \langle q| \mathcal{O} |q'\rangle .$$

5.2 Holomorphic Path Integral

We now use the formalism to derive a path integral representation of the matrix elements of the solution of equation (2.5),

$$\hbar \frac{\partial U}{\partial t}(t,t') = -H(t)U(t,t'), \qquad U(t',t') = \mathbf{1}\,, \tag{5.16}$$

first in the example of the harmonic oscillator and then for more general hamiltonians. We expect the path integral to be related to the form (3.9) by a simple change of variables of the form (5.2) (quantum operators being replaced by classical variables) but the boundary conditions and boundary terms require a more detailed analysis.

5.2.1 The harmonic oscillator

We first expand the exact expression (5.14) for small euclidean time ε:

$$\langle z|\, U_0(\varepsilon)\, |\bar{z}\rangle = \exp\bigl(z\bar{z}(1-\omega\varepsilon) + O(\varepsilon^2)\bigr). \tag{5.17}$$

We then use the group property in the form (5.12) to write the operator U_0 at finite time:

$$\langle z''|\, U_0(t'',t')\, |\bar{z}'\rangle = \lim_{n\to\infty} \int \prod_{k=1}^{n-1} \frac{\mathrm{d}\bar{z}_k \mathrm{d}z_k}{2i\pi} \exp\left[-\mathcal{S}_\varepsilon(z,\bar{z})\right], \tag{5.18}$$

with

$$\mathcal{S}_\varepsilon(z,\bar{z}) = \left[-\sum_{k=1}^{n-1} \bar{z}_k\,(z_{k+1}-z_k) - \bar{z}_0 z_1 + \omega\varepsilon \sum_{k=0}^{n-1} \bar{z}_k z_{k+1}\right], \tag{5.19}$$

where $\varepsilon = (t''-t')/n$ is the time step and the boundary conditions are

$$\bar{z}_0 = \bar{z}'\,, \quad z_n = z''. \tag{5.20}$$

In the formal continuum limit $n\to\infty$, $\varepsilon\to 0$, the expression (5.18) becomes a path integral:

$$\langle z''|\, U_0(t'',t')\, |\bar{z}'\rangle = \int \left[\frac{\mathrm{d}\bar{z}(t)\mathrm{d}z(t)}{2i\pi}\right] \exp\left[-\mathcal{S}_0(z,\bar{z})\right], \tag{5.21}$$
$$\mathcal{S}_0(z,\bar{z}) = -\bar{z}(t')z(t') + \int_{t'}^{t''} \mathrm{d}t\, \bar{z}(t)\left[-\dot{z}(t) + \omega z(t)\right],$$

with 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 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 (5.13) we obtain

$$\mathcal{Z}_0(\beta) = \operatorname{tr} U_0(\hbar\beta/2, -\hbar\beta/2) = \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}z_k \mathrm{d}\bar{z}_k}{2i\pi} \exp\left[-\mathcal{S}_\varepsilon(\bar{z},z)\right], \tag{5.22}$$

where $\mathcal{S}_\varepsilon$ now has different boundary terms:

$$\mathcal{S}_\varepsilon(\bar{z},z)=\sum_{k=1}^{n}\left[-\bar{z}_{k-1}\left(z_k-z_{k-1}\right)+\hbar\omega\varepsilon\bar{z}_{k-1}z_k\right], \tag{5.23}$$

with the identification $\bar{z}_n=\bar{z}_0$ and $z_n=z_0$.

In the continuum limit we obtain a path integral representation of the partition function:

$$\mathcal{Z}_0(\beta)=\int\left[\frac{\mathrm{d}z(t)\mathrm{d}\bar{z}(t)}{2i\pi}\right]\exp\left[-\mathcal{S}_0\left(\bar{z},z\right)\right], \tag{5.24}$$

$$\mathcal{S}_0\left(\bar{z},z\right)=\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\bar{z}(t)\left[-\dot{z}(t)+\hbar\omega z(t)\right], \tag{5.25}$$

with the periodic boundary conditions $z(-\beta/2)=z(\beta/2)$, $\bar{z}(-\beta/2)=\bar{z}(\beta/2)$.

Generating functional. This expression generalizes immediately to a system linearly coupled to external sources $\bar{b}(t)$ and $b(t)$:

$$H=\hbar\omega a^\dagger a-\bar{b}(t)a^\dagger-b(t)a. \tag{5.26}$$

At order ε, the solution $U_{\mathrm{G}}(t+\varepsilon,t)$ of equation (2.5) is $1-\varepsilon H(t)$ and, therefore,

$$\begin{aligned}\langle z|\,U_{\mathrm{G}}(t+\varepsilon,t)\,|\bar{z}\rangle&=\mathrm{e}^{\bar{z}z}\left[1-\varepsilon\big(\omega z\bar{z}-b(t)\bar{z}-\bar{b}(t)z\big)\right]+O(\varepsilon^2)\\&=\exp\left[\bar{z}z-\varepsilon\big(\omega z\bar{z}-b(t)\bar{z}-\bar{b}(t)z\big)\ +O(\varepsilon^2)\right].\end{aligned}$$

In the continuum limit the action $\mathcal{S}_0$ in the path integral (5.21) is replaced by the action $\mathcal{S}_{\mathrm{G}}(z,\bar{z})$:

$$\mathcal{S}_{\mathrm{G}}(z,\bar{z})=-\bar{z}(t')z(t')+\int_{t'}^{t''}\mathrm{d}t\ \left\{\bar{z}(t)\left[-\dot{z}(t)+\omega z(t)\right]-\bar{z}(t)b(t)-\bar{b}(t)z(t)\right\}. \tag{5.27}$$

The trace of U_{G} takes the form (equation (5.13))

$$\begin{aligned}\operatorname{tr}U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2)&=\int\frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\,\mathrm{e}^{-z\bar{z}}\,U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2;z,\bar{z})\\&=\int\left[\frac{\mathrm{d}\bar{z}(t)\mathrm{d}z(t)}{2i\pi}\right]\exp\left[-\mathcal{S}_{\mathrm{G}}(z,\bar{z})\right]\end{aligned} \tag{5.28}$$

with

$$\mathcal{S}_{\mathrm{G}}(z,\bar{z})=\int_{-\beta/2}^{\beta/2}\mathrm{d}t\ \left\{\bar{z}(t)\left[-\dot{z}(t)+\hbar\omega z(t)\right]-\bar{z}(t)b(t)-\bar{b}(t)z(t)\right\},$$

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 will be presented in the fermion case. The result is

$$\operatorname{tr}U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2)=\mathcal{Z}_0(\beta)\exp\left[\int_{\beta/2}^{\beta/2}\mathrm{d}t\,\mathrm{d}u\,\bar{b}(u)\Delta(t-u)b(t)\right], \tag{5.29}$$

where $\mathcal{S}_\varepsilon$ now has different boundary terms:

$$\mathcal{S}_\varepsilon(\bar{z},z)=\sum_{k=1}^{n}\left[-\bar{z}_{k-1}\left(z_k-z_{k-1}\right)+\hbar\omega\varepsilon\bar{z}_{k-1}z_k\right], \tag{5.23}$$

with the identification $\bar{z}_n=\bar{z}_0$ and $z_n=z_0$.

In the continuum limit we obtain a path integral representation of the partition function:

$$\mathcal{Z}_0(\beta)=\int\left[\frac{\mathrm{d}z(t)\mathrm{d}\bar{z}(t)}{2i\pi}\right]\exp\left[-\mathcal{S}_0\left(\bar{z},z\right)\right], \tag{5.24}$$

$$\mathcal{S}_0\left(\bar{z},z\right)=\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\bar{z}(t)\left[-\dot{z}(t)+\hbar\omega z(t)\right], \tag{5.25}$$

with the periodic boundary conditions $z(-\beta/2)=z(\beta/2)$, $\bar{z}(-\beta/2)=\bar{z}(\beta/2)$.

Generating functional. This expression generalizes immediately to a system linearly coupled to external sources $\bar{b}(t)$ and $b(t)$:

$$H=\hbar\omega a^\dagger a-\bar{b}(t)a^\dagger-b(t)a. \tag{5.26}$$

At order ε, the solution $U_{\mathrm{G}}(t+\varepsilon,t)$ of equation (2.5) is $1-\varepsilon H(t)$ and, therefore,

$$\begin{aligned}\langle z|\,U_{\mathrm{G}}(t+\varepsilon,t)\,|\bar{z}\rangle&=\mathrm{e}^{\bar{z}z}\left[1-\varepsilon\big(\omega z\bar{z}-b(t)\bar{z}-\bar{b}(t)z\big)\right]+O(\varepsilon^2)\\&=\exp\left[\bar{z}z-\varepsilon\big(\omega z\bar{z}-b(t)\bar{z}-\bar{b}(t)z\big)\ +O(\varepsilon^2)\right].\end{aligned}$$

In the continuum limit the action $\mathcal{S}_0$ in the path integral (5.21) is replaced by the action $\mathcal{S}_{\mathrm{G}}(z,\bar{z})$:

$$\mathcal{S}_{\mathrm{G}}(z,\bar{z})=-\bar{z}(t')z(t')+\int_{t'}^{t''}\mathrm{d}t\ \left\{\bar{z}(t)\left[-\dot{z}(t)+\omega z(t)\right]-\bar{z}(t)b(t)-\bar{b}(t)z(t)\right\}. \tag{5.27}$$

The trace of U_{G} takes the form (equation (5.13))

$$\begin{aligned}\operatorname{tr}U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2)&=\int\frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\,\mathrm{e}^{-z\bar{z}}\,U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2;z,\bar{z})\\&=\int\left[\frac{\mathrm{d}\bar{z}(t)\mathrm{d}z(t)}{2i\pi}\right]\exp\left[-\mathcal{S}_{\mathrm{G}}(z,\bar{z})\right]\end{aligned} \tag{5.28}$$

with

$$\mathcal{S}_{\mathrm{G}}(z,\bar{z})=\int_{-\beta/2}^{\beta/2}\mathrm{d}t\ \left\{\bar{z}(t)\left[-\dot{z}(t)+\hbar\omega z(t)\right]-\bar{z}(t)b(t)-\bar{b}(t)z(t)\right\},$$

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 will be presented in the fermion case. The result is

$$\operatorname{tr}U_{\mathrm{G}}(\hbar\beta/2,-\hbar\beta/2)=\mathcal{Z}_0(\beta)\exp\left[\int_{\beta/2}^{\beta/2}\mathrm{d}t\,\mathrm{d}u\,\bar{b}(u)\Delta(t-u)b(t)\right], \tag{5.29}$$

where $\mathcal{Z}_0(\beta)$ is the partition function (5.15) of the harmonic oscillator, and the function $\Delta(t)$,

$$\Delta(t) = \tfrac{1}{2}\,\mathrm{e}^{-\hbar\omega t}\left[\epsilon(t) + 1/\tanh(\hbar\omega\beta/2)\right], \tag{5.30}$$

($\epsilon(t) = 1$ for $t > 0$, $\epsilon(-t) = -\epsilon(t)$, $\dot\epsilon = 2\delta$) is the solution of the differential equation

$$\dot\Delta(t) + \hbar\omega\Delta(t) = \delta(t), \tag{5.31}$$

in the interval $[-\beta/2, \beta/2]$ with periodic boundary conditions.

In particular, the gaussian two-point function with weight $\mathrm{e}^{-\mathcal{S}_0}$ (equation (5.25)) or propagator, the basic element of perturbation theory, is

$$\begin{aligned}\langle \bar z(t_2) z(t_1)\rangle_0 &= \mathcal{Z}_0^{-1}(\beta)\,\frac{\delta^2}{\delta b(t_2)\delta\bar b(t_1)}\exp\left[\int_{\beta/2}^{\beta/2}\mathrm{d}t\,\mathrm{d}u\,\bar b(t)\Delta(u-t)b(u)\right]\Bigg|_{b=\bar b\equiv 0}\\ &= \Delta(t_2 - t_1).\end{aligned} \tag{5.32}$$

Verification. The partition function $\mathcal{Z}_0(\beta)$ (equation (5.24)) can be related to the propagator by the trick of Section 2.5.3. Differentiating the path integral (5.24) with respect to ω, one finds

$$\frac{\partial \ln \mathcal{Z}_0(\beta)}{\partial\omega} = -\hbar\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\langle z(t)\bar z(t)\rangle_0 = -\hbar\beta\Delta(0).$$

A difficulty then arises because $\Delta(t)$ is not continuous at $t = 0$; this, again, is the $\epsilon(0)$ problem. Integrating, we obtain (using the limit $\beta \to \infty$ to determine the integration constant)

$$\mathcal{Z}_0(\beta) = \frac{\mathrm{e}^{-\hbar\beta\omega\epsilon(0)/2}}{2\sinh(\hbar\beta\omega/2)}.$$

We see that the ambiguity has the nature of a constant shift of the hamiltonian, a natural consequence of the order problem between the operators $a, a^\dagger$. The rule consistent with the normal order (5.3) is to set $\epsilon(0) = -1$, and yields

$$\mathcal{Z}_0(\beta) = \frac{1}{1 - \mathrm{e}^{-\hbar\beta\omega}}\,. \tag{5.33}$$

However, , such a choice is inconvenient for perturbative calculations, because it breaks time reversal symmetry, and it is better to choose $\epsilon(0) = 0$. This corresponds to a symmetric form $\frac{1}{2}(aa^\dagger + a^\dagger a)$ and yields the standard harmonic oscillator. The additional contribution can then be cancelled, if necessary, by shifting physical parameters.

5.2.2 General hamiltonian: one degree of freedom

For a general quantum hamiltonian one can first write the quantum hamiltonian in terms of creation and annihilation operators, commute all creation operators to the left (normal order) and then replace operators by the corresponding classical variables, as explained in Section 5.1. One obtains the matrix elements $\langle z|\,U(t+\varepsilon, t)\,|\bar z\rangle$ at order ε, and following the method of Section 5.2 derives a path integral representation of the form

$$\langle z''|\,U(t'', t')\,|\bar z'\rangle = \int\left[\frac{\mathrm{d}\bar z(t)\mathrm{d}z(t)}{2i\pi}\right]\exp\left[-\mathcal{S}(z, \bar z)\right], \tag{5.34}$$

$$\mathcal{S}\big(z(t), \bar z(t)\big) = -\bar z(t')z(t') + \int_{t'}^{t''}\mathrm{d}t\,\left[-\bar z(t)\dot z(t) + h\big(z(t), \bar z(t)\big)/\hbar\right], \tag{5.35}$$

where $\mathcal{Z}_0(\beta)$ is the partition function (5.15) of the harmonic oscillator, and the function $\Delta(t)$,

$$\Delta(t) = \tfrac{1}{2}\,\mathrm{e}^{-\hbar\omega t}\left[\epsilon(t) + 1/\tanh(\hbar\omega\beta/2)\right], \tag{5.30}$$

($\epsilon(t) = 1$ for $t > 0$, $\epsilon(-t) = -\epsilon(t)$, $\dot\epsilon = 2\delta$) is the solution of the differential equation

$$\dot\Delta(t) + \hbar\omega\Delta(t) = \delta(t), \tag{5.31}$$

in the interval $[-\beta/2, \beta/2]$ with periodic boundary conditions.

In particular, the gaussian two-point function with weight $\mathrm{e}^{-\mathcal{S}_0}$ (equation (5.25)) or propagator, the basic element of perturbation theory, is

$$\begin{aligned}\langle \bar z(t_2) z(t_1)\rangle_0 &= \mathcal{Z}_0^{-1}(\beta)\,\frac{\delta^2}{\delta b(t_2)\delta\bar b(t_1)}\exp\left[\int_{\beta/2}^{\beta/2}\mathrm{d}t\,\mathrm{d}u\,\bar b(t)\Delta(u-t)b(u)\right]\Bigg|_{b=\bar b\equiv 0}\\ &= \Delta(t_2 - t_1).\end{aligned} \tag{5.32}$$

Verification. The partition function $\mathcal{Z}_0(\beta)$ (equation (5.24)) can be related to the propagator by the trick of Section 2.5.3. Differentiating the path integral (5.24) with respect to ω, one finds

$$\frac{\partial \ln \mathcal{Z}_0(\beta)}{\partial\omega} = -\hbar\int_{-\beta/2}^{\beta/2}\mathrm{d}t\,\langle z(t)\bar z(t)\rangle_0 = -\hbar\beta\Delta(0).$$

A difficulty then arises because $\Delta(t)$ is not continuous at $t = 0$; this, again, is the $\epsilon(0)$ problem. Integrating, we obtain (using the limit $\beta \to \infty$ to determine the integration constant)

$$\mathcal{Z}_0(\beta) = \frac{\mathrm{e}^{-\hbar\beta\omega\epsilon(0)/2}}{2\sinh(\hbar\beta\omega/2)}.$$

We see that the ambiguity has the nature of a constant shift of the hamiltonian, a natural consequence of the order problem between the operators $a, a^\dagger$. The rule consistent with the normal order (5.3) is to set $\epsilon(0) = -1$, and yields

$$\mathcal{Z}_0(\beta) = \frac{1}{1 - \mathrm{e}^{-\hbar\beta\omega}}\,. \tag{5.33}$$

However, , such a choice is inconvenient for perturbative calculations, because it breaks time reversal symmetry, and it is better to choose $\epsilon(0) = 0$. This corresponds to a symmetric form $\frac{1}{2}(aa^\dagger + a^\dagger a)$ and yields the standard harmonic oscillator. The additional contribution can then be cancelled, if necessary, by shifting physical parameters.

5.2.2 General hamiltonian: one degree of freedom

For a general quantum hamiltonian one can first write the quantum hamiltonian in terms of creation and annihilation operators, commute all creation operators to the left (normal order) and then replace operators by the corresponding classical variables, as explained in Section 5.1. One obtains the matrix elements $\langle z|\,U(t+\varepsilon, t)\,|\bar z\rangle$ at order ε, and following the method of Section 5.2 derives a path integral representation of the form

$$\langle z''|\,U(t'', t')\,|\bar z'\rangle = \int\left[\frac{\mathrm{d}\bar z(t)\mathrm{d}z(t)}{2i\pi}\right]\exp\left[-\mathcal{S}(z,\bar z)\right], \tag{5.34}$$

$$\mathcal{S}\big(z(t),\bar z(t)\big) = -\bar z(t')z(t') + \int_{t'}^{t''}\mathrm{d}t\,\left[-\bar z(t)\dot z(t) + h\big(z(t),\bar z(t)\big)/\hbar\right], \tag{5.35}$$

with the boundary conditions $z(t'') = z''$, $\bar{z}(t') = \bar{z}'$.

Such a path integral can be used to generate a perturbative expansion. The expansion is plagued by singularities which are a reflection of the quantization problem and the order between the quantum operators $a, a^\dagger$. In particular the ill-defined quantity $\Delta(0)$ will appear. As we have already explained, normal order corresponds to the choice of $\epsilon(0) = -1$, which is somewhat inconvenient, and it is preferable to choose $\epsilon(0) = 0$ 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

$$p - i\omega q = -i\sqrt{2\hbar\omega}\,\bar{z}\,, \qquad p + i\omega q = i\sqrt{2\hbar\omega}\,z\,. \tag{5.36}$$

Path integrals, because they involve classical hamiltonians, extend somewhat this correspondence to quantum mechanics. If in the path integral over phase space derived in Chapter 3 (equation (3.9)), we change of variables $(p(t), q(t)) \mapsto (\bar{z}(t), z(t))$, where the variables are related at all times by (5.36), we find a form (5.34,5.35) with

$$h(z, \bar{z}) = H\left(i(z - \bar{z})\sqrt{\hbar\omega/2}, (z + \bar{z})\sqrt{\hbar/2\omega}\right).$$

The differences come only from boundary terms and boundary conditions and this justifies our starting from first principles again in this chapter. But once the modifications are known, one can infer the holomorphic path integral directly from the phase space integral.

Both formalisms have problems generated by ordering operators. Notice, however, that the transformation (5.36) is not always 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

$$\mathcal{Z}(\beta) = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\,\mathrm{e}^{-z\bar{z}}\,\langle z|\,U(\hbar\beta/2, -\hbar\beta/2)\,|\bar{z}\rangle = \int \left[\frac{\mathrm{d}\bar{z}(t)\mathrm{d}z(t)}{2i\pi}\right] \exp\left[-\mathcal{S}(z, \bar{z})\right],$$

$$\mathcal{S}(z, \bar{z}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t\,\left[-\bar{z}(t)\dot{z}(t) + h\big(z(t)\bar{z}(t)\big)\right],$$

with 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, $h(z, \bar{z}) = \omega z\bar{z} + h_{\mathrm{I}}(z, \bar{z})$, and evaluating perturbative terms using Wick's theorem (1.30) together with the two-point function (5.30).

Remark. If the hamiltonian has the form discussed in Chapter 2,

$$H = \tfrac{1}{2}\hat{p}^2 + \tfrac{1}{2}\omega^2\hat{q}^2 + V_{\mathrm{I}}(\hat{q}),$$

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 (5.27) that these 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 (5.29) and (5.30) can be symmetrized in time. After the rescaling, $b(t) \mapsto b(t)\sqrt{\hbar/2\omega}$, one finds

$$\operatorname{tr} U_{\mathrm{G}}(\hbar\beta/2, -\hbar\beta/2) = \frac{1}{2\sinh(\omega\beta/2)} \exp\left[\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}\tau\, b(t)\Delta(t - \tau)b(\tau)\right], \tag{5.37}$$

$$\Delta(t) = \hbar\frac{\cosh\hbar\omega(\beta/2 - |t|)}{2\omega\sinh(\hbar\omega\beta/2)}, \tag{5.38}$$

a result consistent with equation (2.45).

with the boundary conditions $z(t'') = z''$, $\bar{z}(t') = \bar{z}'$.

Such a path integral can be used to generate a perturbative expansion. The expansion is plagued by singularities which are a reflection of the quantization problem and the order between the quantum operators $a, a^\dagger$. In particular the ill-defined quantity $\Delta(0)$ will appear. As we have already explained, normal order corresponds to the choice of $\epsilon(0) = -1$, which is somewhat inconvenient, and it is preferable to choose $\epsilon(0) = 0$ 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

$$p - i\omega q = -i\sqrt{2\hbar\omega}\, \bar{z}\,, \qquad p + i\omega q = i\sqrt{2\hbar\omega}\, z\,. \tag{5.36}$$

Path integrals, because they involve classical hamiltonians, extend somewhat this correspondence to quantum mechanics. If in the path integral over phase space derived in Chapter 3 (equation (3.9)), we change of variables $(p(t), q(t)) \mapsto (\bar{z}(t), z(t))$, where the variables are related at all times by (5.36), we find a form (5.34,5.35) with

$$h(z, \bar{z}) = H\big(i(z - \bar{z})\sqrt{\hbar\omega/2}, (z + \bar{z})\sqrt{\hbar/2\omega}\big).$$

The differences come only from boundary terms and boundary conditions and this justifies our starting from first principles again in this chapter. But once the modifications are known, one can infer the holomorphic path integral directly from the phase space integral.

Both formalisms have problems generated by ordering operators. Notice, however, that the transformation (5.36) is not always 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

$$\mathcal{Z}(\beta) = \int \frac{\mathrm{d}\bar{z}\,\mathrm{d}z}{2i\pi}\, \mathrm{e}^{-z\bar{z}} \left\langle z \right| U(\hbar\beta/2, -\hbar\beta/2) \left| \bar{z} \right\rangle = \int \left[\frac{\mathrm{d}\bar{z}(t)\mathrm{d}z(t)}{2i\pi}\right] \exp\left[-\mathcal{S}(z, \bar{z})\right],$$

$$\mathcal{S}(z, \bar{z}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \left[-\bar{z}(t)\dot{z}(t) + h\big(z(t)\bar{z}(t)\big)\right],$$

with 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, $h(z, \bar{z}) = \omega z\bar{z} + h_{\mathrm{I}}(z, \bar{z})$, and evaluating perturbative terms using Wick's theorem (1.30) together with the two-point function (5.30).

Remark. If the hamiltonian has the form discussed in Chapter 2,

$$H = \tfrac{1}{2}\hat{p}^2 + \tfrac{1}{2}\omega^2\hat{q}^2 + V_{\mathrm{I}}(\hat{q}),$$

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 (5.27) that these 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 (5.29) and (5.30) can be symmetrized in time. After the rescaling, $b(t) \mapsto b(t)\sqrt{\hbar/2\omega}$, one finds

$$\operatorname{tr} U_{\mathrm{G}}(\hbar\beta/2, -\hbar\beta/2) = \frac{1}{2\sinh(\omega\beta/2)} \exp\left[\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}\tau\, b(t)\Delta(t - \tau)b(\tau)\right], \tag{5.37}$$

$$\Delta(t) = \hbar\frac{\cosh\hbar\omega(\beta/2 - |t|)}{2\omega\sinh(\hbar\omega\beta/2)}, \tag{5.38}$$

a result consistent with equation (2.45).

5.2.3 Several degrees of freedom: many-body interpretation

A hamiltonian H_0, sum of N independent harmonic oscillators can be expressed in terms of N complex variables as

$$H_0 = \sum_{i=1}^{N} \hbar\omega_i z_i \frac{\partial}{\partial z_i}\,.$$

Its energy levels take the form $\hbar\sum_i n_i\omega_i$ that, due to the additive character of the spectrum of the harmonic oscillator, has also the interpretation of the total energy of $\sum_i n_i$ independent particles. In this different interpretation one-particle states are associated with the various energies $\hbar\omega_i$ and the integer n_i is the number (also called occupation number) of particles which are in the state i.

A vector $\psi(\mathbf{z})$ in the holomorphic formalism now has the expansion

$$\psi(\mathbf{z}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i_1,i_2,\ldots,i_n} \psi_{i_1 i_2\ldots i_n} z_{i_1} z_{i_2} \ldots z_{i_n}\,. \tag{5.39}$$

The coefficient $\psi_{i_1 i_2\ldots i_n}$ is the component of $\psi(\mathbf{z})$ on an eigenstate of energy $\omega_{i_1} + \omega_{i_2} + \cdots + \omega_{i_n}$. In the many-body interpretation $\psi_{i_1 i_2\ldots i_n}$ is the component of $\psi(\mathbf{z})$ on an n-particle state and, since $\psi_{i_1 i_2\ldots i_n}$ is symmetric in the n indices $i_1, i_2, \ldots, i_n$, these particles obey the Bose–Einstein statistics. The holomorphic formalism, therefore, allows us to describe general boson states which are linear combinations of states with an arbitrary number of particles.

The partition function of free bosons that can occupy N different states of energy $\hbar\omega_i$ is then given by the path integral

$$\mathcal{Z}_0(\beta) = \int \left[\frac{\mathrm{d}z(t)\mathrm{d}\bar{z}(t)}{2i\pi}\right] \exp\left[-\mathcal{S}_0(\bar{\mathbf{z}},\mathbf{z})\right], \tag{5.40}$$

$$\mathcal{S}_0(\bar{\mathbf{z}},\mathbf{z}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \sum_i \bar{z}_i(t)\left[-\dot{z}(t) + \hbar\omega_i z_i(t)\right]. \tag{5.41}$$

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 as will be discussed in Section 5.5.

5.3 Path Integrals with Fermions

The derivation of a path integral representation for the statistical operator of fermion systems follows closely the method of Sections 5.1, 5.2, complex variables being replaced by Grassmann variables. Grassmann algebras together with the rules of differentiation and integration have been described in Sections 1.4–1.7.

5.3.1 Fermions and complex vector spaces

Complex conjugation. We again consider a Grassmann algebra $\mathfrak{C}$ with two sets of generators $\{\theta_i, \bar{\theta}_i\}$, $i = 1, \ldots, n$. In this algebra, the analogue of the complex conjugation is the hermitian conjugation of operators,

$$\theta_i^\dagger = \bar{\theta}_i\,,\quad \bar{\theta}_i^\dagger = \theta_i\,,\quad (A_1A_2)^\dagger = A_2^\dagger A_1^\dagger \quad \forall\, A_1, A_2 \in \mathfrak{C}\,. \tag{5.42}$$

With this definition, the measure $\mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i$ is invariant. A quadratic form is invariant if

$$\sum_{i,j=1}^{n} (\bar{\theta}_i M_{ij} \theta_j)^\dagger = \sum_{i,j=1}^{n} \bar{\theta}_j \overline{M}_{ij} \theta_i = \sum_{i,j=1}^{n} \bar{\theta}_i M_{ij}^\dagger \theta_j ,$$

that is, if the matrix $\mathbf{M}$ is hermitian.

Analytic Grassmann functions. "Analytic" Grassmann functions are functions of θ_i variables only:

$$\frac{\partial f}{\partial \bar{\theta}_i} = 0 , \quad \forall i ,$$

and thus elements of a subalgebra $\mathfrak{A}$.

Analytic Grassmann functions of one variable form a complex vector space of dimension two. This is in direct relation with Fermi–Dirac statistics: a constant corresponds to an empty state, the function θ to a state occupied by one fermion.

Scalar product of analytic functions. In the same way as for analytic complex functions (equation (5.7)) we introduce a scalar product between analytic Grassmann functions. The scalar product of two functions f, g is defined by

$$(f, g) = \int \left(\prod_i \mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i\right) \exp\left(\sum_i \bar{\theta}_i \theta_i\right) \; f^\dagger(\theta) \; g(\theta). \tag{5.43}$$

One verifies that the scalar product defines a positive norm. In the simple example of one variable θ one finds

$$(f, g) = \int \mathrm{d}\theta \, \mathrm{d}\bar{\theta} \; \mathrm{e}^{\bar{\theta}\theta} \; \overline{f(\theta)} \; g(\theta).$$

A function $f(\theta)$ is automatically an affine function of θ, thus

$$f(\theta) = a + b\theta , \; \Rightarrow \; (f, f) = \bar{a}a + \bar{b}b \geq 0 .$$

The scalar product (5.43) allows to construct an orthogonal basis: one verifies that all independent monomials in $\mathfrak{A}$ are orthogonal.

The δ-function. In Grassmann algebras the role of the Dirac δ-function is played by the function θ itself since

$$\int \mathrm{d}\theta \, \theta f(\theta) = f(0),$$

in which $f(0)$ means the constant part of the affine function $f(\theta)$. In general, we shall use an integral representation of $\delta(\theta)$ analogous to a Fourier transform:

$$f(0) = \int \mathrm{d}\theta \, \mathrm{d}\bar{\theta} \; \mathrm{e}^{\bar{\theta}\theta} \; f(\theta). \tag{5.44}$$

With this definition, the measure $\mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i$ is invariant. A quadratic form is invariant if

$$\sum_{i,j=1}^{n} (\bar{\theta}_i M_{ij} \theta_j)^\dagger = \sum_{i,j=1}^{n} \bar{\theta}_j \overline{M}_{ij} \theta_i = \sum_{i,j=1}^{n} \bar{\theta}_i M_{ij}^\dagger \theta_j \,,$$

that is, if the matrix $\mathbf{M}$ is hermitian.

Analytic Grassmann functions. "Analytic" Grassmann functions are functions of θ_i variables only:

$$\frac{\partial f}{\partial \bar{\theta}_i} = 0 \,, \quad \forall i \,,$$

and thus elements of a subalgebra $\mathfrak{A}$.

Analytic Grassmann functions of one variable form a complex vector space of dimension two. This is in direct relation with Fermi–Dirac statistics: a constant corresponds to an empty state, the function θ to a state occupied by one fermion.

Scalar product of analytic functions. In the same way as for analytic complex functions (equation (5.7)) we introduce a scalar product between analytic Grassmann functions. The scalar product of two functions f, g is defined by

$$(f, g) = \int \Big(\prod_i \mathrm{d}\theta_i \mathrm{d}\bar{\theta}_i \Big) \exp \Big(\sum_i \bar{\theta}_i \theta_i \Big) \; f^\dagger(\theta) \; g(\theta). \tag{5.43}$$

One verifies that the scalar product defines a positive norm. In the simple example of one variable θ one finds

$$(f, g) = \int \mathrm{d}\theta \, \mathrm{d}\bar{\theta} \; \mathrm{e}^{\bar{\theta}\theta} \; \overline{f(\theta)} \; g(\theta).$$

A function $f(\theta)$ is automatically an affine function of θ, thus

$$f(\theta) = a + b\theta \,, \; \Rightarrow \; (f, f) = \bar{a}a + \bar{b}b \geq 0 \,.$$

The scalar product (5.43) allows to construct an orthogonal basis: one verifies that all independent monomials in $\mathfrak{A}$ are orthogonal.

The δ-function. In Grassmann algebras the role of the Dirac δ-function is played by the function θ itself since

$$\int \mathrm{d}\theta \, \theta f(\theta) = f(0),$$

in which $f(0)$ means the constant part of the affine function $f(\theta)$. In general, we shall use an integral representation of $\delta(\theta)$ analogous to a Fourier transform:

$$f(0) = \int \mathrm{d}\theta \, \mathrm{d}\bar{\theta} \; \mathrm{e}^{\bar{\theta}\theta} \, f(\theta). \tag{5.44}$$

5.3.2 Operator algebra

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:

$$\mathcal{I}(\theta,\bar\theta)=\prod_i(1+\theta_i\bar\theta_i)=\exp\left(-\sum_i\bar\theta_i\theta_i\right). \tag{5.45}$$

A direct verification relies on the identity (5.44)

$$\begin{aligned}\int\prod_i \mathrm{d}\theta'_i\mathrm{d}\bar\theta'_i\,\mathcal{I}(\theta,\bar\theta')\exp\left(\sum_i\bar\theta'_i\theta'_i\right)f(\theta')&=\int\prod_i \mathrm{d}\theta'_i\mathrm{d}\bar\theta'_i\exp\left(\sum_i\bar\theta'_i(\theta'_i-\theta_i)\right)f(\theta')\\&=f(\theta).\end{aligned} \tag{5.46}$$

Operator algebra. We have already introduced the algebra of operators $\{\theta_i,\partial/\partial\theta_i\}$ (see (1.48)) acting by multiplication and differentiation on the algebra $\mathfrak{A}$ of Grassmann functions. 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 (5.46) to associate to each element of this operator algebra a kernel,

$$\theta_{i_1}\theta_{i_2}\ldots\theta_{i_p}\frac{\partial}{\partial\theta_{j_1}}\frac{\partial}{\partial\theta_{j_2}}\cdots\frac{\partial}{\partial\theta_{j_q}}\mathcal{I}(\theta,\bar\theta)=\theta_{i_1}\theta_{i_2}\ldots\theta_{i_p}\bar\theta_{j_1}\bar\theta_{j_1}\ldots\bar\theta_{j_q}\mathcal{I}(\theta,\bar\theta), \tag{5.47}$$

which belongs to the Grassmann algebra $\mathfrak{C}$.

Such an operator $O(\theta,\partial/\partial\theta)$ defined by its kernel $\mathcal{O}(\theta,\bar\theta)$

$$O(\theta,\partial/\partial\theta)\mapsto\mathcal{O}(\theta,\bar\theta)\equiv\langle\theta|\,\mathcal{O}\,|\bar\theta\rangle=O(\theta,\bar\theta)\mathcal{I}(\theta,\bar\theta), \tag{5.48}$$

then acts on a function as

$$(Of)(\theta)=\int\prod_i \mathrm{d}\theta'_i\mathrm{d}\bar\theta'_i\,\mathcal{O}(\theta,\bar\theta')\exp\left(\sum_i\bar\theta'_i\theta'_i\right)f(\theta'). \tag{5.49}$$

The kernel associated with the product $\mathcal{O}_2\mathcal{O}_1$ is given by

$$\langle\theta|\,\mathcal{O}_2\mathcal{O}_1\,|\bar\theta\rangle=\int\prod_i \mathrm{d}\theta'_i\mathrm{d}\bar\theta'_i\,\langle\theta|\,\mathcal{O}_2\,|\bar\theta'\rangle\exp\left(\sum_i\bar\theta'_i\theta'_i\right)\langle\theta'|\,\mathcal{O}_1\,|\bar\theta\rangle. \tag{5.50}$$

Trace. The trace of an operator is

$$\operatorname{tr}\mathcal{O}=\int\prod_i \mathrm{d}\bar\theta_i\mathrm{d}\theta_i\exp\left(-\sum_i\bar\theta_i\theta_i\right)\langle\theta|\,\mathcal{O}\,|\bar\theta\rangle. \tag{5.51}$$

Note the difference of sign in the exponentials between expressions (5.50) and (5.51). One can verify that the sign in (5.51) is consistent with the cyclic property of the trace.

Hermitian conjugation of operators. The hermitian conjugation for operator kernels is the hermitian conjugation in the algebra $\mathfrak{C}$ (equations (5.42)).

5.3.3 Fermion states and path integrals

One-state hamiltonian and fermions. We now generalize the considerations of Section 5.2.3 to fermions. The Hilbert space for an arbitrary number of particles obeying the Fermi–Dirac statistics (fermions) that can occupy only one state of energy $\hbar\omega$ reduces to a two-dimensional complex vector space, as a consequence of the Pauli principle: the state can only be empty (zero-particle state) or occupied once. Path integrals are certainly not required to solve the corresponding matrix Schrödinger equation. However, the solution in the form of a path integral is useful because it can easily be generalized to an arbitrary number of possible fermion states.

The vector corresponding to a linear superposition of a zero-particle and one-particle states can be represented by an affine function

$$\psi(\theta) = \psi_0 + \psi_1\theta\,, \quad \theta \in \mathfrak{A}\,,$$

where by choosing for θ a generator of a Grassmann algebra we enforce the condition that a state can only be occupied once since $\theta^2 = 0$.

The eigenvalues of the hamiltonian H_0 are 0 for the zero-particle state and $\hbar\omega$ for the one-particle state. It can thus be represented by

$$H_0 = \hbar\omega\theta\frac{\partial}{\partial\theta}\,. \tag{5.52}$$

The corresponding kernel is

$$\langle\theta|H_0|\bar\theta\rangle = \hbar\omega\theta\frac{\partial}{\partial\theta}\,\mathrm{e}^{-\bar\theta\theta} = \hbar\omega\theta\bar\theta\,\mathrm{e}^{-\bar\theta\theta} = -\hbar\omega\bar\theta\theta\,.$$

The matrix elements of the operator $U_0(t) = \mathrm{e}^{-tH_0/\hbar}$ are easily obtained:

$$\langle\theta|\,U_0(t)\,|\bar\theta\rangle = \mathrm{e}^{-\bar\theta\theta\,\mathrm{e}^{-\omega t}}\,. \tag{5.53}$$

With our definition both kernels are invariant under complex conjugation. From the explicit expression (5.53) and the definition (5.51) we infer the partition function, trace of $U_0(\hbar\beta)$,

$$\operatorname{tr} U_0(\hbar\beta) = \int \mathrm{d}\bar\theta\,\mathrm{d}\theta\;\mathrm{e}^{-\bar\theta\theta}\,\mathrm{e}^{-\bar\theta\theta\,\mathrm{e}^{-\hbar\omega}} = 1 + \mathrm{e}^{-\hbar\omega\beta}\,. \tag{5.54}$$

Remark. In the conventional operator formalism one introduces operators a and $a^\dagger$, which annihilate and create fermions, with the anticommutation relations

$$a^2 = a^{\dagger 2} = 0\,, \qquad aa^\dagger + a^\dagger a = 1\,. \tag{5.55}$$

Equations (1.48) exhibit a realization of the commutation relations (5.55) as multiplication and differentiation operations on functions of Grassmann variables with the correspondence

$$\theta \mapsto a^\dagger\,, \quad \frac{\partial}{\partial\theta} \mapsto a\,.$$

The hamiltonian then reads

$$H_0 = \hbar\omega a^\dagger a\,. \tag{5.56}$$

Path integral. We now construct a path integral representation for the statistical operator $U_0(t)$. We expand the exact expression (5.53) for a time ε small:

$$\langle \theta \left| U_0(\varepsilon) \right| \bar{\theta} \rangle = \exp\left[-\bar{\theta}\theta \left(1 - \omega\varepsilon\right) + O\left(\varepsilon^2\right)\right]. \tag{5.57}$$

The group property in the form (5.50) allows us to write the statistical operator at finite time:

$$\langle \theta'' \left| U_0(t'', t') \right| \bar{\theta}' \rangle = \lim_{n\to\infty} \int \prod_{k=1}^{n-1} \mathrm{d}\theta_k \mathrm{d}\bar{\theta}_k \, \exp\left[-\mathcal{S}_\varepsilon(\theta, \bar{\theta})\right], \tag{5.58}$$

with

$$\mathcal{S}_\varepsilon(\theta, \bar{\theta}) = \bar{\theta}_0\theta_1 + \sum_{k=1}^{n-1} \bar{\theta}_k \left(\theta_{k+1} - \theta_k\right) - \varepsilon \sum_{k=0}^{n-1} \omega\bar{\theta}_k\theta_{k+1}\,, \tag{5.59}$$

$\varepsilon = (t'' - t')/n$, and the definitions

$$\bar{\theta}_0 = \bar{\theta}', \quad \theta_n = \theta''. \tag{5.60}$$

The formal large n limit takes the form of a path integral representation for the matrix elements of $U_0(t)$:

$$\langle \theta'' \left| U_0(t'', t') \right| \bar{\theta}' \rangle = \int_{\bar{\theta}(t')=\bar{\theta}'}^{\theta(t'')=\theta''} \left[\mathrm{d}\theta(t)\mathrm{d}\bar{\theta}(t)\right] \exp\left[-\mathcal{S}_0(\theta, \bar{\theta})\right], \tag{5.61}$$

with

$$\mathcal{S}_0(\theta, \bar{\theta}) = \int_{t'}^{t''} \mathrm{d}t\, \bar{\theta}(t) \left[\dot{\theta}(t) - \omega\theta(t)\right] + \bar{\theta}(t')\theta(t'). \tag{5.62}$$

Partition function. As in the holomorphic case, to calculate the trace, using equation (5.51), we first return to the discretized expression. We now find

$$\mathcal{S}_\varepsilon(\theta, \bar{\theta}) = \bar{\theta}_0(\theta_n + \theta_0) + \sum_{k=1}^{n} \left[\bar{\theta}_{k-1}\left(\theta_k - \theta_{k-1}\right) - \hbar\omega\varepsilon\bar{\theta}_{k-1}\theta_k\right], \tag{5.63}$$

with the integration measure

$$\prod_{k=1}^{n-1} \mathrm{d}\theta_k \mathrm{d}\bar{\theta}_k \mathrm{d}\bar{\theta}_0 \mathrm{d}\theta_n\,.$$

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

$$\mathcal{S}_\varepsilon(\theta, \bar{\theta}) = \sum_{k=1}^{n} \left[\bar{\theta}_{k-1}\left(\theta_k - \bar{\theta}_{k-1}\right) - \hbar\omega\varepsilon\bar{\theta}_k\theta_{k-1}\right].$$

In the continuum limit we thus obtain

$$\mathcal{Z}_0(\beta) = \operatorname{tr} U(\hbar\beta/2, -\hbar\beta/2) = \int \left[\mathrm{d}\bar{\theta}(t)\mathrm{d}\theta(t)\right] \exp\left[-\mathcal{S}_0(\theta, \bar{\theta})\right] \tag{5.64}$$

with

$$\mathcal{S}_0(\theta,\bar{\theta}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t\,\bar{\theta}(t)\left[\dot{\theta}(t) - \hbar\omega\theta(t)\right] \tag{5.65}$$

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 consider a general gaussian integral that will be useful later to generate 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 (5.64), we then replace the action $\mathcal{S}_0(\theta,\bar{\theta})$ by $\mathcal{S}_{\mathrm{G}}(\theta,\bar{\theta})$:

$$\mathcal{S}_{\mathrm{G}}(\theta,\bar{\theta}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t\left\{\bar{\theta}(t)\left[\dot{\theta}(t) - \hbar\omega\theta(t)\right] - \bar{\eta}(t)\theta(t) - \bar{\theta}(t)\eta(t)\right\}. \tag{5.66}$$

The path integral can be calculated by a simple extension of the method used in the case of the integral (1.70). One looks for the saddle point, solution of the equations obtained by varying $\bar{\theta}(t)$ and $\theta(t)$, respectively:

$$\dot{\theta}(t) - \hbar\omega\theta(t) - \eta(t) = 0\,, \tag{5.67a}$$

$$\dot{\bar{\theta}}(t) + \hbar\omega\bar{\theta}(t) + \bar{\eta}(t) = 0\,. \tag{5.67b}$$

The solution of equation (5.67a) with anti-periodic boundary conditions can be written as

$$\theta(t) = -\int_{-\beta/2}^{\beta/2} \Delta(u-t)\eta(u)\mathrm{d}u \tag{5.68}$$

with

$$\Delta(t) = \tfrac{1}{2}\,\mathrm{e}^{-\hbar\omega t}\left[\epsilon(t) + \tanh(\hbar\omega\beta/2)\right] \tag{5.69}$$

($\epsilon(t)$ is the sign function, $\epsilon(t) = 1$ for $t > 0$, $\epsilon(-t) = -\epsilon(t)$).

The function $\Delta(t)$ is the solution of the differential equation

$$\dot{\Delta} + \hbar\omega\Delta = \delta(t)$$

with anti-periodic boundary conditions $\Delta(\beta/2) = -\Delta(-\beta/2)$.

The explicit solution of equation (5.67b) is not needed. To calculate the action at the saddle point, we integrate by parts $\bar{\theta}\dot{\theta}$, use equation (5.67b) and obtain

$$\mathcal{S}_{\mathrm{G}} = -\int_{-\beta/2}^{\beta/2} \mathrm{d}t\,\bar{\eta}(t)\theta(t) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t\,\mathrm{d}u\,\bar{\eta}(t)\Delta(u-t)\eta(u). \tag{5.70}$$

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,

$$\operatorname{tr} U_{\mathrm{G}}(\hbar\beta/2, -\hbar\beta/2) = \left(1 + \mathrm{e}^{-\hbar\omega\beta}\right)\exp\left[-\int_{-\beta/2}^{\beta/2} \mathrm{d}t\,\mathrm{d}u\,\bar{\eta}(t)\Delta(u-t)\eta(u)\right]. \tag{5.71}$$

As in the complex case, correlation functions are obtained by differentiating the path integral with the action (5.66) and the expression (5.71) with respect to η and $\bar{\eta}$. For example,

$$\langle\bar{\theta}(t_2)\theta(t_1)\rangle = -\mathcal{Z}_0^{-1}(\beta)\,\frac{\delta}{\delta\eta(t_2)}\frac{\delta}{\delta\bar{\eta}(t_1)}\operatorname{tr} U_{\mathrm{G}}(\hbar\beta/2, -\hbar\beta/2)\Big|_{\eta=\bar{\eta}=0} = \Delta(t_2 - t_1). \tag{5.72}$$

Remark. The path integral with periodic boundary conditions appears instead in the calculation of $\operatorname{tr}(-1)^F \mathrm{e}^{-\beta H}$, in which F is the fermion number, and is obtained by integrating expression (5.61) with $\mathrm{e}^{\theta''\bar{\theta}'}$.

The gaussian integration and the $\epsilon(0)$ problem. Again as in the complex case of Section 5.2 we can calculate, up to a numerical factor, the partition function $\mathcal{Z}_0(\beta)$ by relating it to $\Delta(t)$. From (5.64) one infers

$$\frac{\partial \ln \mathcal{Z}_0(\beta)}{\partial \omega} = \hbar \int \left\langle \bar{\theta}(t)\theta(t) \right\rangle = \hbar\beta\Delta(0).$$

Again the function $\Delta(t)$ is not continuous at $t=0$, and thus the result is ill-defined. The choice $\epsilon(0) = -1$ is consistent with the normal order. Integrating, we then obtain

$$\mathcal{Z}_0(\beta) = 1 + \mathrm{e}^{-\hbar\beta\omega}. \tag{5.73}$$

Another choice corresponds to a constant shift of the hamiltonian, natural consequence of the order problem between the operators $\theta, \partial/\partial\theta$ or $a, a^\dagger$ (defined in (5.55)). For example, if we choose $\epsilon(0) = 0$ we obtain instead

$$\mathcal{Z}_0(\beta) = 2\cosh(\hbar\beta\omega/2),$$

which corresponds to the hamiltonian

$$H_0 = \hbar\omega[a^\dagger, a]/2\,.$$

5.3.4 Fermions and many-body theory

The method can be generalized to N possible states of energies $\hbar\omega_i$, $1 \le i \le N$. One now introduces a Grassmann algebra with generators θ_i and considers vectors represented by

$$\psi(\theta) = \sum_{n=0}^{N} \frac{1}{n!} \sum_{i_1, i_2, \ldots, i_n} \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n}\,.$$

Since the product of Grassmann generators is antisymmetric, the coefficients $\psi_{i_1 i_2 \ldots i_n}$, amplitude of the vector on an n particle state, can be chosen antisymmetric in all indices and this enforces the Pauli principle for fermions.

A hamiltonian H_0 for free fermions that can occupy N possible states of energies $\hbar\omega_i$ is represented by

$$H_0 = \sum_{i=1}^{N} \hbar\omega_i \theta_i \frac{\partial}{\partial\theta_i}\,,$$

as one verifies by explicit calculation,

$$H_0 \sum_{i_1, i_2, \ldots, i_n} \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n} = \hbar \sum_{i_1, i_2, \ldots, i_n} (\omega_{i_1} + \omega_{i_2} + \cdots + \omega_{i_n}) \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n}\,.$$

The corresponding partition function has a path integral representation with an action which is a straightforward generalization of the form (5.65)

$$\mathcal{S}_0(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \sum_i \bar{\theta}_i(t) \left[\dot{\theta}_i(t) - \hbar\omega_i\theta_i(t)\right] \tag{5.74}$$

Remark. The path integral with periodic boundary conditions appears instead in the calculation of $\mathrm{tr}(-1)^F \mathrm{e}^{-\beta H}$, in which F is the fermion number, and is obtained by integrating expression (5.61) with $\mathrm{e}^{\theta''\bar{\theta}'}$.

The gaussian integration and the $\epsilon(0)$ problem. Again as in the complex case of Section 5.2 we can calculate, up to a numerical factor, the partition function $\mathcal{Z}_0(\beta)$ by relating it to $\Delta(t)$. From (5.64) 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 ill-defined. The choice $\epsilon(0) = -1$ is consistent with the normal order. Integrating, we then obtain

$$\mathcal{Z}_0(\beta) = 1 + \mathrm{e}^{-\hbar\beta\omega}. \tag{5.73}$$

Another choice corresponds to a constant shift of the hamiltonian, natural consequence of the order problem between the operators $\theta, \partial/\partial\theta$ or $a, a^\dagger$ (defined in (5.55)). For example, if we choose $\epsilon(0) = 0$ we obtain instead

$$\mathcal{Z}_0(\beta) = 2\cosh(\hbar\beta\omega/2),$$

which corresponds to the hamiltonian

$$H_0 = \hbar\omega[a^\dagger, a]/2.$$

5.3.4 Fermions and many-body theory

The method can be generalized to N possible states of energies $\hbar\omega_i$, $1 \le i \le N$. One now introduces a Grassmann algebra with generators θ_i and considers vectors represented by

$$\psi(\theta) = \sum_{n=0}^{N} \frac{1}{n!} \sum_{i_1, i_2, \ldots, i_n} \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n}.$$

Since the product of Grassmann generators is antisymmetric, the coefficients $\psi_{i_1 i_2 \ldots i_n}$, amplitude of the vector on an n particle state, can be chosen antisymmetric in all indices and this enforces the Pauli principle for fermions.

A hamiltonian H_0 for free fermions that can occupy N possible states of energies $\hbar\omega_i$ is represented by

$$H_0 = \sum_{i=1}^{N} \hbar\omega_i\theta_i \frac{\partial}{\partial\theta_i},$$

as one verifies by explicit calculation,

$$H_0 \sum_{i_1, i_2, \ldots, i_n} \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n} = \hbar \sum_{i_1, i_2, \ldots, i_n} (\omega_{i_1} + \omega_{i_2} + \cdots + \omega_{i_n}) \psi_{i_1 i_2 \ldots i_n} \theta_{i_1}\theta_{i_2}\ldots\theta_{i_n}.$$

The corresponding partition function has a path integral representation with an action which is a straightforward generalization of the form (5.65)

$$\mathcal{S}_0(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \sum_i \bar{\theta}_i(t) \left[\dot{\theta}_i(t) - \hbar\omega_i\theta_i(t)\right] \tag{5.74}$$

and with *anti-periodic boundary conditions* $\boldsymbol{\theta}(-\beta/2) = -\boldsymbol{\theta}(\beta/2)$, $\bar{\boldsymbol{\theta}}(-\beta/2) = -\bar{\boldsymbol{\theta}}(\beta/2)$.

Interactions. Interactions between particles can be introduced by considering hamiltonians which are more general elements of the operator algebra considered in Section 5.3.2. 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

$$H(\boldsymbol{\theta}, \partial/\partial\boldsymbol{\theta}) \mapsto H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})\mathcal{I}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}).$$

The equation (5.57) generalizes in the form

$$\langle \boldsymbol{\theta} \left| U(t) \right| \bar{\boldsymbol{\theta}} \rangle = \exp\left[-\sum_i \theta_i \bar{\theta}_i - t\, H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) + O\left(t^2\right) \right]. \tag{5.75}$$

At finite time, a path integral representation follows:

$$\langle \boldsymbol{\theta}'' \left| U(t'', t') \right| \bar{\boldsymbol{\theta}}' \rangle = \int_{\bar{\boldsymbol{\theta}}(t')=\bar{\boldsymbol{\theta}}'}^{\boldsymbol{\theta}(t'')=\boldsymbol{\theta}''} \left[\mathrm{d}\boldsymbol{\theta}(t) \mathrm{d}\bar{\boldsymbol{\theta}}(t) \right] \exp\left[-\mathcal{S}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) \right] \tag{5.76}$$

with

$$\mathcal{S}(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = \sum_i \bar{\theta}_i(t')\theta_i(t') + \int_{t'}^{t''} \mathrm{d}t \left\{ \sum_i \bar{\theta}_i(t)\dot{\theta}_i(t) + H\left[\boldsymbol{\theta}(t), \bar{\boldsymbol{\theta}}(t) \right] \right\}. \tag{5.77}$$

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 (5.77) in perturbation theory. However, again problems due to operator ordering arise. In particular, as in the case of the holomorphic path integral, perturbative calculations involve $\epsilon(0)$. The correct ansatz, consistent with our construction, is to again set $\epsilon(0) = -1$. The more convenient, left—right symmetric choice $\epsilon(0) = 0$ requires an appropriate modification of the function $H(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}})$.

Remark. Alternatively, one can introduce creation and annihilation operators for fermions

$$a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger = a_i a_j + a_j a_i = 0 \quad \text{and} \quad a_i^\dagger a_j + a_j a_i^\dagger = \delta_{ij}\,. \tag{5.78}$$

The relations (1.48) realize the anticommutation relations of the operators $\{a_i^\dagger, a_j\}$:

$$a_i^\dagger \mapsto \theta_i\,, \quad a_i \mapsto \partial/\partial\theta_i \mapsto \bar{\theta}_i\,.$$

A general hamiltonian H takes the form

$$H = H(\mathbf{a}^\dagger, \mathbf{a}) \mapsto H(\boldsymbol{\theta}, \partial/\partial\boldsymbol{\theta}).$$

5.4 Quantum Statistical Physics: Fixed Number of Particles

In this section, we consider n-particle systems as relevant for Quantum Statistical Physics. We first briefly recall how the density matrix at equilibrium for a fixed number of identical particles can be obtained from a path integral, simple extension of the form derived in Chapter 2, where only the Pauli principle has to be added. In the next section we calculate the partition function in the grand canonical formulation with a formalism that generalizes directly the holomorphic path integral.

We consider the quantum hamiltonian H_n for a system of n identical non-relativistic quantum particles of mass m:

$$H_n = \sum_{i=1}^{n} \frac{1}{2m}\hat{p}_i^2 + V(\hat{q}_1, \ldots, \hat{q}_n), \tag{5.79}$$

where $V(\hat{q}_1, \ldots, \hat{q}_n)$ is a symmetric function of the n variables.

The partition function is still the trace of the statistical operator $\mathrm{e}^{-\beta H}$. However, in the trace the statistical properties of the quantum particles have to be taken into account. For simplicity we assume that the particles have no internal degrees of freedom like spin. Then the corresponding wave functions $\psi_n(q_1, q_2, \ldots, q_n)$ are either totally symmetric for bosons, or antisymmetric for fermions, in the positions q_i. 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 still be expressed as a path integral,

$$\mathcal{Z}(n, \beta) = \int [\mathrm{d}q_i(t)]\, \mathrm{e}^{-\mathcal{S}(q)/\hbar}, \tag{5.80}$$

with

$$\mathcal{S}(q)/\hbar = \int_0^\beta \mathrm{d}t \left[\sum_i \frac{1}{2m\,\hbar^2}(\dot{q}_i)^2 + V\bigl(q(t)\bigr)\right], \tag{5.81}$$

but averaged over $n!$ different boundary conditions corresponding to all permutations P.

$$q_i(\beta) = q_{P(i)}(0).$$

Bosons and fermions differ by a sign: in the case of fermions each contribution in the sum has to be multiplied by a factor $\epsilon(P)$ which is the signature of the permutation.

This formalism is quite useful for models where the n-body problem can be solved exactly, but more cumbersome in general. Often it is more convenient to employ, instead, a formalism where the number of particles can vary and is fixed only on average.

5.5 The Bose Gas. Functional Integrals

We discuss now the thermodynamics of a gas of quantum particles in a formalism (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 the chemical potential. In this section, we consider non-relativistic massive particles obeying the Bose–Einstein statistics. We use a representation of bosons often referred to as second quantization. We derive an expression for the partition function which has the form of a functional integral, a rather straightforward generalization of the holomorphic path integral of Section 5.2.3.

5.4 Quantum Statistical Physics: Fixed Number of Particles

In this section, we consider n-particle systems as relevant for Quantum Statistical Physics. We first briefly recall how the density matrix at equilibrium for a fixed number of identical particles can be obtained from a path integral, simple extension of the form derived in Chapter 2, where only the Pauli principle has to be added. In the next section we calculate the partition function in the grand canonical formulation with a formalism that generalizes directly the holomorphic path integral.

We consider the quantum hamiltonian H_n for a system of n identical non-relativistic quantum particles of mass m:

$$H_n = \sum_{i=1}^{n} \frac{1}{2m}\hat{p}_i^2 + V(\hat{q}_1, \ldots, \hat{q}_n), \tag{5.79}$$

where $V(\hat{q}_1, \ldots, \hat{q}_n)$ is a symmetric function of the n variables.

The partition function is still the trace of the statistical operator $\mathrm{e}^{-\beta H}$. However, in the trace the statistical properties of the quantum particles have to be taken into account. For simplicity we assume that the particles have no internal degrees of freedom like spin. Then the corresponding wave functions $\psi_n(q_1, q_2, \ldots, q_n)$ are either totally symmetric for bosons, or antisymmetric for fermions, in the positions q_i. 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 still be expressed as a path integral,

$$\mathcal{Z}(n,\beta) = \int [\mathrm{d}q_i(t)]\, \mathrm{e}^{-\mathcal{S}(q)/\hbar}, \tag{5.80}$$

with

$$\mathcal{S}(q)/\hbar = \int_0^\beta \mathrm{d}t \left[\sum_i \frac{1}{2m\,\hbar^2}(\dot{q}_i)^2 + V\bigl(q(t)\bigr)\right], \tag{5.81}$$

but averaged over $n!$ different boundary conditions corresponding to all permutations P.

$$q_i(\beta) = q_{P(i)}(0).$$

Bosons and fermions differ by a sign: in the case of fermions each contribution in the sum has to be multiplied by a factor $\epsilon(P)$ which is the signature of the permutation.

This formalism is quite useful for models where the n-body problem can be solved exactly, but more cumbersome in general. Often it is more convenient to employ, instead, a formalism where the number of particles can vary and is fixed only on average.

5.5 The Bose Gas. Functional Integrals

We discuss now the thermodynamics of a gas of quantum particles in a formalism (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 the chemical potential. In this section, we consider non-relativistic massive particles obeying the Bose–Einstein statistics. We use a representation of bosons often referred to as second quantization. We derive an expression for the partition function which has the form of a functional integral, a rather straightforward generalization of the holomorphic path integral of Section 5.2.3.

5.5.1 Fock's space and hamiltonian

We denote by $\mathfrak{H}_n$ the Hilbert space of n-particle boson states in d space dimensions and H_n the corresponding hamiltonian,

$$H_n = T_n + \mathcal{V}_n(x_1, \ldots, x_n), \tag{5.82}$$

where T_n is the kinetic term,

$$T_n = -\frac{\hbar^2}{2m} \sum_{i=1}^{n} \nabla_i^2 , \tag{5.83}$$

$\nabla_i \equiv \nabla_{x_i}$, and the potential $\mathcal{V}_n(x_1, \ldots, x_n)$ is a symmetric function. We restrict ourselves in what follows to potentials that are sums of one-body and two-body potentials:

$$\mathcal{V}_n(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} V_1(x_i) + \sum_{i<j\le n} V_2(x_i, x_j), \tag{5.84}$$

$(V(x,y) = V(y,x))$ because the generalization is simple.

We again assume the absence of internal degrees of freedom and denote the n-particle wave function by $\psi_n(x_1, \ldots, x_n)$, a symmetric function of its arguments due to Bose symmetry. The space $\mathfrak{H}_n$ is thus the space of square integrable symmetric functions.

We now consider the direct sum $\oplus\mathfrak{H}_n$, $n = 1, \ldots, \infty$ of Hilbert spaces. Such a space obtained by linear superposition of n-particle states and Cauchy completion is called a 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:

$$\mathbf{H} = \mathbf{T} + \mathbf{V}_1 + \mathbf{V}_2 . \tag{5.85}$$

Generating functional of wave functions and hamiltonian. Since the wave functions are symmetric, we introduce a generating functional $\Psi(\varphi)$ of n-particle wave functions (see Section 1.9),

$$\Psi(\varphi) = \sum_{n=0} \frac{1}{n!} \left(\int \prod_i \mathrm{d}^d x_i \, \varphi(x_i) \right) \psi_n(x_1, \ldots, x_n). \tag{5.86}$$

We then represent the kinetic term $\mathbf{T}$ and the potential terms $\mathbf{V}_1$ and $\mathbf{V}_2$ as operators acting on Ψ by functional differentiation.

We first calculate

$$\int \mathrm{d}^d x \, \varphi(x) \nabla_x^2 \frac{\delta}{\delta\varphi(x)} \Psi(\varphi) = \int \mathrm{d}^d x \, \varphi(x) \nabla_x^2 \sum_n \frac{1}{(n-1)!} \times \int \left(\prod_{i<n} \mathrm{d}^d x_i \, \varphi(x_i) \right) \psi_n(x_1, \ldots, x_{n-1}, x).$$

The variable x in the r.h.s. can be renamed x_n. A symmetrization of the coefficient of $\prod_{i<n} \varphi(x_i)$ then yields a factor $1/n$ and the sum of all gradients squared. We have reconstructed, up to a factor, the kinetic term (5.83). We conclude that

$$\mathbf{T}\Psi(\varphi) = -\frac{\hbar^2}{2m} \int \mathrm{d}^d x \, \varphi(x) \nabla_x^2 \frac{\delta}{\delta\varphi(x)} \Psi(\varphi). \tag{5.87}$$

For the one-body potential $\mathbf{V}_1$ an analogous argument shows that

$$\mathbf{V}_1\Psi(\varphi) = \int \mathrm{d}^d x\, \varphi(x)V_1(x)\frac{\delta}{\delta\varphi(x)}\Psi(\varphi).$$

To generate a two-body potential it is necessary to differentiate twice with respect to φ at two different points. One verifies

$$\mathbf{V}_2\Psi(\varphi) = \frac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}^d y\, \varphi(x)\varphi(y)V_2(x,y)\frac{\delta^2}{\delta\varphi(x)\delta\varphi(y)}\Psi(\varphi). \tag{5.88}$$

We now have a representation of the full hamiltonian $\mathbf{H}$ acting on generating functionals. We still need an operator for the number of particles, $\mathbf{N}$. It is given by

$$\mathbf{N} = \int \mathrm{d}^d x\, \varphi(x)\frac{\delta}{\delta\varphi(x)} \quad \Rightarrow \quad [\mathbf{N},\mathbf{H}] = 0\,. \tag{5.89}$$

In what follows we consider as total 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 have constructed here what is called a ***quantum field theory***, more precisely a non-relativistic, non-local, quantum field theory.

5.5.2 Functional integral

We now combine functional methods and the holomorphic formalism to find a representation of matrix elements of the statistical operator as functional integrals, generalization to an infinite number of complex paths of the path integrals discussed in Section 5.2. In the functional integral fields then replace the paths of ordinary quantum mechanics. The form of the functional integral actually follows rather directly from results already obtained in quantum mechanics, 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 5.1. 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 $\varphi(x)$. As a kernel the identity operator $\mathcal{I}(\varphi,\bar\varphi)$ takes the form

$$\mathcal{I}(\varphi,\bar\varphi) = \exp\left[\int \mathrm{d}^d x\, \varphi(x)\bar\varphi(x)\right].$$

Note that the hamiltonian here is directly written in normal order. Its matrix elements are

$$\begin{aligned}\langle\varphi|\mathbf{H}|\bar\varphi\rangle = \mathcal{I}(\varphi,\bar\varphi)\left\{\int \mathrm{d}^d x\, \varphi(x)\left[-\frac{\hbar^2}{2m}\nabla^2 + V_1(x)\right]\bar\varphi(x)\right.\\ \left. +\frac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}^d y\, \varphi(x)\varphi(y)V_2(x,y)\bar\varphi(x)\bar\varphi(y)\right\}.\end{aligned} \tag{5.90}$$

The particle number operator is proportional to $\int \mathrm{d}^d x\, \varphi(x)\bar\varphi(x)$:

$$\langle\varphi|\,\mathbf{N}\,|\bar\varphi\rangle = \int \mathrm{d}^d x\, \varphi(x)\bar\varphi(x)\left(\exp\int \mathrm{d}^d x\, \varphi(x)\bar\varphi(x)\right). \tag{5.91}$$

If we now consider the space coordinate x as a continuous index, we can adapt the expressions of Section 5.2, in particular equation (5.35), 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 functional integral, because we no longer integrate over paths, but, instead, over *fields* $\{\varphi(t,x),\bar{\varphi}(t,x)\}$. Here, we find

$$\langle\varphi''|\,\mathbf{U}(t'',t')\,|\bar{\varphi}'\rangle = \langle\varphi''|\,\mathrm{e}^{-(t''-t)(\mathbf{H}-\mu\mathbf{N})/\hbar}\,|\bar{\varphi}'\rangle = \int[\mathrm{d}\bar{\varphi}(t,x)\mathrm{d}\varphi(t,x)]\exp[-\mathcal{S}(\varphi,\bar{\varphi})] \tag{5.92}$$

with the boundary conditions

$$\bar{\varphi}(t,x') \equiv \bar{\varphi}'(x), \quad \varphi(t,x'') \equiv \varphi''(x),$$

and the euclidean action

$$\begin{aligned}\mathcal{S}(\varphi,\bar{\varphi}) = {}& -\bar{\varphi}(t,x')\varphi(t,x') + \int \mathrm{d}t\,\mathrm{d}^d x\,\bar{\varphi}(t,x)\left(-\frac{\partial}{\partial t} - \frac{\hbar}{2m}\nabla_x^2 + \frac{V_1(x)-\mu}{\hbar}\right)\varphi(t,x) \\ & + \frac{1}{2\hbar}\int \mathrm{d}t\,\mathrm{d}^d x\,\mathrm{d}^d y\,\bar{\varphi}(t,x)\varphi(t,x)V_2(x,y)\bar{\varphi}(t,y)\varphi(t,y).\end{aligned} \tag{5.93}$$

The statistical operator is obtained by setting $t = \hbar\beta$. The partition function is then given by

$$\mathcal{Z}(\beta) = \int[\mathrm{d}\bar{\varphi}(t,x)\mathrm{d}\varphi(t,x)]\exp[-\mathcal{S}(\varphi,\bar{\varphi})], \tag{5.94}$$

with the periodic boundary conditions

$$\bar{\varphi}(-\beta/2,x) = \bar{\varphi}(\beta/2,x), \quad \varphi(-\beta/2,x) = \varphi(\beta/2,x),$$

and the euclidean action $\mathcal{S}(\varphi,\bar{\varphi})$ is

$$\begin{aligned}\mathcal{S}(\varphi,\bar{\varphi}) = {}& \int \mathrm{d}t\,\mathrm{d}^d x\,\bar{\varphi}(t,x)\left(-\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\nabla_x^2 + V_1(x)-\mu\right)\varphi(t,x) \\ & + \frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^d x\,\mathrm{d}^d y\,\bar{\varphi}(t,x)\varphi(t,x)V_2(x,y)\bar{\varphi}(t,y)\varphi(t,y).\end{aligned} \tag{5.95}$$

We have thus derived a representation of the partition function in the form of a functional integral for a *non-relativistic quantum field theory.* The generalization to a relativistic theory will mainly be a matter of implementing relativistic invariance.

Remark. As a consequence of particle number conservation the functional integral has a $U(1)$ phase symmetry:

$$\varphi(t,x) \mapsto \mathrm{e}^{i\theta}\,\varphi(t,x), \quad \bar{\varphi}(t,x) \mapsto \mathrm{e}^{-i\theta}\,\bar{\varphi}(t,x).$$

We show in Chapter 28, devoted to phase transitions, that interesting physics is associated with the spontaneous symmetry breaking of this symmetry.

5.5.3 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. Of course the functional formalism is not required for these simple examples, but it then allows to study more easily the effect of additional interactions.

Free Bose gas. In the absence of potential, the partition function can easily be calculated. To take advantage of translation invariance, we expand the fields $\varphi, \bar{\varphi}$ in Fourier modes. In d space dimensions

$$\bar{\varphi}(t, \mathbf{x}) = \int \mathrm{d}^d p \, \mathrm{e}^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \, \tilde{\varphi}^*(t, \mathbf{p}), \quad \varphi(t, \mathbf{x}) = \int \mathrm{d}^d p \, \mathrm{e}^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \, \tilde{\varphi}(t, \mathbf{p}). \tag{5.96}$$

The jacobian is trivial and the action becomes

$$\mathcal{S}_0(\tilde{\varphi}, \tilde{\varphi}^*) = (2\pi\hbar)^d \int \mathrm{d}t \, \mathrm{d}^d p \, \tilde{\varphi}^*(t, \mathbf{p}) \left(-\frac{\partial}{\partial t} + \frac{p^2}{2m} - \mu \right) \tilde{\varphi}(t, \mathbf{p}). \tag{5.97}$$

It is convenient to first calculate the partition function in a hypercubic box of linear size L with periodic boundary conditions and then take the thermodynamic limit. Then the fields $\varphi(t, \mathbf{x}), \bar{\varphi}(t, \mathbf{x})$ are periodic functions of all space variables, of period L. The corresponding Fourier variables belong to a lattice:

$$\mathbf{p} = \frac{2\pi\hbar}{L} \mathbf{n}, \quad \mathbf{n} \in \mathbb{Z}^d,$$

and the action reads

$$\mathcal{S}_0(\tilde{\varphi}^*, \tilde{\varphi}) = (2\pi\hbar)^{2d} \int \mathrm{d}t \sum_{\mathbf{n}} \tilde{\varphi}^*(t, \mathbf{p}) \left(-\frac{\partial}{\partial t} + \frac{p^2}{2m} - \mu \right) \tilde{\varphi}(t, \mathbf{p}). \tag{5.98}$$

The functional integral factorizes and the result is a product over all values of $\mathbf{n}$ of a partition function of the form (5.33). The free energy, $\mathcal{W}_0(\beta) = \beta^{-1} \ln \mathcal{Z}_0(\beta)$, is

$$\mathcal{W}_0(\beta) = -\frac{1}{\beta} \sum_{\mathbf{n} \in \mathbb{Z}^d} \ln \left(1 - \mathrm{e}^{-\beta(p^2/2m - \mu)} \right). \tag{5.99}$$

In the thermodynamic limit $L \to \infty$, the free energy per unit volume, which is the pressure Π becomes ($\mathrm{d}\mathbf{n} = \mathrm{d}\mathbf{p} L / 2\pi\hbar$)

$$\Pi = L^{-d} \mathcal{W}_0(\beta) = -\frac{1}{\beta} \int \frac{\mathrm{d}^d p}{(2\pi\hbar)^d} \ln \left(1 - \mathrm{e}^{-\beta(p^2/2m - \mu)} \right). \tag{5.100}$$

Note that the Bose gas is stable only if the chemical potential is non-positive.

The average energy density is obtained from the derivative of $\ln \mathcal{Z}$ with respect to β, at $\beta\mu$ fixed:

$$\langle H \rangle = -\frac{1}{L^d} \frac{\partial \ln \mathcal{Z}}{\partial \beta} \bigg|_{\beta\mu \text{ fixed}} = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p (p^2/2m)}{\left(\mathrm{e}^{\beta(p^2/2m - \mu)} - 1 \right)} \,. \tag{5.101}$$

Taking the derivative of $\ln \mathcal{Z}$ with respect to $\beta\mu$ (β fixed), one obtains the average particle number and thus the gas density ρ:

$$\rho = L^{-d} \langle N \rangle = \frac{1}{L^d \beta} \int \mathrm{d}t \, \mathrm{d}^d x \, \langle \varphi(t,x) \bar{\varphi}(t,x) \rangle = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p}{\mathrm{e}^{\beta(p^2/2m - \mu)} - 1} \,. \tag{5.102}$$

As is well known, the equation of state (5.102) exhibits 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 two, since μ cannot become positive, ρ is bounded by the value ρ_c of the integral calculated for $\mu = 0$:

$$\rho \le \rho_c = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p}{\mathrm{e}^{\beta p^2/2m} - 1} = \zeta(d/2) \left(\frac{mT}{2\pi\hbar^2} \right)^{d/2} ,$$

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:

$$T_0 = \frac{2\pi\hbar^2}{m} \left(\frac{\rho}{\zeta(d/2)} \right)^{2/d} .$$

To understand the physics below T_0 one has 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 here is the zero momentum mode.

In two dimensions, because ρ_c diverges, there is no condensation.

General one-body potential. To calculate the gaussian functional integral we replace the Fourier expansion (5.96) by an expansion on the eigenfunctions of the one-body hamiltonian H_1. The same arguments then lead to

$$\mathcal{W}(\beta) = -\operatorname{tr} \ln \left(1 - \mathrm{e}^{\beta\mu - \beta H_1} \right) .$$

Similarly

$$\langle N \rangle = \operatorname{tr} \left[1 - \mathrm{e}^{\beta\mu - \beta H_1} \right]^{-1} , \tag{5.103}$$

$$\langle H \rangle = \operatorname{tr} H_1 \left[1 - \mathrm{e}^{\beta\mu - \beta H_1} \right]^{-1} . \tag{5.104}$$

An analysis based on the arguments of Section 2.6 shows that in the semi-classical limit the free energy becomes

$$\mathcal{W}(\beta) = -\frac{1}{\beta} \int \frac{\mathrm{d}^d p \, \mathrm{d}^d x}{(2\pi\hbar)^d} \ln \left(1 - \mathrm{e}^{\beta\mu - \beta H_1(p,x)} \right) ,$$

where $H_1(p,x)$ is now the classical hamiltonian.

A simple example is provided by a harmonic well, of the kind relevant for the magnetic traps of actual Bose–Einstein condensation experiments. For simplicity, we assume that the trap is spherical:

$$H_1 = \frac{1}{2m} \mathbf{p}^2 + \frac{1}{2} m\omega^2 \mathbf{x}^2 .$$

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

$$T_0 = \frac{1}{\beta} = \hbar\omega(\langle N \rangle)^{1/d}.$$

At lower temperatures the discrete nature of the quantum spectrum becomes relevant, the relation between particle numbers and chemical potential can be satisfied when $\mu \to d\hbar\omega/2$, and particles accumulate in the ground state.

5.5.4 *Pair-potentials: an example*

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 $V_1 \equiv 0$):

$$\mathcal{S}(\varphi, \bar{\varphi}) = \int \mathrm{d}t\, \mathrm{d}^d x \left[\bar{\varphi}(t,x) \left(-\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\nabla^2 - \mu \right) \varphi(t,x) + \frac{g}{2}\left(\bar{\varphi}(t,x)\varphi(t,x)\right)^2 \right]. \quad (5.105)$$

For simple potentials, the strength g of the interaction is proportional to $\hbar^2 a^{d-2}/m$, where a is the scattering length. It 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. Instead, for $\mu > 0$ the minimum of the $\varphi, \bar{\varphi}$ potential now corresponds to a non-vanishing value of $|\varphi|$. This indicates 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. It, therefore, describes the physics of the He superfluid transition (see Section 28.3).

Note, however, that a hamiltonian with a δ-function potential is well defined only in one dimension. In higher dimensions it leads to divergences that have to be dealt with (regularized, see Chapter 9). In particular, in two dimensions the model is renormalizable, as power counting shows, and non-renormalizable in higher dimensions (see Chapters 9 and 17).

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.

Perturbation theory. As an exercise, we calculate the correction of order g to the free energy. As a basic element, we need the gaussian two-point $\langle \varphi\bar{\varphi} \rangle$ correlation function. In Fourier space, from (5.97,5.30,5.32), we infer ($\vartheta(t) = (1 + \epsilon(t))/2$)

$$\Delta(t, \mathbf{p}) = \langle \tilde{\varphi}^*(t, \mathbf{p})\tilde{\varphi}(0, \mathbf{p}) \rangle = (2\pi\hbar)^{-d}\, \mathrm{e}^{-\omega(\mathbf{p})t} \left(\vartheta(t) + \frac{1}{\mathrm{e}^{\beta\omega(\mathbf{p})} - 1} \right)$$

with $\omega(\mathbf{p}) = p^2/2m - \mu$. We recognize the sum of the two contributions generated by quantum and thermal fluctuations.

Applying Wick's theorem we obtain

$$\mathcal{W}(\beta) = \mathcal{W}_0(\beta) - \frac{g}{\beta} \int \mathrm{d}t\, \mathrm{d}^d x\, \left(\langle \varphi(t,x)\bar{\varphi}(t,x)\rangle\right)^2 + O(g^2).$$

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 we face the $\epsilon(0)$ problem. The convention corresponding to normal order is

$$\Delta(\mathbf{p}, 0) = \frac{1}{\mathrm{e}^{\beta(p^2/2m-\mu)} - 1},$$

and the result is simply the square of the leading order density

$$\mathcal{W}(\beta) = \mathcal{W}_0(\beta) - gL^d \left[\frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p}{\mathrm{e}^{\beta(p^2/2m-\mu)} - 1}\right]^2 + O(g^2).$$

High temperature. Finally, as a consequence of the periodic boundary conditions in the euclidean time direction, the field can be expanded on a basis of periodic functions of period β,

$$\varphi(t,x) = \sum_{\nu\in\mathbb{Z}} \mathrm{e}^{2i\pi\nu t/\beta} \varphi_\nu(x), \quad \bar{\varphi}(t,x) = \sum_{\nu\in\mathbb{Z}} \mathrm{e}^{2i\pi\nu t/\beta} \bar{\varphi}_\nu(x).$$

At high temperature, $\beta \to 0$, the time derivative in the action 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 wave length λ_{th} the condition can be written as

$$\lambda_{\text{th}} = \hbar\sqrt{\beta/m} = \sqrt{\hbar^2/mT} \ll \xi,$$

and 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 (5.105) can then be approximated by

$$\mathcal{S}(\varphi,\bar{\varphi}) = \beta \int \mathrm{d}^d x \left[\bar{\varphi}_0(x)\left(-\frac{\hbar^2}{2m}\nabla^2 - \mu\right)\varphi_0(x) + \frac{g}{2}\left(\bar{\varphi}_0(x)\varphi_0(x)\right)^2\right], \tag{5.106}$$

an action that we will study systematically in the context of euclidean quantum field theory, and which is relevant to the superfluid phase transition (see Section 28.3).

5.6 The Fermi Gas. Functional Integrals

The path integral for fermions constructed in Section 5.3 and, in particular, the considerations of Section 5.3.4 allow for a simple generalization of the preceding method to fermions.

We start from the same hamiltonian $\mathbf{H}$ as in the preceding section, with the form (5.82) in the n-particle subspace. Since we now deal with fermions (for simplicity without internal degrees of freedom) the wave functions ψ_n are antisymmetric. Therefore, the construction of a generating functional requires the introduction of functions $\varphi(x)$, generators of a Grassmann algebra:

$$\varphi(x)\varphi(x') + \varphi(x')\varphi(x) = 0 .$$

We then define the functional Ψ:

$$\Psi(\varphi) = \sum_{n=0} \frac{1}{n!} \left(\int \prod_i \mathrm{d}^d x_i \, \varphi(x_i) \right) \psi_n(x_1, \ldots, x_n). \tag{5.107}$$

The derivation of the functional 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 appear.

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 5.3, one finds for $\mathbf{H}$,

$$\begin{aligned} \langle \varphi | \mathbf{H} | \bar\varphi \rangle = & \left(\exp - \int \mathrm{d}^d x \, \bar\varphi(x)\varphi(x) \right) \left\{ \int \mathrm{d}^d x \, \varphi(x) \left[-\frac{\hbar^2}{2m} \nabla^2 + V_1(x) \right] \bar\varphi(x) \right. \\ & \left. + \frac{1}{2} \int \mathrm{d}^d x \, \mathrm{d}^d y \, \varphi(x)\varphi(y) V_2(x,y) \bar\varphi(y) \bar\varphi(x) \right\} . \end{aligned} \tag{5.108}$$

The partition function $\mathcal{Z}(\beta)$ is then given by a functional integral,

$$\mathcal{Z}(\beta) = \int [\mathrm{d}\varphi(t,x) \mathrm{d}\bar\varphi(t,x)] \, \mathrm{e}^{-\mathcal{S}(\varphi,\bar\varphi)}, \tag{5.109}$$

where the euclidean action $\mathcal{S}(\varphi, \bar\varphi)$ is

$$\begin{aligned} \mathcal{S}(\varphi,\bar\varphi) = & \int \mathrm{d}t \, \mathrm{d}^d x \, \bar\varphi(t,x) \left(\frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 - V_1(x) + \mu \right) \varphi(t,x) \\ & + \frac{1}{2} \int \mathrm{d}t \, \mathrm{d}^d x \, \mathrm{d}^d y \, \bar\varphi(t,x)\varphi(t,x) V_2(x,y) \bar\varphi(t,y)\varphi(t,y), \end{aligned} \tag{5.110}$$

and the fields $\bar\varphi(t,x), \varphi(t,x)$ satisfy *antiperiodic boundary conditions* in the euclidean time direction.

5.6.1 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 (5.54), which gives the partition function for one degree of freedom, one obtains

$$L^{-d} \mathcal{W}(\beta) = \frac{1}{(2\pi\hbar)^d} \int \mathrm{d}^d p \, \ln \left(1 + \mathrm{e}^{-\beta(p^2/2m - \mu)} \right) . \tag{5.111}$$

Clearly, the sign of the chemical potential is now arbitrary; the Pauli principle prevents the collapse of the fermion system.

The density takes the form

$$\rho = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p}{\mathrm{e}^{\beta(p^2/2m-\mu)}+1} . \tag{5.112}$$

The mean energy density is

$$\langle H \rangle = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p (p^2/2m)}{\left(\mathrm{e}^{\beta(p^2/2m-\mu)}+1\right)} . \tag{5.113}$$

The $\delta(x)$-function potential. When the two-body potential is short-range, and we are interested in long wavelength phenomena, we 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:

$$\mathcal{S}(\varphi, \bar\varphi) = \int \mathrm{d}t\, \mathrm{d}^d x \left[\bar\varphi^\alpha(t,x) \left(\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2 + \mu\right) \varphi^\alpha(t,x) + \frac{g}{2}\left(\bar\varphi^\alpha(t,x)\varphi^\alpha(t,x)\right)^2\right] . \tag{5.114}$$

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 one space dimension the model with the action (5.114) is also integrable. There it is the non-relativistic limit of the massive Thirring model, which is also integrable.

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 31.7).

5.6.2 Non-relativistic fermion gas at low temperature, in one dimension

The equation of state (5.112) shows that at low temperature (β large) and $\mu > 0$ only the states with momenta p below the Fermi surface are occupied

$$|p| \le k_{\mathrm{F}} = \sqrt{2m\mu} . \tag{5.115}$$

Moreover, in the presence of weak interactions only excitations with momenta close to the Fermi surface (5.115) are important.

In one dimension this implies that momenta k are close to $\pm k_{\mathrm{F}}$. It is then convenient to set

$$k = \pm k_{\mathrm{F}} + q , \quad |q| \ll k_{\mathrm{F}} .$$

The Fourier transform of the gaussian two-point function or propagator can thus be approximated by

$$\Delta(\omega, k) = \frac{1}{-i\omega + k^2/2m - \mu} \sim \sum_{\epsilon=\pm 1} \frac{1}{-i\omega + \epsilon q k_{\mathrm{F}}/m} .$$

Clearly, the sign of the chemical potential is now arbitrary; the Pauli principle prevents the collapse of the fermion system.

The density takes the form

$$\rho = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p}{\mathrm{e}^{\beta(p^2/2m-\mu)}+1} . \tag{5.112}$$

The mean energy density is

$$\langle H \rangle = \frac{1}{(2\pi\hbar)^d} \int \frac{\mathrm{d}^d p (p^2/2m)}{\left(\mathrm{e}^{\beta(p^2/2m-\mu)}+1\right)} . \tag{5.113}$$

The $\delta(x)$-function potential. When the two-body potential is short-range, and we are interested in long wavelength phenomena, we 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:

$$\mathcal{S}(\varphi,\bar\varphi) = \int \mathrm{d}t\,\mathrm{d}^d x \left[\bar\varphi^\alpha(t,x)\left(\frac{\partial}{\partial t} + \frac{\hbar^2}{2m}\nabla^2 + \mu\right)\varphi^\alpha(t,x) + \frac{g}{2}\left(\bar\varphi^\alpha(t,x)\varphi^\alpha(t,x)\right)^2\right] . \tag{5.114}$$

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 one space dimension the model with the action (5.114) is also integrable. There it is the non-relativistic limit of the massive Thirring model, which is also integrable.

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 31.7).

5.6.2 Non-relativistic fermion gas at low temperature, in one dimension

The equation of state (5.112) shows that at low temperature (β large) and $\mu > 0$ only the states with momenta p below the Fermi surface are occupied

$$|p| \le k_{\mathrm{F}} = \sqrt{2m\mu} . \tag{5.115}$$

Moreover, in the presence of weak interactions only excitations with momenta close to the Fermi surface (5.115) are important.

In one dimension this implies that momenta k are close to $\pm k_{\mathrm{F}}$. It is then convenient to set

$$k = \pm k_{\mathrm{F}} + q , \quad |q| \ll k_{\mathrm{F}} .$$

The Fourier transform of the gaussian two-point function or propagator can thus be approximated by

$$\Delta(\omega,k) = \frac{1}{-i\omega + k^2/2m - \mu} \sim \sum_{\epsilon=\pm1} \frac{1}{-i\omega + \epsilon q k_{\mathrm{F}}/m} .$$

This corresponds to approximating, locally near the Fermi points, the dispersion curve by straight lines. We recognize the propagators of a massless relativistic fermion, the two values $\epsilon = \pm 1$ corresponding to right and left movers. The quantity $m/\hbar k_{\mathrm{F}}$ 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 only have small momentum components:

$$\begin{aligned}\varphi(t,x) &= \mathrm{e}^{-ik_{\mathrm{F}}x/\hbar}\,\eta_1(t,x) + \mathrm{e}^{ik_{\mathrm{F}}x/\hbar}\,\eta_2(t,x),\\ \bar\varphi(t,x) &= -i\,\mathrm{e}^{-ik_{\mathrm{F}}x/\hbar}\,\bar\eta_1(t,x) + i\,\mathrm{e}^{ik_{\mathrm{F}}x/\hbar}\,\bar\eta_2(t,x).\end{aligned}$$

After substitution, all terms in the action with explicit exponential factors $\mathrm{e}^{\pm ik_{\mathrm{F}}x/\hbar}$, which correspond to large momentum exchanges, can be neglected at leading order. The free fermion action then becomes

$$\mathcal{S}_0(\bar{\boldsymbol\eta},\boldsymbol\eta) = \int \mathrm{d}t\,\mathrm{d}x\,\bar{\boldsymbol\eta}(t,x)\left[\sigma_2\partial_t + (\hbar k_{\mathrm{F}}/m)\sigma_1\partial_x\right]\boldsymbol\eta(t,x),$$

where the two-component fermion fields $\{\bar{\boldsymbol\eta},\boldsymbol\eta\}$ can be identified with a massless Dirac fermion field, and σ_μ, $\mu = 1,2,3$, are the Pauli matrices (see Appendix A8). Note that in these covariant relativistic conventions $\boldsymbol\eta^\dagger = \bar{\boldsymbol\eta}\sigma_2$.

We now again consider N-component fermion fields φ_α and the example of the action (5.114) with a two-body δ-function potential. After the same substitution (again all terms proportional to non-trivial powers of $\mathrm{e}^{\pm ik_{\mathrm{F}}x/\hbar}$ being neglected) the action becomes

$$\begin{aligned}\mathcal{S}(\bar{\boldsymbol\eta},\boldsymbol\eta) = \int \mathrm{d}t\,\mathrm{d}x\Big[&\bar\eta^\alpha(t,x)\left(\sigma_2\partial_t + (\hbar k_{\mathrm{F}}/m)\sigma_1\partial_x\right)\eta^\alpha(t,x) + \frac{g}{2}\left(\bar\eta^\alpha(t,x)\sigma_2\eta^\alpha(t,x)\right)^2\\ &+\frac{g}{4}\left(\bar\eta^\alpha(t,x)\eta^\alpha(t,x)\right)^2 - \frac{g}{4}\left(\bar\eta^\alpha(t,x)\sigma_3\eta^\alpha(t,x)\right)^2\Big].\end{aligned}$$

The two-component fermion fields $\{\bar\eta^\alpha,\eta^\alpha\}$ represent N massless Dirac fermions. The action has a chiral symmetry

$$\eta^\alpha \mapsto \mathrm{e}^{i\sigma_3\theta}\,\eta^\alpha\,,\quad \bar\eta^\alpha \mapsto \bar\eta^\alpha\,\mathrm{e}^{i\sigma_3\theta}\,,$$

which prevents the addition of a fermion mass term (see Appendix A8).

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 will be considered in Sections 32.6, 33.6, thus describes self-interacting one-dimensional non-relativistic fermions at low temperature. All results derived from the point of view of the Kosterlitz–Thouless phase transition have an interpretation in systems like the Luttinger liquid (the equivalent of the Fermi liquid of higher dimensions), or one-dimensional conductors. For a systematic discussion of such models the reader is referred to the literature.

In higher space dimensions, the situation is no longer that simple and does not lead to massless relativistic fermions.

This corresponds to approximating, locally near the Fermi points, the dispersion curve by straight lines. We recognize the propagators of a massless relativistic fermion, the two values $\epsilon = \pm 1$ corresponding to right and left movers. The quantity $m/\hbar k_{\rm F}$ 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 only have small momentum components:

$$\begin{aligned}\varphi(t,x) &= \mathrm{e}^{-ik_{\rm F}x/\hbar}\,\eta_1(t,x) + \mathrm{e}^{ik_{\rm F}x/\hbar}\,\eta_2(t,x),\\ \bar\varphi(t,x) &= -i\,\mathrm{e}^{-ik_{\rm F}x/\hbar}\,\bar\eta_1(t,x) + i\,\mathrm{e}^{ik_{\rm F}x/\hbar}\,\bar\eta_2(t,x).\end{aligned}$$

After substitution, all terms in the action with explicit exponential factors $\mathrm{e}^{\pm ik_{\rm F}x/\hbar}$, which correspond to large momentum exchanges, can be neglected at leading order. The free fermion action then becomes

$$\mathcal{S}_0(\bar{\boldsymbol\eta},\boldsymbol\eta) = \int \mathrm{d}t\,\mathrm{d}x\,\bar{\boldsymbol\eta}(t,x)\left[\sigma_2\partial_t + (\hbar k_{\rm F}/m)\sigma_1\partial_x\right]\boldsymbol\eta(t,x),$$

where the two-component fermion fields $\{\bar{\boldsymbol\eta},\boldsymbol\eta\}$ can be identified with a massless Dirac fermion field, and σ_μ, $\mu=1,2,3$, are the Pauli matrices (see Appendix A8). Note that in these covariant relativistic conventions $\boldsymbol\eta^\dagger = \bar{\boldsymbol\eta}\sigma_2$.

We now again consider N-component fermion fields φ_α and the example of the action (5.114) with a two-body δ-function potential. After the same substitution (again all terms proportional to non-trivial powers of $\mathrm{e}^{\pm ik_{\rm F}x/\hbar}$ being neglected) the action becomes

$$\begin{aligned}\mathcal{S}(\bar{\boldsymbol\eta},\boldsymbol\eta) = \int \mathrm{d}t\,\mathrm{d}x\Big[&\bar\eta^\alpha(t,x)\left(\sigma_2\partial_t + (\hbar k_{\rm F}/m)\sigma_1\partial_x\right)\eta^\alpha(t,x) + \frac{g}{2}\left(\bar\eta^\alpha(t,x)\sigma_2\eta^\alpha(t,x)\right)^2\\ &+\frac{g}{4}\left(\bar\eta^\alpha(t,x)\eta^\alpha(t,x)\right)^2 - \frac{g}{4}\left(\bar\eta^\alpha(t,x)\sigma_3\eta^\alpha(t,x)\right)^2\Big].\end{aligned}$$

The two-component fermion fields $\{\bar\eta^\alpha,\eta^\alpha\}$ represent N massless Dirac fermions. The action has a chiral symmetry

$$\eta^\alpha \mapsto \mathrm{e}^{i\sigma_3\theta}\,\eta^\alpha\,,\quad \bar\eta^\alpha \mapsto \bar\eta^\alpha\,\mathrm{e}^{i\sigma_3\theta},$$

which prevents the addition of a fermion mass term (see Appendix A8).

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 will be considered in Sections 32.6, 33.6, thus describes self-interacting one-dimensional non-relativistic fermions at low temperature. All results derived from the point of view of the Kosterlitz–Thouless phase transition have an interpretation in systems like the Luttinger liquid (the equivalent of the Fermi liquid of higher dimensions), or one-dimensional conductors. For a systematic discussion of such models the reader is referred to the literature.

In higher space dimensions, the situation is no longer that simple and does not lead to massless relativistic fermions.

Bibliographical Notes

For more details about the holomorphic representation, see

L. D. Faddeev in *Methods in Field Theory*, Les Houches School 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

The fermionic path integral is discussed in

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

For a discussion of the use of functional integrals in Quantum Statistical Physics see also

V.N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics*, Reidel Publishing (Dordrecht 1983); L.S. Brown, *Quantum Field Theory* (Cambridge University Press, Cambridge 1992).

For a review of the application of functional integral methods to classical and quantum Bose fluids, see

F.W. Wiegel, *Phys. Rep.* 16C (1975) 57,

For the application of these ideas of Section 5.6.2 to one-dimensional conductors, see the review

J. Solyom, *Advances in Physics* 28 (1979) 201.

6 QUANTUM EVOLUTION: FROM PARTICLES TO FIELDS

Up to now we have discussed only quantum statistical operators, and thus, formally, evolution in imaginary or euclidean time. To calculate scattering S-matrix elements, quantities relevant to Particle Physics, it is necessary to consider instead the quantum evolution operator in real time.

Therefore, we first derive the path integral representation of the evolution operator and the S-matrix in simple quantum mechanics. To illustrate the power of the formalism we 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 5.1 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 functional integral.

Using the parallel formalism of Section 5.6 we derive an analogous representation for the evolution operator of a system of non-relativistic fermions.

We then begin our study of *relativistic quantum field theory* with the example of the self-coupled neutral scalar boson. We show that the holomorphic formalism, in a form that extends the construction of Section 5.5 to relativistic real time evolution, leads to various representations of the S-matrix in terms of functional integrals. We relate S-matrix elements to the continuation to real time of various kinds of euclidean correlation functions. We discuss only scalar fields and postpone the study of relativistic fermions to Chapter 8 because we need the representations of the spin group.

Though relativistic quantum field theory appears as a natural formal generalization of quantum mechanics, it is not quite obvious how ordinary quantum mechanics is recovered in the non-relativistic limit. Indeed one could have also thought about developing a formalism of relativistic quantum particles. The known existence of the electromagnetic field, which predates quantum mechanics, made this alternative possibility less attractive since it would not have unified fields and particles. We, therefore, briefly indicate how non-relativistic quantum mechanics emerges in the low-energy, low-momentum limit of a massive relativistic quantum field theory. The limiting theory takes the form of a quantum many-body theory of the kind discussed in Section 5.5.

As a warning, we also have to stress that several quantities we meet in this chapter are really infinite, or at least infinite in high enough dimensions. Therefore, the discussion is sometimes rather formal. However, the divergence problems will be carefully studied later, in the coming chapters. Regularization by a space lattice would have rendered most quantities meaningful but would have complicated the analysis.

Quantum Evolution and Scattering Matrix. Evolution in quantum mechanics is associated with an operator acting linearly on the Hilbert space of states. Conservation of probabilities implies that the evolution operator must be unitary. Finally, one assumes that the evolution of an isolated system is markovian (without memory effects). This translates into a group structure for the evolution operator $U(t'', t')$ from time t' to time t''

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1)\,. \tag{6.1}$$

Assuming the evolution differentiable, we expand $U(t+\varepsilon,t)$ for an infinitesimal time interval ε:

$$U(t+\varepsilon,t) = \mathbf{1} - i\varepsilon H(t)/\hbar + O(\varepsilon^2)\,,$$

where $H(t)$ is the quantum hamiltonian. Equation (6.1) then implies

$$i\hbar\frac{\partial U}{\partial t}(t,t') = H(t)Ut,t'), \quad U(t',t') = \mathbf{1}\,. \tag{6.2}$$

When the operator $H(t)$ is time-independent $U(t'',t') = U(t''-t') = \mathrm{e}^{-i(t''-t')H/\hbar}$. 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 with 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:

$$S = \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \mathrm{e}^{iH_0t''/\hbar}\,U(t'',t')\,\mathrm{e}^{-iH_0t'/\hbar}, \tag{6.3}$$

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.

6.1 Time Evolution and Scattering Matrix in Quantum Mechanics

In the simple quantum mechanics of a particle in a potential, examples of free and interacting hamiltonians are simply

$$H_0 = \mathbf{p}^2/2m, \qquad H = \mathbf{p}^2/2m + V(\mathbf{q},t). \tag{6.4}$$

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 $|t|\to\infty$ the evolution converges fast enough toward the free evolution corresponding to H_0.

The free evolution operator $U_0 = \mathrm{e}^{-i(t''-t')H_0/\hbar}$, corresponding to the hamiltonian H_0 in (6.4), in d space dimensions, is given by

$$\langle\mathbf{q}''|\,U_0(t'',t')\,|\mathbf{q}'\rangle = \frac{1}{(2\pi)^d}\int \mathrm{d}^d p\, \exp\frac{i}{\hbar}\left[\mathbf{p}\cdot(\mathbf{q}''-\mathbf{q}') - \mathbf{p}^2(t''-t')/2m\right] \tag{6.5a}$$

$$= \left(\frac{m}{2i\pi\hbar(t''-t')}\right)^{d/2} \exp\left[\frac{i}{\hbar}\frac{m\,(\mathbf{q}''-\mathbf{q}')^2}{2(t''-t')}\right]. \tag{6.5b}$$

In the momentum basis the relation (6.3) between the S-matrix and the evolution operator takes the form

$$\langle\mathbf{p}''|\,S\,|\mathbf{p}'\rangle = \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \mathrm{e}^{iE''t''/\hbar}\,\langle\mathbf{p}''|\,U(t'',t')\,|\mathbf{p}'\rangle\,\mathrm{e}^{-iE't'/\hbar}, \tag{6.6}$$

where

$$E' = E(\mathbf{p}'), \quad E'' = E(\mathbf{p}''), \qquad E(\mathbf{p}) = \mathbf{p}^2/2m\,.$$

Assuming the evolution differentiable, we expand $U(t+\varepsilon,t)$ for an infinitesimal time interval ε:

$$U(t+\varepsilon,t) = \mathbf{1} - i\varepsilon H(t)/\hbar + O(\varepsilon^2)\,,$$

where $H(t)$ is the quantum hamiltonian. Equation (6.1) then implies

$$i\hbar\frac{\partial U}{\partial t}(t,t') = H(t)Ut,t'), \quad U(t',t') = \mathbf{1}\,. \tag{6.2}$$

When the operator $H(t)$ is time-independent $U(t'',t') = U(t''-t') = \mathrm{e}^{-i(t''-t')H/\hbar}$. 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 with 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:

$$S = \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \mathrm{e}^{iH_0t''/\hbar}\, U(t'',t')\, \mathrm{e}^{-iH_0t'/\hbar}, \tag{6.3}$$

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.

6.1 Time Evolution and Scattering Matrix in Quantum Mechanics

In the simple quantum mechanics of a particle in a potential, examples of free and interacting hamiltonians are simply

$$H_0 = \mathbf{p}^2/2m, \qquad H = \mathbf{p}^2/2m + V(\mathbf{q},t). \tag{6.4}$$

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 $|t|\to\infty$ the evolution converges fast enough toward the free evolution corresponding to H_0.

The free evolution operator $U_0 = \mathrm{e}^{-i(t''-t')H_0/\hbar}$, corresponding to the hamiltonian H_0 in (6.4), in d space dimensions, is given by

$$\langle\mathbf{q}''|\,U_0(t'',t')\,|\mathbf{q}'\rangle = \frac{1}{(2\pi)^d}\int \mathrm{d}^d p \exp\frac{i}{\hbar}\left[\mathbf{p}\cdot(\mathbf{q}''-\mathbf{q}') - \mathbf{p}^2(t''-t')/2m\right] \tag{6.5a}$$

$$= \left(\frac{m}{2i\pi\hbar(t''-t')}\right)^{d/2} \exp\left[\frac{i}{\hbar}\frac{m\,(\mathbf{q}''-\mathbf{q}')^2}{2(t''-t')}\right]. \tag{6.5b}$$

In the momentum basis the relation (6.3) between the S-matrix and the evolution operator takes the form

$$\langle\mathbf{p}''|\,S\,|\mathbf{p}'\rangle = \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \mathrm{e}^{iE''t''/\hbar}\,\langle\mathbf{p}''|\,U(t'',t')\,|\mathbf{p}'\rangle\,\mathrm{e}^{-iE't'/\hbar}, \tag{6.6}$$

where

$$E' = E(\mathbf{p}'), \quad E'' = E(\mathbf{p}''), \qquad E(\mathbf{p}) = \mathbf{p}^2/2m\,.$$

The limits $t' \to -\infty$, $t'' \to +\infty$ have to be understood mathematically in the sense of distributions (one should use test functions in the form of wave packets).

The S-matrix in the momentum basis is in general parametrized in terms of the scattering matrix $\mathcal{T}$:

$$S = 1 - i\mathcal{T}, \Rightarrow \langle \mathbf{p}'' | S | \mathbf{p}' \rangle = (2\pi\hbar)^d \delta^{(d)}(\mathbf{p}'' - \mathbf{p}') - i \langle \mathbf{p}'' | \mathcal{T} | \mathbf{p}' \rangle . \tag{6.7}$$

When the potential is time-independent, energy is conserved and one can write

$$\langle \mathbf{p}'' | \mathcal{T} | \mathbf{p}' \rangle = -2\pi\delta(E'' - E') T(\mathbf{p}'', \mathbf{p}'). \tag{6.8}$$

Path integrals. To calculate matrix elements of the evolution operator, we can start from expression (2.19) and proceed by analytic continuation replacing t by $t\,\mathrm{e}^{i\varphi}$ in all expressions and rotating in the complex t plane from $\varphi = 0$ to $\varphi = \pi/2$ in the positive direction (but leaving unchanged the argument of $H(t)$).

For example, a hamiltonian of the form

$$H = p^2/2m + V(q,t),$$

leads to the representation

$$\langle q'' | U(t'', t') | q' \rangle = \int_{q(t')=q'}^{q(t'')=q''} [\mathrm{d}q(t)] \exp\left[i\mathcal{A}(q)/\hbar\right]. \tag{6.9}$$

The function $\mathcal{A}(q)$, which replaces the euclidean action is now the usual classical action, integral of the lagrangian:

$$\mathcal{A}(q) = \int_{t'}^{t''} \mathrm{d}t \left[\tfrac{1}{2} m\dot{q}^2 - V(q,t)\right]. \tag{6.10}$$

The expression (6.9) establishes a beautiful relation between classical and quantum mechanics. In the quantum evolution all paths contribute but they are weighted with the complex weight $\mathrm{e}^{i\mathcal{A}/\hbar}$. Therefore, paths close to extrema of the action, that is, paths which are solutions of the classical equations of motion, give the largest contributions to the path integral. In particular, if the value of the classical action on classical paths is large compared to $\hbar$, paths close to the classical paths completely dominate the path integral.

Phase space formulation. In the real time formulation the path integral over phase space (3.9) is replaced by

$$\langle q'' | U(t'', t') | q' \rangle = \int [\mathrm{d}p(t)\mathrm{d}q(t)] \exp\left[i\mathcal{A}(p,q)/\hbar\right]. \tag{6.11}$$

The quantity $\mathcal{A}(p,q)$ which replaces the euclidean action in the path integral is again the classical action in the hamiltonian formalism:

$$\mathcal{A}(p,q) = \int_{t'}^{t''} [p(t)\dot{q}(t) - H(p,q,t)]\, \mathrm{d}t . \tag{6.12}$$

Even in this more general situation the quantum evolution is obtained by summing over all paths weighted with the complex weight $\mathrm{e}^{i\mathcal{A}/\hbar}$, where $\mathcal{A}$ is the classical action in the phase space formalism.

6.2 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 A6.2. 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

$$H = p^2/2m + V(x) \tag{6.13}$$

(setting for convenience $\hbar = 1$). The classical actions corresponding to the free hamiltonian H_0 and to H, respectively, are

$$\mathcal{A}_0(x) = \int_{t'}^{t''} \frac{1}{2} m\dot{x}^2(t)\mathrm{d}t\,, \qquad \mathcal{A}(x) = \int_{t'}^{t''} \left[\tfrac{1}{2}m\dot{x}^2(t) - V\bigl(x(t)\bigr)\right]\mathrm{d}t\,. \tag{6.14}$$

We now expand the path integral (6.9) in powers of V:

$$\begin{aligned}\langle x''\,|U(t'',t')|\,x'\rangle &= \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)] \exp\left[i\mathcal{A}(x)\right] = \sum_\ell \langle x''|\,U^{(\ell)}(t'',t')\,|x'\rangle\,, \\ \langle x''|\,U^{(\ell)}(t'',t')\,|x'\rangle &= \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)]\,\mathrm{e}^{i\mathcal{A}_0(x)} \sum_\ell \frac{(-i)^\ell}{\ell!}\left[\int_{t'}^{t''} V\bigl(x(t)\bigr)\mathrm{d}t\right]^\ell.\end{aligned} \tag{6.15}$$

Since the potential must vanish at large distances it makes sense to assume that it has a Fourier representation:

$$V(x) = (2\pi)^{-d}\int \mathrm{d}^d k\, \tilde{V}(k)\,\mathrm{e}^{ikx}\,. \tag{6.16}$$

Introducing the representation (6.16) in the path integral,

$$\begin{aligned}\langle x''|\,U^{(\ell)}(t'',t')\,|x'\rangle = \frac{(-i)^\ell}{\ell!}\int_{t'}^{t''}\prod_j \mathrm{d}\tau_j \int \prod_{j=1}^{\ell} \tilde{V}(k_j)\frac{\mathrm{d}^d k_j}{(2\pi)^d} \\ \times \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)] \exp i\left[\int_{t'}^{t''}\frac{1}{2}m\dot{x}^2(t)\mathrm{d}t + \sum_j k_j x(\tau_j)\right],\end{aligned} \tag{6.17}$$

we see that each term in the perturbative expansion then involves only the calculation of a gaussian path integral. The integrand in expression (6.17) is symmetric in the times $\tau_1, \tau_2, \ldots, \tau_\ell$. We can, therefore, order them as $t'' \ge \tau_\ell \ge \tau_{\ell-1} \cdots \ge \tau_1 \ge t'$ 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 $\tau_1, \ldots, \tau_\ell$

6.2 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 A6.2. 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

$$H = p^2/2m + V(x) \tag{6.13}$$

(setting for convenience $\hbar = 1$). The classical actions corresponding to the free hamiltonian H_0 and to H, respectively, are

$$\mathcal{A}_0(x) = \int_{t'}^{t''} \frac{1}{2} m\dot{x}^2(t)\mathrm{d}t\,, \qquad \mathcal{A}(x) = \int_{t'}^{t''} \left[\tfrac{1}{2}m\dot{x}^2(t) - V\big(x(t)\big)\right]\mathrm{d}t\,. \tag{6.14}$$

We now expand the path integral (6.9) in powers of V:

$$\begin{aligned}\langle x''\,|U(t'',t')|\,x'\rangle &= \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)] \exp\left[i\mathcal{A}(x)\right] = \sum_{\ell} \langle x''|\, U^{(\ell)}(t'',t')\,|x'\rangle\,,\\ \langle x''|\, U^{(\ell)}(t'',t')\,|x'\rangle &= \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)]\, \mathrm{e}^{i\mathcal{A}_0(x)} \sum_{\ell} \frac{(-i)^{\ell}}{\ell!} \left[\int_{t'}^{t''} V\big(x(t)\big)\mathrm{d}t\right]^{\ell}.\end{aligned} \tag{6.15}$$

Since the potential must vanish at large distances it makes sense to assume that it has a Fourier representation:

$$V(x) = (2\pi)^{-d} \int \mathrm{d}^d k\, \tilde{V}(k)\, \mathrm{e}^{ikx}\,. \tag{6.16}$$

Introducing the representation (6.16) in the path integral,

$$\begin{aligned}\langle x''|\, U^{(\ell)}(t'',t')\,|x'\rangle = {}& \frac{(-i)^{\ell}}{\ell!} \int_{t'}^{t''} \prod_j \mathrm{d}\tau_j \int \prod_{j=1}^{\ell} \tilde{V}(k_j) \frac{\mathrm{d}^d k_j}{(2\pi)^d}\\ &\times \int_{x(t')=x'}^{x(t'')=x''} [\mathrm{d}x(t)] \exp i \left[\int_{t'}^{t''} \frac{1}{2} m\dot{x}^2(t)\mathrm{d}t + \sum_j k_j x(\tau_j)\right],\end{aligned} \tag{6.17}$$

we see that each term in the perturbative expansion then involves only the calculation of a gaussian path integral. The integrand in expression (6.17) is symmetric in the times $\tau_1, \tau_2, \ldots, \tau_\ell$. We can, therefore, order them as $t'' \geq \tau_\ell \geq \tau_{\ell-1} \cdots \geq \tau_1 \geq t'$ 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 $\tau_1, \ldots, \tau_\ell$

the momentum changes by amounts $k_1, \dots, k_\ell$. One then integrates 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:

$$\exp[ik_j x(\tau_j)] = \int \mathrm{d}^d x_j \, \delta\big(x_j - x(\tau_j)\big) \exp\left(ik_j x_j\right).$$

Then in each interval $\tau_j \le t \le \tau_{j+1}$ we recognize the matrix elements of the free evolution operator, which we write in the Fourier representation (6.5a). Moreover, we calculate the Fourier transform with respect to x' and x'', calling p' and p'', respectively, the corresponding momenta. We then find

$$\begin{aligned}\langle p''| \, U^{(\ell)}(t'',t') \, |p'\rangle = (-i)^\ell \int \prod_{j=1}^{\ell} \mathrm{d}\tau_j \tilde{V}(k_j) \frac{\mathrm{d}^d k_j}{(2\pi)^d} \mathrm{d}^d x_j \prod_{j=2}^{\ell} \frac{\mathrm{d}^d p_j}{(2\pi)^d} \\ \times \exp\left[\sum_{j=1}^{\ell+1} -ip_j^2(\tau_j - \tau_{j-1})/2m + ip_j(x_j - x_{j-1}) + ik_j x_j\right],\end{aligned}$$

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 $t'' \to +\infty$, $t' \to -\infty$ can be taken. The corresponding S-matrix elements follow:

$$\begin{aligned}\langle p''| \, S^{(\ell)} \, |p'\rangle = (-i)^\ell \int \prod_{j=1}^{\ell} \mathrm{d}\tau_j \prod_{j=2}^{\ell} \frac{\mathrm{d}^d p_j}{(2\pi)^d} \, \mathrm{e}^{ip''^2 \tau_\ell/2m} \, \tilde{V}(p'' - p_\ell) \, \mathrm{e}^{ip_\ell^2(\tau_{\ell-1} - \tau_\ell)/2m} \times \\ \cdots \times \mathrm{e}^{ip_2^2(\tau_1 - \tau_2)/2m} \, \tilde{V}(p_2 - p') \, \mathrm{e}^{-ip'^2 \tau_1/2m} .\end{aligned} \tag{6.18}$$

We still have to integrate over the times τ_j. We set

$$\tau_{j+1} = \tau_j + u_j \,, \quad u_j \ge 0 .$$

The remaining integral over τ_1 yields a δ-function of energy conservation: $2\pi\delta(E'' - E')$, which follows from time translation invariance. The integrals over the variables u_j on the positive axis yield mathematical distributions:

$$\int_0^{+\infty} \mathrm{d}u_j \, \mathrm{e}^{i(E'' - E(p_j))u_j} = \frac{i}{E'' - E(p_j) + i\epsilon} \,, \quad E(p) \equiv p^2/2m \,;$$

the $i\epsilon$, $\epsilon \to 0_+$, contribution 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

$$\langle p''| \, S^{(\ell)} \, |p'\rangle = -2i\pi\delta(E'' - E') \int \tilde{V}(p'' - p_\ell) \prod_j \frac{\mathrm{d}^d p_j}{(2\pi)^d} \frac{\tilde{V}(p_j - p_{j-1})}{E'' + i\epsilon - E(p_j)} .$$

In this form the perturbation series is a geometric series whose sum satisfies an integral equation, called Lippman–Schwinger's equation. In terms of the operator $T(E)$, where E is generically complex, solution of

$$T(E) = V - VG_0(E)T(E) \,, \quad \text{with} \quad G_0(E) = (H_0 - E)^{-1},$$

the quantity $T(p'', p')$ which appears in equation (6.8) is given by

$$T(p'', p') = \langle p''| \, T(E + i\epsilon) \, |p'\rangle \ \text{for} \ E = p'^2/2m = p''^2/2m.$$

the momentum changes by amounts $k_1, \dots, k_\ell$. One then integrates 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:

$$\exp[ik_j x(\tau_j)] = \int \mathrm{d}^d x_j \, \delta\big(x_j - x(\tau_j)\big) \exp\left(ik_j x_j\right).$$

Then in each interval $\tau_j \le t \le \tau_{j+1}$ we recognize the matrix elements of the free evolution operator, which we write in the Fourier representation (6.5a). Moreover, we calculate the Fourier transform with respect to x' and x'', calling p' and p'', respectively, the corresponding momenta. We then find

$$\begin{aligned}\langle p''| \, U^{(\ell)}(t'', t') \, |p'\rangle = (-i)^\ell \int \prod_{j=1}^{\ell} \mathrm{d}\tau_j \tilde{V}(k_j) \frac{\mathrm{d}^d k_j}{(2\pi)^d} \mathrm{d}^d x_j \prod_{j=2}^{\ell} \frac{\mathrm{d}^d p_j}{(2\pi)^d} \\ \times \exp\left[\sum_{j=1}^{\ell+1} -ip_j^2(\tau_j - \tau_{j-1})/2m + ip_j(x_j - x_{j-1}) + ik_j x_j\right],\end{aligned}$$

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 $t'' \to +\infty$, $t' \to -\infty$ can be taken. The corresponding S-matrix elements follow:

$$\begin{aligned}\langle p''| \, S^{(\ell)} \, |p'\rangle = (-i)^\ell \int \prod_{j=1}^{\ell} \mathrm{d}\tau_j \prod_{j=2}^{\ell} \frac{\mathrm{d}^d p_j}{(2\pi)^d} \, \mathrm{e}^{ip''^2 \tau_\ell/2m} \, \tilde{V}(p'' - p_\ell) \, \mathrm{e}^{ip_\ell^2(\tau_{\ell-1} - \tau_\ell)/2m} \times \\ \cdots \times \mathrm{e}^{ip_2^2(\tau_1 - \tau_2)/2m} \, \tilde{V}(p_2 - p') \, \mathrm{e}^{-ip'^2 \tau_1/2m} . \end{aligned} \tag{6.18}$$

We still have to integrate over the times τ_j. We set

$$\tau_{j+1} = \tau_j + u_j \,, \quad u_j \ge 0 .$$

The remaining integral over τ_1 yields a δ-function of energy conservation: $2\pi\delta(E'' - E')$, which follows from time translation invariance. The integrals over the variables u_j on the positive axis yield mathematical distributions:

$$\int_0^{+\infty} \mathrm{d}u_j \, \mathrm{e}^{i(E'' - E(p_j))u_j} = \frac{i}{E'' - E(p_j) + i\epsilon}, \quad E(p) \equiv p^2/2m ;$$

the $i\epsilon$, $\epsilon \to 0_+$, contribution 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

$$\langle p''| \, S^{(\ell)} \, |p'\rangle = -2i\pi\delta(E'' - E') \int \tilde{V}(p'' - p_\ell) \prod_j \frac{\mathrm{d}^d p_j}{(2\pi)^d} \frac{\tilde{V}(p_j - p_{j-1})}{E'' + i\epsilon - E(p_j)} .$$

In this form the perturbation series is a geometric series whose sum satisfies an integral equation, called Lippman–Schwinger's equation. In terms of the operator $T(E)$, where E is generically complex, solution of

$$T(E) = V - VG_0(E)T(E) \,, \quad \text{with} \quad G_0(E) = (H_0 - E)^{-1},$$

the quantity $T(p'', p')$ which appears in equation (6.8) is given by

$$T(p'', p') = \langle p''| \, T(E + i\epsilon) \, |p'\rangle \ \text{ for } E = p'^2/2m = p''^2/2m.$$

6.3 Path Integral and S-Matrix: Semi-Classical Expansions

The representation (6.9) 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.

6.3.1 Path integral and S-matrix

Let us calculate the elements of the S-matrix between two wave packets:

$$\langle\psi_2|\,S\,|\psi_1\rangle = \lim_{\substack{t'\to-\infty\\ t''\to+\infty}} \int \mathrm{d}q'\mathrm{d}q''\, \langle\psi_2|\,\mathrm{e}^{iH_0 t''/\hbar}\,|\mathbf{q}''\rangle\,\langle\mathbf{q}''|\,U(t'',t')\,|\mathbf{q}'\rangle\,\langle\mathbf{q}'|\,\mathrm{e}^{-iH_0 t'/\hbar}\,|\psi_1\rangle\,. \tag{6.19}$$

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

$$\psi_1(\mathbf{q},t) = \langle q|\,\mathrm{e}^{-iH_0 t/\hbar}\,|\psi_1\rangle = \int \frac{\mathrm{d}^d p}{(2\pi)^d}\tilde{\psi}_1(\mathbf{p}) \exp\left[i\left(\mathbf{p}\cdot\mathbf{q} - t\frac{\mathbf{p}^2}{2m}\right)/\hbar\right], \tag{6.20}$$

and a similar expression for ψ_2.

When t becomes large, the phase in expression (6.20) varies rapidly, and the integral is then dominated by the stationary points of the phase:

$$\frac{\partial}{\partial\mathbf{p}}\left(\mathbf{p}\cdot\mathbf{q} - t\frac{\mathbf{p}^2}{2m}\right) = 0 \quad\Longrightarrow\quad \mathbf{q} = t\frac{\mathbf{p}}{m}\,. \tag{6.21}$$

Integral (6.20) is thus equivalent to

$$\psi_1(\mathbf{q},t) \underset{|t|\to\infty}{\sim} \tilde{\psi}_1(\mathbf{p})\frac{1}{(2\pi)^{d/2}}\left(\frac{m\hbar}{|t|}\right)^{d/2} \exp\left(\frac{i\pi}{4}\operatorname{sgn} t + it\frac{\mathbf{p}^2}{2m\hbar}\right), \tag{6.22}$$

with

$$\mathbf{p} = \frac{m}{t}\mathbf{q}\,.$$

We then change variables in integral (6.19), setting

$$\mathbf{q}' = \frac{t'}{m}\mathbf{p}', \qquad \mathbf{q}'' = \frac{t''}{m}\mathbf{p}'', \tag{6.23}$$

and obtain

$$\begin{aligned}\langle\psi_2|\,S\,|\psi_1\rangle \propto \lim_{\substack{t'\to-\infty\\ t''\to+\infty}} &\int \mathrm{d}p'\mathrm{d}p''\tilde{\psi}_2^*(\mathbf{p}'')\tilde{\psi}_1(\mathbf{p}') \exp\left[\frac{i}{\hbar}\left(t''\frac{\mathbf{p}''^2}{2m} - t'\frac{\mathbf{p}'^2}{2m}\right)\right] \\ &\times \langle t''\mathbf{p}''/m|\,U(t'',t')\,|t'\mathbf{p}'/m\rangle\,.\end{aligned} \tag{6.24}$$

In this equation, we now introduce the path integral representation (6.9) of the evolution operator:

$$\langle t''\mathbf{p}''/m|\,U(t'',t')\,|t'\mathbf{p}'/m\rangle = \int_{q(t')=t'p'/m}^{q(t'')=t''p''/m} [\mathrm{d}q(t)] \exp\bigl(i\mathcal{A}(q)/\hbar\bigr).$$

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 time of the free classical equation of motion. 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 below.

6.3 Path Integral and S-Matrix: Semi-Classical Expansions

The representation (6.9) 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.

6.3.1 Path integral and S-matrix

Let us calculate the elements of the S-matrix between two wave packets:

$$\langle\psi_2|\,S\,|\psi_1\rangle = \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \int \mathrm{d}q'\mathrm{d}q''\,\langle\psi_2|\,\mathrm{e}^{iH_0t''/\hbar}\,|\mathbf{q}''\rangle\,\langle\mathbf{q}''|\,U(t'',t')\,|\mathbf{q}'\rangle\,\langle\mathbf{q}'|\,\mathrm{e}^{-iH_0t'/\hbar}\,|\psi_1\rangle\,. \tag{6.19}$$

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

$$\psi_1(\mathbf{q},t) = \langle q|\,\mathrm{e}^{-iH_0t/\hbar}\,|\psi_1\rangle = \int \frac{\mathrm{d}^d p}{(2\pi)^d}\tilde{\psi}_1(\mathbf{p})\exp\left[i\left(\mathbf{p}\cdot\mathbf{q} - t\frac{\mathbf{p}^2}{2m}\right)/\hbar\right], \tag{6.20}$$

and a similar expression for ψ_2.

When t becomes large, the phase in expression (6.20) varies rapidly, and the integral is then dominated by the stationary points of the phase:

$$\frac{\partial}{\partial\mathbf{p}}\left(\mathbf{p}\cdot\mathbf{q} - t\frac{\mathbf{p}^2}{2m}\right) = 0 \quad\Longrightarrow\quad \mathbf{q} = t\frac{\mathbf{p}}{m}\,. \tag{6.21}$$

Integral (6.20) is thus equivalent to

$$\psi_1(\mathbf{q},t) \underset{|t|\to\infty}{\sim} \tilde{\psi}_1(\mathbf{p})\frac{1}{(2\pi)^{d/2}}\left(\frac{m\hbar}{|t|}\right)^{d/2} \exp\left(\frac{i\pi}{4}\operatorname{sgn} t + it\frac{\mathbf{p}^2}{2m\hbar}\right), \tag{6.22}$$

with

$$\mathbf{p} = \frac{m}{t}\mathbf{q}\,.$$

We then change variables in integral (6.19), setting

$$\mathbf{q}' = \frac{t'}{m}\mathbf{p}', \qquad \mathbf{q}'' = \frac{t''}{m}\mathbf{p}'', \tag{6.23}$$

and obtain

$$\begin{aligned}\langle\psi_2|\,S\,|\psi_1\rangle \propto \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \int \mathrm{d}p'\mathrm{d}p''\tilde{\psi}_2^*(\mathbf{p}'')\tilde{\psi}_1(\mathbf{p}')\exp\left[\frac{i}{\hbar}\left(t''\frac{\mathbf{p}''^2}{2m} - t'\frac{\mathbf{p}'^2}{2m}\right)\right] \\ \times\,\langle t''\mathbf{p}''/m|\,U(t'',t')\,|t'\mathbf{p}'/m\rangle\,. \end{aligned} \tag{6.24}$$

In this equation, we now introduce the path integral representation (6.9) of the evolution operator:

$$\langle t''\mathbf{p}''/m|\,U(t'',t')\,|t'\mathbf{p}'/m\rangle = \int_{q(t')=t'p'/m}^{q(t'')=t''p''/m} [\mathrm{d}q(t)]\exp\bigl(i\mathcal{A}(q)/\hbar\bigr).$$

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 time of the free classical equation of motion. 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 below.

6.3.2 One-dimension: semi-classical limit

To illustrate previous considerations we consider the hamiltonian

$$H = p^2/2m + V(x),$$

and perform the explicit calculation for $\hbar \to 0$ of the S-matrix using the path integral representation and the steepest descent method. We assume a generic analytic potential, decreasing fast enough at large distance.

Forward scattering. We assume that the energy is larger than the maximal value of the potential so that classical forward scattering is possible. We then solve the equation of motion. Integrating once we obtain

$$\tfrac{1}{2}m\dot{x}^2(\tau) + V(x) = \tfrac{1}{2}\kappa^2/m,$$

with the boundary conditions

$$x(\tau') = x', \quad x(\tau'') = x''.$$

We set $X = x'' - x'$, $T = \tau'' - \tau'$. We know from the general analysis of Section 6.3.1 that relevant trajectories correspond to classical scattering, with $mX/T = k$ finite when $\tau' \to -\infty$, $\tau'' \to \infty$. In these limits, the boundary condition

$$T = m\int_{x'}^{x''} \frac{\mathrm{d}x}{\sqrt{\kappa^2 - 2mV(x)}},$$

leads to

$$\kappa = k + \frac{m}{T}\int_{-\infty}^{+\infty} \mathrm{d}x \left(\frac{k}{\sqrt{k^2 - 2mV(x)}} - 1\right) + O\left(T^{-2}\right).$$

The value of the action for the trajectory is then

$$\begin{aligned} \mathcal{A} &= \frac{\kappa^2 T}{2m} - 2\int_{t'}^{t''} V\big(x(\tau)\big)\mathrm{d}\tau = \frac{\kappa^2 T}{2m} - 2m\int_{x'}^{x''} \frac{V(x)\mathrm{d}x}{\sqrt{\kappa^2 - 2mV(x)}} \\ &= \frac{\kappa^2 T}{2m} + \int_{-\infty}^{+\infty} \mathrm{d}x \left(\sqrt{k^2 - 2mV(x)} - k\right) + O\left(T^{-1}\right). \end{aligned}$$

We then calculate the Fourier transform. The result depends only on $x'' - x'$ and thus a factor of momentum conservation $\delta(k'' - k')$ follows. The remaining integral over X is calculated by steepest descent. At leading order, for T large, we need only consider the terms of order T for the saddle point. We find

$$X = k'T/m \ \Rightarrow\ k' = k.$$

As shown in Section 6.1 the terms proportional to T cancel with the free motion factors, and we finally obtain the semi-classical result,

$$\ln S_+(k) \sim \frac{i}{\hbar}\int_{-\infty}^{+\infty} \mathrm{d}x \left(\sqrt{k^2 - 2mV(x)} - k\right).$$

Backward scattering. We now assume instead 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:

$$\mathcal{A} = \frac{m}{2T}\left(x' + x'' - 2x_0\right)^2 + 2\int_{-\infty}^{x_0} \mathrm{d}x \left(\sqrt{k^2 - 2mV(x)} - k\right) + O\left(T^{-1}\right),$$

where x_0 is the reflection point $k^2 = 2mV(x_0)$, and $k = m(2x_0 - x' - x'')/T$.

Since the result depends only on the combination $x' + x''$, we find after Fourier transformation the expected factor $\delta(k'' + k')$. The remaining integral over $X = x' + x''$, calculated by steepest descent, yields

$$2x_0 - X = k'T/m\,, \quad k' = k\,,$$

and thus

$$\ln S_-(k) = \frac{2i}{\hbar}\int_{-\infty}^{x_0} \mathrm{d}x \left(\sqrt{k^2 - 2V(x)} - k\right) + \frac{2ikx_0}{\hbar}\,.$$

Note, finally, 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.

6.3.3 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.

In the absence of a potential, the evolution operator is given by a gaussian path integral that 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 (6.9) and corresponds to the free hamiltonian $H_0 = \mathbf{p}^2/2m$, with $\mathbf{p} \in \mathbb{R}^d$, is

$$\mathbf{q}(t) = \mathbf{q}' + (\mathbf{q}'' - \mathbf{q}')\frac{t - t'}{t'' - t'}\,. \tag{6.25}$$

Translating the integration variables $\mathbf{q}(t)$ by the classical solution (6.25) we still have to calculate a normalization integral which can be obtained by comparing with the exact result ($\hbar = 1$) (6.5b).

The idea of the eikonal approximation is that in the momentum regime,

$$\mathbf{p}' = \mathbf{p} - \mathbf{k}/2, \quad \mathbf{p}'' = \mathbf{p} + \mathbf{k}/2, \qquad \mathbf{p}^2 \to \infty, \quad \mathbf{p}^2 \gg \mathbf{k}^2,$$

the kinetic term $\mathbf{p}^2/2m$ dominates the action (a situation already encountered in Section 2.6.1) and, therefore, the leading contributions to the path integral come from paths close to the straight lines of the free motion (6.25).

We thus calculate the action for the trajectories (6.25), 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 $\mathbf{q}'' - \mathbf{q}' = \mathbf{s}$, $(\mathbf{q}'' + \mathbf{q}')/2 = \mathbf{x}$ we find

$$\langle \mathbf{p} + \mathbf{k}/2|\, U(t'', t')\, |\mathbf{p} - \mathbf{k}/2\rangle \propto \int \mathrm{d}^d s\, \mathrm{d}^d x\, \exp\left[-i\left(\mathbf{p}\cdot\mathbf{s} + \mathbf{k}\cdot\mathbf{x}\right) + i\mathcal{A}(\mathbf{s}, \mathbf{x})\right], \tag{6.26}$$

Backward scattering. We now assume instead 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:

$$\mathcal{A} = \frac{m}{2T}\left(x' + x'' - 2x_0\right)^2 + 2\int_{-\infty}^{x_0} \mathrm{d}x \left(\sqrt{k^2 - 2mV(x)} - k\right) + O\left(T^{-1}\right),$$

where x_0 is the reflection point $k^2 = 2mV(x_0)$, and $k = m(2x_0 - x' - x'')/T$.

Since the result depends only on the combination $x' + x''$, we find after Fourier transformation the expected factor $\delta(k'' + k')$. The remaining integral over $X = x' + x''$, calculated by steepest descent, yields

$$2x_0 - X = k'T/m\,, \quad k' = k\,,$$

and thus

$$\ln S_-(k) = \frac{2i}{\hbar}\int_{-\infty}^{x_0} \mathrm{d}x \left(\sqrt{k^2 - 2V(x)} - k\right) + \frac{2ikx_0}{\hbar}\,.$$

Note, finally, 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.

6.3.3 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.

In the absence of a potential, the evolution operator is given by a gaussian path integral that 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 (6.9) and corresponds to the free hamiltonian $H_0 = \mathbf{p}^2/2m$, with $\mathbf{p} \in \mathbb{R}^d$, is

$$\mathbf{q}(t) = \mathbf{q}' + (\mathbf{q}'' - \mathbf{q}')\frac{t - t'}{t'' - t'}\,. \tag{6.25}$$

Translating the integration variables $\mathbf{q}(t)$ by the classical solution (6.25) we still have to calculate a normalization integral which can be obtained by comparing with the exact result ($\hbar = 1$) (6.5*b*).

The idea of the eikonal approximation is that in the momentum regime,

$$\mathbf{p}' = \mathbf{p} - \mathbf{k}/2, \quad \mathbf{p}'' = \mathbf{p} + \mathbf{k}/2, \qquad \mathbf{p}^2 \to \infty, \quad \mathbf{p}^2 \gg \mathbf{k}^2,$$

the kinetic term $\mathbf{p}^2/2m$ dominates the action (a situation already encountered in Section 2.6.1) and, therefore, the leading contributions to the path integral come from paths close to the straight lines of the free motion (6.25).

We thus calculate the action for the trajectories (6.25), 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 $\mathbf{q}'' - \mathbf{q}' = \mathbf{s}$, $(\mathbf{q}'' + \mathbf{q}')/2 = \mathbf{x}$ we find

$$\langle \mathbf{p} + \mathbf{k}/2|\, U(t'', t')\, |\mathbf{p} - \mathbf{k}/2\rangle \propto \int \mathrm{d}^d s\, \mathrm{d}^d x\, \exp\left[-i\left(\mathbf{p}\cdot\mathbf{s} + \mathbf{k}\cdot\mathbf{x}\right) + i\mathcal{A}(\mathbf{s}, \mathbf{x})\right], \tag{6.26}$$

in which the classical action now is

$$\mathcal{A}(\mathbf{s},\mathbf{x}) = \frac{im}{2}\frac{\mathbf{s}^2}{t''-t'} - i\int_{t'}^{t''} \mathrm{d}t\, V\left(\mathbf{x} - \frac{\mathbf{s}}{2} + \frac{t-t'}{t''-t'}\mathbf{s}\right). \tag{6.27}$$

The normalization in equation (6.26) is determined by comparing with the result (6.5*b*) for the free motion.

Taking the large time limit, and neglecting the contribution of the potential, we find that the integral over $\mathbf{s}$ is dominated by the saddle point:

$$\mathbf{s} = (t''-t')\mathbf{p}/m\,. \tag{6.28}$$

After the substitution (6.28) and the shift $t-(t'+t'')/2 \mapsto t$, and thus $|t| \le (t''-t')/2$, the argument in the potential becomes $\mathbf{x}+t\mathbf{p}/m$. We assume that the potential decreases fast enough for the integral in (6.27) to have a large $t''-t' \to \infty$ limit. The contribution of the potential to the action then becomes

$$-i\int_{-\infty}^{+\infty} \mathrm{d}t\, V\left(\mathbf{x}+t\mathbf{p}/m\right).$$

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 $\mathbf{p}$:

$$\mathbf{b} = \mathbf{x} - \mathbf{p}\left(\mathbf{x}\cdot\mathbf{p}/\mathbf{p}^2\right). \tag{6.29}$$

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.

We finally obtain

$$\langle \mathbf{p}+\mathbf{k}/2|\, U(t'',t')\,|\mathbf{p}-\mathbf{k}/2\rangle \simeq \delta(\mathbf{p}\cdot\mathbf{k})\mathcal{N}(\mathbf{p})\int \mathrm{d}^{d-1}b\,\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{b}}\exp\left[-i\int \mathrm{d}t\, V\left(\frac{\mathbf{p}t}{m}+\mathbf{b}\right)\right] \tag{6.30}$$

with

$$\mathcal{N}(\mathbf{p}) \propto \exp\left[i(t''-t')\frac{\mathbf{p}^2}{2m}\right]. \tag{6.31}$$

The insertion of the result (6.30) into equation (6.6) 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 (6.7,6.8), in the eikonal approximation is

$$T(\mathbf{p}+\mathbf{k}/2,\mathbf{p}-\mathbf{k}/2) \simeq \frac{i\,|\mathbf{p}|}{m}\int\frac{\mathrm{d}^{d-1}b}{(2\pi)^d}\,\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{b}}\left\{\exp\left[-i\int_{-\infty}^{+\infty}\mathrm{d}t\, V\left(\frac{\mathbf{p}t}{m}+\mathbf{b}\right)\right]-1\right\}. \tag{6.32}$$

Application to the Coulomb potential. We now apply the eikonal approximation to the scattering amplitude of a $1/q$ Coulomb-like potential

$$V(\mathbf{q}) = \alpha/\,|\mathbf{q}| \tag{6.33}$$

We immediately note that the integral over the potential in expression (6.32) diverges because the potential decreases too slowly at large distance. Integrating over a finite time interval we find

$$\int_{(t'-t'')/2}^{(t''-t')/2}\mathrm{d}t\, V\left(\frac{\mathbf{p}t}{m}+\mathbf{x}\right) \underset{t''-t'\to\infty}{\sim} \frac{2\alpha m}{p}\ln\left((t''-t')p/mb\right). \tag{6.34}$$

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. In the case of the Coulomb potential only cross-sections are well-defined, not amplitudes.

Factorizing the infinite phase, we can complete the calculation of the scattering amplitude. Integrating over the vector $\mathbf{b}$ we obtain

$$T(\mathbf{p}+\mathbf{k}/2,\mathbf{p}-\mathbf{k}/2) \simeq \frac{i\pi^{(d-1)/2}}{(2\pi)^d}\frac{p}{m}\exp\left[-i\frac{2\alpha m}{p}\ln\bigl((t''-t')p/mb\bigr)\right] \\ \times \frac{\Gamma\left[\frac{1}{2}(d-1)-\theta\right]}{\Gamma(\theta)}\left(\frac{\mathbf{k}^2}{4}\right)^{[\theta+(1-d)/2]} \tag{6.35}$$

with

$$\theta = -i\alpha m/p\,. \tag{6.36}$$

In three dimensions the expression (6.35) is identical to the exact result. It contains also, for $\alpha < 0$, the correct Coulomb bound state energies E_n which are given by the poles of the scattering amplitude:

$$\theta = \frac{d-1}{2}+n \quad \Rightarrow E_n = \frac{\mathbf{p}^2}{2m} = -\frac{2\alpha^2 m}{(d-1+2n)^2}. \tag{6.37}$$

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 there by an approximate summation of ladder and crossed ladder Feynman diagrams.

6.4 S-Matrix and Holomorphic Formalism

The holomorphic formalism has been discussed in Section 5.1, and we adapt here the expressions to real time evolution. 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 5.2.3) and relativistic quantum field theory, as we start discussing in Section 6.6.2.

6.4.1 Path integrals

The path integral representation of the evolution operator is formally obtained by the continuation $t \mapsto it$. Then

$$U(z'',\bar z';t'',t') = \int \left[\frac{\mathrm{d}\bar z(t)\mathrm{d}z(t)}{2i\pi}\right]\exp\left[i\mathcal{A}(z,\bar z)\right],$$

$$\mathcal{A}(z,\bar z) = -i\bar z(t')z(t') - \int_{t'}^{t''}\mathrm{d}t\,\left[i\bar z(t)\dot z(t) + h\bigl(z(t),\bar z(t)\bigr)\right],$$

with the boundary conditions $z(t'') = z''$, $\bar z(t') = \bar z'$.

From the evolution operator we can derive the corresponding S-matrix. Defining the S-matrix by expression (6.3), where H_0 is the hamiltonian of the harmonic oscillator (5.1),

$$H_0 = \tfrac{1}{2}\hat p^2 + \tfrac{1}{2}\omega^2\hat q^2,$$

($\omega > 0$) we find

$$S(z,\bar{z}) = \lim_{\substack{t' \to -\infty \\ t'' \to +\infty}} \int \frac{\mathrm{d}z''\mathrm{d}\bar{z}''}{2i\pi} \frac{\mathrm{d}z'\mathrm{d}\bar{z}'}{2i\pi} \, \mathrm{e}^{-z''\bar{z}''} \, \mathrm{e}^{-z'\bar{z}'} \, \mathrm{e}^{i\omega t''/2} \exp\left(z\bar{z}'' \, \mathrm{e}^{i\omega t''}\right)$$
$$\times U(z'',\bar{z}';t'',t') \, \mathrm{e}^{-i\omega t'/2} \exp\left(z'\bar{z}\, \mathrm{e}^{-i\omega t'}\right).$$

Using equation (5.6), or integrating directly, we obtain

$$S(z,\bar{z}) = \lim_{\substack{t' \to -\infty \\ t'' \to +\infty}} \mathrm{e}^{i\omega t''/2} \, U(z\,\mathrm{e}^{i\omega t''}, \bar{z}\,\mathrm{e}^{-i\omega t'}; t'', t') \, \mathrm{e}^{-i\omega t'/2} . \tag{6.38}$$

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 6.1), the configurations in the path integral which contribute to the S-matrix are, for large time, asymptotic to the solutions of the classical equation of motion. For the harmonic oscillator H_0 this means

$$z(t'') \underset{t'' \to +\infty}{\sim} z\,\mathrm{e}^{i\omega t''}, \qquad \bar{z}(t') \underset{t' \to -\infty}{\sim} \bar{z}\,\mathrm{e}^{-i\omega t'} .$$

6.4.2 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 (6.38) to the hamiltonian (5.26), 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 $t \to \pm\infty$. After a short calculation, one finds

$$S(z,\bar{z}) = \exp\left[z\bar{z} + i\int_{-\infty}^{+\infty} \mathrm{d}t \left(z\,\mathrm{e}^{i\omega t}\,\bar{b}(t) + \bar{z}\,\mathrm{e}^{-i\omega t}\,b(t)\right)\right.$$
$$\left. - \int_{-\infty}^{+\infty} \mathrm{d}t_1\,\mathrm{d}t_2\,\bar{b}(t_1)\theta(t_2 - t_1)\,\mathrm{e}^{-i\omega(t_2-t_1)}\,b(t_2)\right] . \tag{6.39}$$

Using the formalism of Section 5.1 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

$$b(t) = \int_{-\infty}^{+\infty} \mathrm{d}\nu \; \mathrm{e}^{i\nu t}\,\tilde{b}(\nu),$$

one obtains

$$S(z,\bar{z}) = \exp\left[z\bar{z} + 4\pi i\,\mathrm{Re}\big(\bar{z}\tilde{b}(\omega)\big) - 2\pi\int_{-\infty}^{+\infty} \mathrm{d}\nu \frac{i}{\nu - \omega + i\varepsilon} |b(\nu)|^2\right], \tag{6.40}$$

($\omega > 0$) we find

$$S(z,\bar{z}) = \lim_{\substack{t' \to -\infty \\ t'' \to +\infty}} \int \frac{\mathrm{d}z''\mathrm{d}\bar{z}''}{2i\pi} \frac{\mathrm{d}z'\mathrm{d}\bar{z}'}{2i\pi} \,\mathrm{e}^{-z''\bar{z}''}\,\mathrm{e}^{-z'\bar{z}'}\,\mathrm{e}^{i\omega t''/2} \exp\left(z\bar{z}''\,\mathrm{e}^{i\omega t''}\right)$$
$$\times\, U(z'',\bar{z}';t'',t')\,\mathrm{e}^{-i\omega t'/2} \exp\left(z'\bar{z}\,\mathrm{e}^{-i\omega t'}\right).$$

Using equation (5.6), or integrating directly, we obtain

$$S(z,\bar{z}) = \lim_{\substack{t' \to -\infty \\ t'' \to +\infty}} \mathrm{e}^{i\omega t''/2}\, U(z\,\mathrm{e}^{i\omega t''}, \bar{z}\,\mathrm{e}^{-i\omega t'}; t'', t')\,\mathrm{e}^{-i\omega t'/2}. \tag{6.38}$$

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 6.1), the configurations in the path integral which contribute to the S-matrix are, for large time, asymptotic to the solutions of the classical equation of motion. For the harmonic oscillator H_0 this means

$$z(t'') \underset{t'' \to +\infty}{\sim} z\,\mathrm{e}^{i\omega t''}, \qquad \bar{z}(t') \underset{t' \to -\infty}{\sim} \bar{z}\,\mathrm{e}^{-i\omega t'}.$$

6.4.2 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 (6.38) to the hamiltonian (5.26), 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 $t \to \pm\infty$. After a short calculation, one finds

$$S(z,\bar{z}) = \exp\left[z\bar{z} + i\int_{-\infty}^{+\infty} \mathrm{d}t\,\left(z\,\mathrm{e}^{i\omega t}\,\bar{b}(t) + \bar{z}\,\mathrm{e}^{-i\omega t}\,b(t)\right)\right.$$
$$\left. - \int_{-\infty}^{+\infty} \mathrm{d}t_1\,\mathrm{d}t_2\,\bar{b}(t_1)\theta(t_2 - t_1)\,\mathrm{e}^{-i\omega(t_2-t_1)}\,b(t_2)\right]. \tag{6.39}$$

Using the formalism of Section 5.1 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

$$b(t) = \int_{-\infty}^{+\infty} \mathrm{d}\nu\,\mathrm{e}^{i\nu t}\,\tilde{b}(\nu),$$

one obtains

$$S(z,\bar{z}) = \exp\left[z\bar{z} + 4\pi i\,\mathrm{Re}\big(\bar{z}\tilde{b}(\omega)\big) - 2\pi\int_{-\infty}^{+\infty} \mathrm{d}\nu \frac{i}{\nu - \omega + i\varepsilon} |b(\nu)|^2\right], \tag{6.40}$$

where $i\varepsilon$ means limit of $i\varepsilon$, with ε real, when $\varepsilon \to 0_+$.

Coupling to position only. If the function $b(t)$ is real the perturbation is coupled to the position $q(t)$ only. From (5.36) we infer that the coefficient of $b(t)$ is $\sqrt{2\omega}q(t)$. Then the expression (6.39) can be symmetrized in time and becomes

$$\begin{aligned} S(z,\bar z) = \exp\biggl[z\bar z + i\int_{-\infty}^{+\infty} \mathrm{d}t \left(z\,\mathrm{e}^{i\omega t} + \bar z\,\mathrm{e}^{-i\omega t}\right) b(t) \\ -\frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1\,\mathrm{d}t_2\, b(t_1)\,\mathrm{e}^{-i\omega|t_2-t_1|}\, b(t_2)\biggr]. \end{aligned} \tag{6.41}$$

In terms of Fourier components one finds

$$S(z,\bar z) = \exp\left[z\bar z + 2\pi i\left(z\tilde b(-\omega) + \bar z\tilde b(\omega)\right) - \pi\int_{-\infty}^{+\infty} \mathrm{d}\nu \frac{2i\omega}{\nu^2-\omega^2+i\varepsilon}|b(\nu)|^2\right]. \tag{6.42}$$

It is interesting to compare this expression with a direct calculation of the real path integral

$$\mathcal{Z}(b) = \int [\mathrm{d}q]\exp i\int_{-\infty}^{+\infty} \mathrm{d}t\left(\tfrac{1}{2}\dot q^2(t) - \tfrac{1}{2}\omega^2 q^2(t) + \sqrt{2\omega}b(t)q(t)\right).$$

As such the path integral is ill-defined 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 $t \mapsto t\,\mathrm{e}^{i\theta}$ where θ varies between 0 (the euclidean theory) and $\pi/2$, (the real time theory). In the Fourier variable ν the corresponding rotation is $\nu \mapsto \nu\,\mathrm{e}^{-i\theta}$. 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 $\varepsilon \to 0_+$. The $i\varepsilon$ term then ensures the convergence of the gaussian path integral and one finds

$$\mathcal{Z}(b) = \mathcal{Z}(0)\exp\left[-\frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t_1\,\mathrm{d}t_2\, b(t_1)\,\mathrm{e}^{-i\omega|t_2-t_1|}\, b(t_2)\right].$$

We recognize the contribution quadratic in b in the expression (6.41). Therefore, with this prescription the S-matrix is given in terms of a modified path integral

$$\mathcal{Z}(b) = \int [\mathrm{d}q]\exp i\int_{-\infty}^{+\infty} \mathrm{d}t\left(\tfrac{1}{2}\dot q^2(t) - \tfrac{1}{2}\omega^2 q^2(t) + \sqrt{2\omega}b(t)\left(q(t)+q_0(t)\right)\right) \tag{6.43}$$

$$q_0(t) = (2\omega)^{-1/2}\left(z\,\mathrm{e}^{i\omega t} + \bar z\,\mathrm{e}^{-i\omega t}\right).$$

We note that the function $q_0(t)$ is the most general solution of the equation of motion of the classical harmonic oscillator. Then,

$$S(z,\bar z) = \mathrm{e}^{z\bar z}\,\mathcal{Z}(b)/\mathcal{Z}(0).$$

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 (6.43) can be written as

$$\mathcal{A}(q) = \int_{-\infty}^{+\infty} \mathrm{d}t\left(\tfrac{1}{2}\dot q^2(t) - \tfrac{1}{2}\omega^2 q^2(t) - \tfrac{1}{2}\dot q_0^2 + \tfrac{1}{2}\omega^2 q_0^2 + \sqrt{2\omega}b(t)q(t)\right),$$

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.

where $i\varepsilon$ means limit of $i\varepsilon$, with ε real, when $\varepsilon \to 0_+$.

Coupling to position only. If the function $b(t)$ is real the perturbation is coupled to the position $q(t)$ only. From (5.36) we infer that the coefficient of $b(t)$ is $\sqrt{2\omega}q(t)$. Then the expression (6.39) can be symmetrized in time and becomes

$$\begin{aligned} S(z,\bar{z}) = \exp \Bigl[z\bar{z} + i \int_{-\infty}^{+\infty} \mathrm{d}t \left(z\,\mathrm{e}^{i\omega t} + \bar{z}\,\mathrm{e}^{-i\omega t} \right) b(t) \\ - \frac{1}{2} \int_{-\infty}^{+\infty} \mathrm{d}t_1\, \mathrm{d}t_2\, b(t_1)\, \mathrm{e}^{-i\omega|t_2-t_1|}\, b(t_2) \Bigr] . \end{aligned} \tag{6.41}$$

In terms of Fourier components one finds

$$S(z,\bar{z}) = \exp \left[z\bar{z} + 2\pi i \bigl(z\tilde{b}(-\omega) + \bar{z}\tilde{b}(\omega) \bigr) - \pi \int_{-\infty}^{+\infty} \mathrm{d}\nu \frac{2i\omega}{\nu^2 - \omega^2 + i\varepsilon} |b(\nu)|^2 \right] . \tag{6.42}$$

It is interesting to compare this expression with a direct calculation of the real path integral

$$\mathcal{Z}(b) = \int [\mathrm{d}q] \exp i \int_{-\infty}^{+\infty} \mathrm{d}t \left(\tfrac{1}{2}\dot{q}^2(t) - \tfrac{1}{2}\omega^2 q^2(t) + \sqrt{2\omega} b(t) q(t) \right) .$$

As such the path integral is ill-defined 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 $t \mapsto t\,\mathrm{e}^{i\theta}$ where θ varies between 0 (the euclidean theory) and $\pi/2$, (the real time theory). In the Fourier variable ν the corresponding rotation is $\nu \mapsto \nu\,\mathrm{e}^{-i\theta}$. 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 $\varepsilon \to 0_+$. The $i\varepsilon$ term then ensures the convergence of the gaussian path integral and one finds

$$\mathcal{Z}(b) = \mathcal{Z}(0) \exp \left[-\frac{1}{2} \int_{-\infty}^{+\infty} \mathrm{d}t_1\, \mathrm{d}t_2\, b(t_1)\, \mathrm{e}^{-i\omega|t_2-t_1|}\, b(t_2) \right] .$$

We recognize the contribution quadratic in b in the expression (6.41). Therefore, with this prescription the S-matrix is given in terms of a modified path integral

$$\begin{aligned} \mathcal{Z}(b) &= \int [\mathrm{d}q] \exp i \int_{-\infty}^{+\infty} \mathrm{d}t \left(\tfrac{1}{2}\dot{q}^2(t) - \tfrac{1}{2}\omega^2 q^2(t) + \sqrt{2\omega} b(t) \bigl(q(t) + q_0(t) \bigr) \right) \\ q_0(t) &= (2\omega)^{-1/2} \left(z\,\mathrm{e}^{i\omega t} + \bar{z}\,\mathrm{e}^{-i\omega t} \right) . \end{aligned} \tag{6.43}$$

We note that the function $q_0(t)$ is the most general solution of the equation of motion of the classical harmonic oscillator. Then,

$$S(z,\bar{z}) = \mathrm{e}^{z\bar{z}}\, \mathcal{Z}(b)/\mathcal{Z}(0).$$

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 (6.43) can be written as

$$\mathcal{A}(q) = \int_{-\infty}^{+\infty} \mathrm{d}t \left(\tfrac{1}{2}\dot{q}^2(t) - \tfrac{1}{2}\omega^2 q^2(t) - \tfrac{1}{2}\dot{q}_0^2 + \tfrac{1}{2}\omega^2 q_0^2 + \sqrt{2\omega} b(t) q(t) \right) ,$$

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.

6.4.3 The Bose gas

The preceding formalism extends to the Bose gas discussed in Section 5.5. The evolution operator, in the formalism of second quantization in the presence of a chemical potential μ coupled to the particle number $\mathbf{N}$, is given by a functional integral, continuation to real time of the expression (5.92). Here we find

$$\langle\varphi''|\,\mathbf{U}(t'',t')\,|\bar\varphi'\rangle = \langle\varphi''|\,\mathrm{e}^{-i(t''-t)(\mathbf{H}-\mu\mathbf{N})/\hbar}\,|\bar\varphi'\rangle = \int[\mathrm{d}\bar\varphi(t,x)\mathrm{d}\varphi(t,x)]\exp[i\mathcal{A}(\varphi,\bar\varphi)/\hbar], \tag{6.44}$$

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 $\mathcal{A}(\varphi,\bar\varphi)$ is the continuation of the expression (5.93),

$$\begin{aligned}\mathcal{A}(\varphi,\bar\varphi) &= -i\hbar\bar\varphi(t,x')\varphi(t,x') + \int \mathrm{d}t\,\mathrm{d}^dx\,\bar\varphi(t,x)\left(-i\hbar\frac{\partial}{\partial t}-\frac{\hbar^2}{2m}\nabla_x^2 - V_1(x)+\mu\right)\varphi(t,x)\\ &\quad -\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^dx\,\mathrm{d}^dy\,\bar\varphi(t,x)\varphi(t,x)V_2(x,y)\bar\varphi(t,y)\varphi(t,y).\end{aligned} \tag{6.45}$$

In the absence of an external potential V_1 and for a pseudo-potential $V_2 = G\delta(x-y)$ the action simplifies and becomes local:

$$\begin{aligned}\mathcal{A}(\varphi,\bar\varphi) &= -i\hbar\bar\varphi(t,x')\varphi(t,x') + \int \mathrm{d}t\,\mathrm{d}^dx\left[\bar\varphi(t,x)\left(-i\hbar\frac{\partial}{\partial t}-\frac{\hbar^2}{2m}\nabla_x^2+\mu\right)\varphi(t,x)\right.\\ &\quad\left.-\frac{1}{2}G\big(\bar\varphi(t,x)\varphi(t,x)\big)^2\right].\end{aligned} \tag{6.46}$$

Note that below, but near, the transition temperature of the Bose gas discussed in 5.5.4 (see also Section 28.3), the field φ is almost classical for small coupling. The functional integral (6.44) in the limit $\hbar\to 0$ can then be evaluated by the stationary phase approximation, that is, by replacing the field φ by a solution of the equation $\delta\mathcal{A}/\delta\varphi = \delta\mathcal{A}/\delta\bar\varphi = 0$. 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. This equation has the form of a non-linear Schrödinger equation and is called the Gross–Pitaevski equation.

6.5 Fermi Gas: Evolution Operator

With the conventions of Section 5.3, for a finite number of fermions and a normal-ordered hamiltonian of the form $h(\boldsymbol{\theta}, \partial/\partial\boldsymbol{\theta})$, the evolution operator is given by the path integral

$$\langle \boldsymbol{\theta}''| U(t'',t') |\bar{\boldsymbol{\theta}}'\rangle = \int \left[\prod_\alpha \mathrm{d}\theta_\alpha(t)\mathrm{d}\bar\theta_\alpha(t)\right] \exp i\mathcal{A}(\boldsymbol{\theta},\bar{\boldsymbol{\theta}}), \tag{6.47a}$$

$$\mathcal{A}(\boldsymbol{\theta},\bar{\boldsymbol{\theta}}) = \int_{t'}^{t''} \mathrm{d}t \left\{ i\bar{\boldsymbol{\theta}}(t)\cdot\dot{\boldsymbol{\theta}}(t) - h\left[\boldsymbol{\theta}(t),\bar{\boldsymbol{\theta}}(t)\right]\right\} + i\bar{\boldsymbol{\theta}}(t')\cdot\boldsymbol{\theta}(t'), \tag{6.47b}$$

real time continuation of the representation (5.77). We recall that hermiticity now is equivalent to $\mathcal{A} = \mathcal{A}^\dagger$.

Note that in the case of a free hamiltonian

$$h(\theta,\bar\theta) = \omega\theta\bar\theta = -\omega\bar\theta\theta\,,$$

if ω is negative the one-particle state has an energy lower than the vacuum or ground state. This simply means that the two states have been misidentified; they have to be interchanged. A simple transformation deals with the problem. We set

$$\eta = \bar\theta\,, \quad \bar\eta = \theta\,.$$

Then, after integration by parts,

$$\mathcal{A}(\eta,\bar\eta) = \int_{t'}^{t''} \mathrm{d}t\left[i\bar\eta(t)\dot\eta(t) + |\omega|\bar\eta(t)\eta(t)\right] - i\bar\eta(t'')\cdot\eta(t''),$$

which now corresponds to one-particle states with positive energy.

The Fermi gas. The generalization of this formalism to fields allows to describe the time evolution of the Fermi gas in the formalism of second quantization introduced in Section 5.6. 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 functional integral, continuation to real time of the expression (5.109). Here we find

$$\langle\varphi''|\,\mathbf{U}(t'',t')\,|\bar\varphi'\rangle = \langle\varphi''|\,\mathrm{e}^{-i(t''-t)(\mathbf{H}-\mu\mathbf{N})/\hbar}\,|\bar\varphi'\rangle = \int[\mathrm{d}\bar\varphi(t,x)\mathrm{d}\varphi(t,x)]\exp[i\mathcal{A}(\varphi,\bar\varphi)/\hbar], \tag{6.48}$$

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 the expression (5.110):

$$\begin{aligned}\mathcal{A}(\varphi,\bar\varphi) = {}& -i\hbar\bar\varphi(t,x')\varphi(t,x') + \int \mathrm{d}t\,\mathrm{d}^d x\,\bar\varphi(t,x)\left(-i\hbar\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\nabla_x^2 - V_1(x) + \mu\right)\varphi(t,x)\\ & -\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^d x\,\mathrm{d}^d y\,\bar\varphi(t,x)\varphi(t,x)V_2(x,y)\bar\varphi(t,y)\varphi(t,y).\end{aligned} \tag{6.49}$$

For example, in the absence of an external potential V_1, the action for N component Fermi particles and a pseudo-potential $V_2 = G\delta_{\alpha\beta}\delta(x-y)$ becomes local:

$$\begin{aligned}\mathcal{A}(\varphi,\bar\varphi) = {}& -i\hbar\bar\varphi^\alpha(t,x')\varphi^\alpha(t,x') + \int \mathrm{d}t\,\mathrm{d}^d x\left[\bar\varphi^\alpha(t,x)\left(-i\hbar\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\nabla_x^2 + \mu\right)\varphi^\alpha(t,x)\right.\\ & \left. -\tfrac{1}{2}G\big(\bar\varphi^\alpha(t,x)\varphi^\alpha(t,x)\big)^2\right].\end{aligned} \tag{6.50}$$

6.6 Relativistic Quantum Field Theory: The Scalar Field

With this section, we begin our study of *local, relativistic* quantum field theory with the simplest example, the scalar field. It is useful to describe its properties first in real time, before returning to the euclidean theory, because some aspects like the relation between fields and particles are easier to understand. The holomorphic formalism will be specially useful in this respect. Various expressions for the scattering S-matrix will follow.

6.6.1 The neutral scalar field

We first discuss the neutral self-coupled scalar (boson) field. An important example is provided by the so-called ϕ^4 theory, ϕ being the scalar field, which as we argue in Section 6.11, has the theory of spinless bosons interacting through pair potentials, described in Section 5.5.4, as a non-relativistic limit.

The classical field $\phi(t, \mathbf{x})$ is real and depends on time t and a $(d-1)$-dimensional space coordinate $\mathbf{x}$. The classical field equations derive from a lagrangian density $\mathcal{L}(\phi)$ of the form (almost everywhere in this work the speed of light c will be set to one)

$$\mathcal{L}(\phi) = \tfrac{1}{2}\left(\partial_t\phi(t,x)\right) - \tfrac{1}{2}\left(\nabla_x\phi(t,x)\right)^2 - V\left(\phi(t,x)\right), \tag{6.51}$$

where $V(\phi)$ is a function expandable in powers of ϕ, a polynomial in the simplest examples like

$$V(\phi) = \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{4!}g\phi^4 .$$

For $g = 0$ (free field theory) the parameter m is the physical mass of the particle associated with the field ϕ. Note, however, that for $g \neq 0$ the parameter m^2 which appears in the lagrangian is not necessarily positive, and thus this traditional notation is slightly misleading.

The expression (6.51) provides a simple example of a lagrangian density having the following properties:

(i) It is *local* in time and space because it depends only on the field $\phi(t, x)$ and its partial derivatives (and not on the product of fields at different points). This property, *locality*, plays a central role in most of this work.

(ii) It is invariant under space and time translations since space and time do not appear explicitly in the expression (6.51).

(iii) It is relativistic invariant, that is, invariant under the pseudo-orthogonal group $O(1, d-1)$ acting linearly on t and $\mathbf{x}$.

(iv) It leads after quantization, for a suitable class of potentials $V(\phi)$, to a hermitian quantum hamiltonian bounded from below, as we verify now.

Quantization and functional integrals. To quantize the classical theory we first calculate the hamiltonian density corresponding to the lagrangian density (6.51). Lagrangian and hamiltonian densities are related by a Legendre transformation (see also Section 1.8) involving $\partial_t\phi$ and $\pi(x)$, conjugate momentum of $\phi(x)$:

$$\mathcal{H}(\pi,\phi) + \mathcal{L}(\partial_t\phi,\phi) - \pi(x)\partial_t\phi(t,x) = 0\,, \tag{6.52}$$

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\partial_t\phi(t,x)} \quad \Leftrightarrow \quad \partial_t\phi(t,x) = \frac{\partial\mathcal{H}}{\partial\pi(x)}. \tag{6.53}$$

The total hamiltonian $\mathbf{H}$ is the integral of the hamiltonian density:

$$\mathbf{H} = \int \mathrm{d}^{d-1}x\, \mathcal{H}\left[\pi(x),\phi(x)\right]. \tag{6.54}$$

The coordinates q_i of quantum mechanics are replaced here by the field $\phi(x)$. The transition from quantum mechanics to field theory can be understood in much the same way as the transition between the discretized action (2.17) and the continuum time limit (2.18).

The quantum hamiltonian $\hat{\mathbf{H}}$ is then obtained by replacing classical fields $\{\pi, \phi\}$ by quantum operators $\{\hat{\pi}, \hat{\phi}\}$ that satisfy the commutation relations

$$\left[\hat{\pi}(x), \hat{\phi}(x')\right] = \frac{\hbar}{i}\delta^{d-1}(x - x'). \tag{6.55}$$

The properties of such a quantum theory can then be studied using the standard methods of operator quantum mechanics. Instead, we now introduce the formalism of functional integrals, which generalize the path integrals studied in Chapter 2 and are closely related to the functional integrals introduced in Section 5.5.2.

In the example of the lagrangian density (6.51) the hamiltonian density reads

$$\mathcal{H}(\pi, \phi) = \tfrac{1}{2}\pi^2(x) + \tfrac{1}{2}\left[\nabla\phi(x)\right]^2 + V\left(\phi(x)\right). \tag{6.56}$$

This hamiltonian has an important property: it is quadratic in the momentum $\pi(x)$. It is a continuum generalization of hamiltonians of the form $\sum_i p_i^2 + V(q)$. Extending to field theory the ideas presented in Section 6.1, we thus immediately obtain a functional integral representation for the matrix elements of the evolution operator $U(t_2, t_1) = \mathrm{e}^{-i(t_2-t_1)\mathbf{H}/\hbar}$:

$$\langle\phi_2|\, U(t_2, t_1)\, |\phi_1\rangle = \int [\mathrm{d}\phi(t, x)] \exp\left[i\mathcal{A}(\phi)/\hbar\right], \tag{6.57}$$

with the boundary conditions $\phi(t_1, \mathbf{x}) = \phi_1(\mathbf{x})$, $\phi(t_2, \mathbf{x}) = \phi_2(\mathbf{x})$, and where $\mathcal{A}(\phi)$ is the classical action, space–time integral of the lagrangian density (6.51):

$$\mathcal{A}(\phi) = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^{d-1}x \left\{\tfrac{1}{2}\left[\left(\partial_t\phi(t, x)\right)^2 - \left(\nabla_x\phi(t, x)\right)^2\right] - V\left(\phi(t, x)\right)\right\}. \tag{6.58}$$

Note that in Schrödinger's formulation of quantum mechanics, wave functions now become functionals of classical fields like $\phi_1(x)$ or $\phi_2(x)$ in expression (6.57), which corresponds to an infinite number of usual variables.

6.6.2 Free field theory and the holomorphic formalism

From the preceding formal construction we learn little about the physical interpretation of relativistic fields, and the relation with quantum particles. As in the example of the non-relativistic field theory of Section 5.5, the holomorphic formalism will provide the proper tool to clarify this point.

To explain the role of the holomorphic formalism in the relativistic theory it is useful to first briefly return to the construction of the free field theory. The free action $\mathcal{A}_0(\phi)$ for the scalar field ϕ is obtained from the general expression (6.58) by specializing to $V(\phi) = m^2\phi^2/2$:

$$\mathcal{A}_0(\phi) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\tfrac{1}{2}(\partial_t\phi)^2 - \tfrac{1}{2}(\nabla_x\phi)^2 - \tfrac{1}{2}m^2\phi^2(t, x)\right]. \tag{6.59}$$

Because the action is quadratic in $\phi(t, x)$, the field ϕ can be considered as a collection of harmonic oscillators. It is thus natural to introduce the holomorphic formalism. Following the remarks of Section 5.2.2 we first write the phase space integral, generalization of the path integral (6.11), and then change variables. The action in the hamiltonian formalism is

$$\mathcal{A}_0(\Pi,\phi)=\int \mathrm{d}t\,\mathrm{d}^{d-1}x\left\{\Pi(t,x)\partial_t\phi(t,x)-\tfrac{1}{2}\Pi^2(t,x)-\tfrac{1}{2}\left[\nabla\phi(t,x)\right]^2-\tfrac{1}{2}m^2\phi^2(t,x)\right\}. \tag{6.60}$$

The different harmonic oscillators decouple in the momentum basis. We thus introduce the fields $\varphi(t,\hat{p}),\bar{\varphi}(t,\hat{p})$, analogues of the complex functions $\bar{z}(t), z(t)$ of Section 6.4.1:

$$\phi(t,x)=\int\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\left[\mathrm{e}^{i\hat{p}x}\,\varphi(t,\hat{p})+\mathrm{e}^{-i\hat{p}x}\,\bar{\varphi}(t,\hat{p})\right], \tag{6.61a}$$

$$\Pi(t,x)=i\int\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\omega(\hat{p})\left[\mathrm{e}^{i\hat{p}x}\,\varphi(t,\hat{p})-\mathrm{e}^{-i\hat{p}x}\,\bar{\varphi}(t,\hat{p})\right] \tag{6.61b}$$

with $\omega(\hat{p})=\sqrt{\hat{p}^2+m^2}$. The sign conventions ensure that when $\bar{\varphi}$ and φ are complex conjugated both Π and ϕ are real.

We have chosen the integration measure $\mathrm{d}^{d-1}\hat{p}/2\omega(\hat{p})$ because it is $O(1,d-1)$ covariant. Indeed, let us introduce the notation (convenient but slightly inconsistent with our euclidean notation) $p=\{p_0,\hat{p}\}$, where p_0 is the energy and $\hat{p}$ the momentum, and

$$p^2=p_0^2-\hat{p}^2,\qquad \delta_+(p^2-m^2)=\delta(p^2-m^2)\theta(p_0), \tag{6.62}$$

where $\theta(s)$ is the step function: $\theta(s)=0$ for $s<0$, $\theta(s)=1$ for $s>0$. Then,

$$\int\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}f(\hat{p})=\int\mathrm{d}^d p\,\delta_+(p^2-m^2)f(\hat{p}).$$

In terms of $\varphi,\bar{\varphi}$ the free action (6.60) becomes

$$\mathcal{A}_0(\varphi,\bar{\varphi})=-(2\pi)^{d-1}\int_{t'}^{t''}\mathrm{d}t\,\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\left[i\bar{\varphi}(t,\hat{p})\partial_t\varphi(t,\hat{p})+\omega(\hat{p})\bar{\varphi}(t,\hat{p})\varphi(t,\hat{p})\right]. \tag{6.63}$$

This formalism leads to a simple interpretation (see also Sections 5.2.3, 5.5, 6.4.1): one particle states are relativistic particles of momentum $\hat{p}$ and energy $\omega(\hat{p})$. Note, however, that the transformation (6.61) is not local in the space. As a consequence the main drawback of the holomorphic formalism is that the locality of the action is no longer apparent.

Functional integral. A representation of the free evolution operator U_0 as a functional integral in the holomorphic formalism follows

$$U_0(t'',t';\varphi'',\bar{\varphi}')=\int\left[\omega^{-1}(\hat{p})\mathrm{d}\varphi(t,\hat{p})\mathrm{d}\bar{\varphi}(t,\hat{p})\right]\exp\left[i\mathcal{A}_0(\varphi,\bar{\varphi})\right],$$

where $\mathcal{A}_0(\varphi,\bar{\varphi})$ is the action (6.63).

Fock's space. In the holomorphic formalism the differences with the non-relativistic example discussed in Section 5.5 are of kinematic nature. For completeness we, therefore, briefly review the construction of the corresponding Fock's space. We work in the momentum representation, where the hamiltonian is diagonal.

Let $\psi(\hat{p})$ 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

$$\langle \Psi_1 | \Psi_2 \rangle = (2\pi)^{d-1} \int \frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})} \psi_1^*(\hat{p})\psi_2(\hat{p}). \tag{6.64}$$

We now introduce a complex field $\varphi(\hat{p})$ and the generating functional $\mathbf{\Psi}(\varphi)$ of general n-particle wave functions for bosons

$$\mathbf{\Psi}(\varphi) = \sum_{n=0}^{\infty} \frac{(2\pi)^{n(d-1)}}{n!} \int \psi(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n) \prod_{i=1}^{n} \frac{\mathrm{d}^{d-1}\hat{p}_i}{2\omega(\hat{p}_i)} \varphi(\hat{p}_i), \tag{6.65}$$

where $\psi(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n)$ is a wave function totally symmetric in the momenta $\hat{p}_i$. The direct sum of all n-particle spaces is called Fock space.

The scalar product of two vectors $\mathbf{\Psi}_1$ and $\mathbf{\Psi}_2$ takes the form

$$\langle \mathbf{\Psi}_2 | \mathbf{\Psi}_1 \rangle = \int [\omega^{-1}(\hat{p})\mathrm{d}\varphi(\hat{p})\mathrm{d}\bar{\varphi}(\hat{p})] \exp\left[-(2\pi)^{d-1} \int \frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})} \bar{\varphi}(\hat{p})\varphi(\hat{p})\right] \overline{\mathbf{\Psi}_2(\varphi)} \mathbf{\Psi}_1(\varphi). \tag{6.66}$$

Operators. The free action (6.63) shows that the one-particle hamiltonian has an energy spectrum of the form $\omega(\hat{p})$. Acting on $\mathbf{\Psi}(\varphi)$ the free hamiltonian is thus represented by the operator

$$\mathbf{H}_0 = \int \mathrm{d}^{d-1}\hat{p}\, \varphi(\hat{p})\omega(\hat{p}) \frac{\delta}{\delta\varphi(\hat{p})} + E_0 \,,$$

where E_0 is the ground state or vacuum, that is, the zero-particle state energy.

The kernel representing the identity which corresponds to the scalar product (6.66) is

$$\mathcal{I}(\varphi, \bar{\varphi}) = \exp\left[(2\pi)^{d-1} \int \frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})} \bar{\varphi}(\hat{p})\varphi(\hat{p})\right]. \tag{6.67}$$

The kernel associated with the hamiltonian follows

$$\mathbf{H}_0 \mapsto \left[\frac{1}{2}(2\pi)^{d-1} \int \mathrm{d}^{d-1}\hat{p}\, \bar{\varphi}(\hat{p})\varphi(\hat{p})\right] \mathcal{I}(\varphi, \bar{\varphi}).$$

Note that the free hamiltonian commutes with the particle number operator $\mathbf{N}$:

$$\mathbf{N} = \int \mathrm{d}^{d-1}\hat{p}\, \varphi(\hat{p}) \frac{\delta}{\delta\varphi(\hat{p})} \mapsto (2\pi)^{d-1} \int \frac{\mathrm{d}^{p-1}\hat{p}}{2\omega(\hat{p})} \varphi(\hat{p})\bar{\varphi}(\hat{p})\, \mathcal{I}(\varphi, \bar{\varphi}) \quad \Rightarrow \quad [\mathbf{N}, \mathbf{H}_0] = 0\,,$$

a property that, in general, no longer holds in the presence of local interactions.

The vacuum energy. The ground state $|0\rangle$ of the hamiltonian H_0 is the zero-particle state, also called the vacuum. 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

$$E_0 = \frac{1}{2} \sum_{\hat{p}} \sqrt{m^2 + \hat{p}^2}\,,$$

Let $\psi(\hat{p})$ 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

$$\langle\Psi_1|\Psi_2\rangle = (2\pi)^{d-1}\int \frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\psi_1^*(\hat{p})\psi_2(\hat{p}). \tag{6.64}$$

We now introduce a complex field $\varphi(\hat{p})$ and the generating functional $\Psi(\varphi)$ of general n-particle wave functions for bosons

$$\Psi(\varphi) = \sum_{n=0}^{\infty}\frac{(2\pi)^{n(d-1)}}{n!}\int \psi(\hat{p}_1,\hat{p}_2,\ldots,\hat{p}_n)\prod_{i=1}^{n}\frac{\mathrm{d}^{d-1}\hat{p}_i}{2\omega(\hat{p}_i)}\varphi(\hat{p}_i), \tag{6.65}$$

where $\psi(\hat{p}_1,\hat{p}_2,\ldots,\hat{p}_n)$ is a wave function totally symmetric in the momenta $\hat{p}_i$. 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

$$\langle\Psi_2|\Psi_1\rangle = \int[\omega^{-1}(\hat{p})\mathrm{d}\varphi(\hat{p})\mathrm{d}\bar{\varphi}(\hat{p})]\exp\left[-(2\pi)^{d-1}\int\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\bar{\varphi}(\hat{p})\varphi(\hat{p})\right]\overline{\Psi_2(\varphi)}\Psi_1(\varphi). \tag{6.66}$$

Operators. The free action (6.63) shows that the one-particle hamiltonian has an energy spectrum of the form $\omega(\hat{p})$. Acting on $\Psi(\varphi)$ the free hamiltonian is thus represented by the operator

$$\mathbf{H}_0 = \int \mathrm{d}^{d-1}\hat{p}\,\varphi(\hat{p})\omega(\hat{p})\frac{\delta}{\delta\varphi(\hat{p})} + E_0\,,$$

where E_0 is the ground state or vacuum, that is, the zero-particle state energy.

The kernel representing the identity which corresponds to the scalar product (6.66) is

$$\mathcal{I}(\varphi,\bar{\varphi}) = \exp\left[(2\pi)^{d-1}\int\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}\bar{\varphi}(\hat{p})\varphi(\hat{p})\right]. \tag{6.67}$$

The kernel associated with the hamiltonian follows

$$\mathbf{H}_0 \mapsto \left[\frac{1}{2}(2\pi)^{d-1}\int \mathrm{d}^{d-1}\hat{p}\,\bar{\varphi}(\hat{p})\varphi(\hat{p})\right]\mathcal{I}(\varphi,\bar{\varphi}).$$

Note that the free hamiltonian commutes with the particle number operator $\mathbf{N}$:

$$\mathbf{N} = \int \mathrm{d}^{d-1}\hat{p}\,\varphi(\hat{p})\frac{\delta}{\delta\varphi(\hat{p})} \mapsto (2\pi)^{d-1}\int\frac{\mathrm{d}^{p-1}\hat{p}}{2\omega(\hat{p})}\varphi(\hat{p})\bar{\varphi}(\hat{p})\,\mathcal{I}(\varphi,\bar{\varphi}) \quad\Rightarrow\quad [\mathbf{N},\mathbf{H}_0]=0\,,$$

a property that, in general, no longer holds in the presence of local interactions.

The vacuum energy. The ground state $|0\rangle$ of the hamiltonian H_0 is the zero-particle state, also called the vacuum. 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

$$E_0 = \frac{1}{2}\sum_{\hat{p}}\sqrt{m^2+\hat{p}^2}\,,$$

which is ill-defined. To give a precise meaning to E_0, it is necessary to quantize in a large box of linear size L and to modify the theory at short distance or at large momenta so that the Fourier modes are cut-off at some momentum scale Λ (a space lattice would provide such a cut-off). The Fourier variables $\hat{p}$ are then quantized:

$$\hat{p} = 2\pi\mathbf{n}/L\,, \qquad \mathbf{n} \in \mathbb{Z}^{d-1},$$

and the vacuum energy becomes

$$E_0 = \frac{1}{2}\sum_{\mathbf{n}} \omega(\hat{p}).$$

For L large, sums can be replaced by integrals and $\mathrm{d}\mathbf{n} = L^{d-1}/(2\pi)^{d-1}\mathrm{d}\hat{p}$. The space volume factorizes, showing as expected that the energy is an extensive quantity:

$$E_0/L^{d-1} = \frac{1}{2}\int^{\Lambda} \frac{\mathrm{d}^{d-1}\hat{p}}{(2\pi)^{d-1}}\sqrt{m^2+\hat{p}^2}\,, \tag{6.68}$$

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 would no longer be the case if the field theory is coupled to the gravitational field). Note, however, that 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 Appendix A18.1).

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 2.4 and Appendix A6.1), is given by

$$\langle 0|\,\mathrm{T}[\tilde{\phi}(t,\hat{p})\tilde{\phi}(\hat{p}',0)]\,|0\rangle = \langle 0|\,\tilde{\phi}(\hat{p})\,\mathrm{e}^{-i\mathbf{H}_0|t|}\,\tilde{\phi}(\hat{p}')]\,|0\rangle = (2\pi)^{1-d}\frac{1}{2\omega(\hat{p})}\delta^{d-1}(\hat{p}+\hat{p}')\,\mathrm{e}^{-i\omega(\hat{p})|t|}\,. \tag{6.69}$$

After Fourier transformation over time one finds

$$\frac{1}{2\pi}\int \mathrm{e}^{ip_0 t}\,\mathrm{d}t\,\langle 0|\,\mathrm{T}[\tilde{\phi}(\hat{p},t)\tilde{\phi}(0,\hat{p}')]\,|0\rangle = \frac{1}{(2\pi)^d}\delta^{d-1}(\hat{p}+\hat{p}')\frac{i}{p_0^2-\omega^2(\hat{p})+i\varepsilon}\,, \tag{6.70}$$

where the $i\varepsilon$ term in the denominator indicates that we have to add a small positive imaginary part. The real time two-point function is a distribution in the mathematical sense, boundary value of an analytic function

$$\frac{i}{p_0^2-\omega^2(\hat{p})+i\varepsilon} \equiv 2\pi\delta(p_0^2-\omega^2(\hat{p})) + i\mathrm{PP}\frac{1}{p_0^2-\omega^2(\hat{p})}\,,$$

where PP means principal part.

6.7 The S-Matrix

Having explored the relation between fields and particles, we can now define the scattering S-matrix. We first calculate explicitly the scattering by an external source. The result will lead to a perturbative expression for the S-matrix in a general interacting theory.

6.7.1 Scattering by an external source

To the free action (6.59) we now add a source term, corresponding to the linear coupling of the field ϕ to an external classical source $J(t,x)$. The resulting action $\mathcal{A}_{\mathrm{G}}$ then takes the form

$$\mathcal{A}_{\mathrm{G}}(\phi)=\mathcal{A}_0(\phi)+\int \mathrm{d}t\,\mathrm{d}x\,J(t,x)\phi(t,x).$$

In terms of the fields $\varphi,\bar\varphi$ (equation (6.61a)) and the Fourier components of the source,

$$J(t,x)=\int \mathrm{e}^{i\hat p x}\,\tilde J(t,\hat p)\mathrm{d}\hat p\,,$$

the action $\mathcal{A}_{\mathrm{G}}$ reads

$$\mathcal{A}_{\mathrm{G}}(\varphi,\bar\varphi)=\mathcal{A}_0(\varphi,\bar\varphi)+(2\pi)^{d-1}\int \mathrm{d}t\frac{\mathrm{d}\hat p}{2\omega(\hat p)}\left[\tilde J(t,-\hat p)\varphi(t,\hat p)+\tilde J(t,\hat p)\bar\varphi(t,\hat p)\right],$$

where $\mathcal{A}_0(\varphi,\bar\varphi)$ is the action (6.63).

A simple adaptation of the expression (6.39) then yields the holomorphic S-matrix S_{G}:

$$\begin{aligned}
S_{\mathrm{G}}(J,\varphi,\bar\varphi)&=\exp\left[(2\pi)^{d-1}\int\frac{\mathrm{d}\hat p}{2\omega(\hat p)}\,K(\hat p)\right],\\
K(\hat p)&=\varphi(\hat p)\bar\varphi(\hat p)+i\int\mathrm{d}t\left[\varphi(\hat p)\,\mathrm{e}^{i\omega(\hat p)t}\,\tilde J(t,-\hat p)+\bar\varphi(\hat p)\,\mathrm{e}^{-i\omega(\hat p)t}\,\tilde J(t,\hat p)\right]\\
&\quad-\frac12\int\mathrm{d}t_1\,\mathrm{d}t_2\,\tilde J(t,-\hat p_1)\,\mathrm{e}^{-i|t_2-t_1|\omega(\hat p)}\,\tilde J(t_2,\hat p),
\end{aligned}$$

where in the last term we have symmetrized in $\hat p\mapsto-\hat p$ and then $t_1\leftrightarrow t_2$.

As explained in Section 6.4.1, the expansion of this functional in powers of φ and $\bar\varphi$ then yields the coefficients of the scattering matrix.

A more useful expression is obtained by introducing the time Fourier components of $\tilde J(t,\hat p)$

$$\tilde J(t,\hat p)=\int\mathrm{d}p_0\ \mathrm{e}^{-ip_0t}\,\tilde J(p_0,\hat p).$$

Then,

$$\begin{aligned}
\ln S_{\mathrm{G}}(J,\varphi,\bar\varphi)&=(2\pi)^{d-1}\int\frac{\mathrm{d}\hat p}{2\omega(\hat p)}\,\varphi(\hat p)\bar\varphi(\hat p)+i(2\pi)^d\int\mathrm{d}\hat p\,\mathrm{d}p_0\,\tilde J(p_0,\hat p)\\
&\quad\times\left[\delta_+(p_0^2-\hat p^2-m^2)\varphi(-\hat p)+\delta_-(p_0^2-\hat p^2-m^2)\bar\varphi(\hat p)\right]\\
&\quad-\frac12(2\pi)^d\int\mathrm{d}\hat p\,\mathrm{d}p_0\,\tilde J(-p_0,-\hat p)\frac{i}{p_0^2-\hat p^2-m^2+i\varepsilon}\tilde J(p_0,\hat p),
\end{aligned}\tag{6.71}$$

where the notation (6.62) and

$$\delta_-(p_0^2-\hat p^2-m^2)=\theta(-p_0)\delta(p_0^2-\hat p^2-m^2),$$

has been used.

In the coefficient of the term quadratic in J we recognize the free two-point function (6.70).

6.7.2 General interacting theory

An interaction term $V_{\rm I}(\phi)$ can then be added to the free action, where ϕ has to be expressed in terms of $\varphi, \bar{\varphi}$,

$$\mathcal{A}(\phi) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\tfrac{1}{2}(\partial_t\phi)^2 - \tfrac{1}{2}(\nabla_x\phi)^2 - \tfrac{1}{2}m^2\phi^2(t,x)\right] - V_{\rm I}(\phi). \tag{6.72}$$

Using one of the functional expressions for the perturbative expansion that we prove in Section 7.2 we obtain the form of the S-matrix for the interacting theory:

$$S(\varphi, \bar{\varphi}) = \exp\left[-iV_{\rm I}\left(\frac{1}{i}\frac{\delta}{\delta J}\right)\right] S_{\rm G}(J, \varphi, \bar{\varphi})\bigg|_{J=0}. \tag{6.73}$$

The S-matrix thus has a Feynman diagram expansion with internal propagators Δ given by the quadratic term in J in expression (6.71) (Section 7.2):

$$\Delta(p_0, \hat{p}) = \frac{i}{p_0^2 - \hat{p}^2 - m^2 + i\varepsilon} \equiv \frac{i}{p^2 - m^2 + i\varepsilon}. \tag{6.74}$$

We note that we have indeed obtained a propagator, which otherwise would be singular on the mass-shell $p^2 = p_0^2 - \hat{p}^2 = m^2$, with the well-known $i\varepsilon$ prescription.

Unitarity. With our conventions the unitarity of the S-matrix takes the functional form

$$\begin{aligned}\int [\mathrm{d}\bar{\varphi}'(\hat{p})\mathrm{d}\varphi'(\hat{p})] S^*(\varphi', \varphi) S(\varphi', \bar{\varphi}) \exp\left[-(2\pi)^{d-1}\int \frac{\mathrm{d}\hat{p}}{2\omega(\hat{p})}\,\varphi'(\hat{p})\bar{\varphi}'(\hat{p})\right] \\ = \exp\left[(2\pi)^{d-1}\int \frac{\mathrm{d}\hat{p}}{2\omega(\hat{p})}\,\varphi(\hat{p})\bar{\varphi}(\hat{p})\right].\end{aligned} \tag{6.75}$$

Discussion. 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 to the interaction terms quadratic in the field. 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 23).

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 $\hat{p}$ and energy $\omega(\hat{p})$. Additional eigenstates have energies at least equal to $2m$.

We have, therefore, to take as vacuum state the true ground state of the complete hamiltonian, 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 ensures 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.

Note, finally, that the functional 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.

6.8 S-Matrix and Field Asymptotic Conditions

Since the action is only local when written in terms of the initial real field ϕ, it would be convenient to find 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 (6.73) of the S-matrix as derived from the holomorphic representation.

6.8.1 The gaussian integral in an external source and S-matrix

We first consider the gaussian theory in an external source

$$\mathcal{Z}_{\mathrm{G}}(J) = \int [\mathrm{d}\phi] \exp\left[i\mathcal{A}_0(\phi) + i\int \mathrm{d}t\,\mathrm{d}^{d-1}x\, J(t,x)\phi(t,x)\right]. \tag{6.76}$$

The action can be written in terms of the Fourier components of the source J and the field ϕ

$$\phi(t,x) = \int \mathrm{d}p\,\mathrm{e}^{i\hat{p}x - ip_0 t}\,\tilde{\phi}(p_0,\hat{p}).$$

One obtains

$$\mathcal{A}_0(\phi) = (2\pi)^d \int \mathrm{d}p_0\,\mathrm{d}\hat{p}\left[\tfrac{1}{2}\tilde{\phi}(-p_0,-\hat{p})(p_0^2 - \hat{p}^2 - m^2)\tilde{\phi}(p_0,\hat{p}) + J(-p_0,-\hat{p})\tilde{\phi}(p_0,\hat{p})\right]. \tag{6.77}$$

Unlike the euclidean functional integral, the functional integral for the evolution operator has convergence problems because classical field equations have non-trivial solutions. This problem has already been discussed in Section 6.4.2, and we use the same strategy here. We define the functional integral as the analytic continuation in time of the euclidean path integral. We perform a rotation in the time complex plane $t \mapsto t\,\mathrm{e}^{i\theta}$ where θ varies between 0 (the euclidean theory) and $\pi/2$, (the Minkowsky theory). In the energy variable p_0 the corresponding rotation is $p_0 \mapsto p_0\,\mathrm{e}^{-i\theta}$. As we have indicated this amounts to adding to m^2 an infinitesimal negative imaginary part which ensures the convergence of the functional integral (6.76) for large fields. The generating functional $\mathcal{Z}_{\mathrm{G}}(J)$ can then be calculated and one finds

$$\ln \mathcal{Z}_{\mathrm{G}}(J) = -\frac{1}{2}(2\pi)^d \int \mathrm{d}p_0\mathrm{d}\hat{p}\, J(p_0,\hat{p})\Delta(p_0,\hat{p})J(-p_0,-\hat{p}), \tag{6.78}$$

where $\Delta(p_0,\hat{p})$ is the free propagator

$$\Delta(p_0,\hat{p}) = \frac{i}{p_0^2 - \hat{p}^2 - m^2 + i\varepsilon} \equiv \frac{i}{p^2 - m^2 + i\varepsilon}. \tag{6.79}$$

We note that the propagator obtained by this prescription is identical to the internal propagator (6.74) which 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 (6.78) and (6.71) we see that the quadratic term is reproduced but not the term linear in the source. This term corresponds to an addition to the field $\tilde{\phi}$:

$$\tilde{\phi}(p_0,\hat{p}) \mapsto \tilde{\phi}(p_0,\hat{p}) + \tilde{\phi}_0(p_0,\hat{p})$$

with

$$\tilde{\phi}_0(p_0,\hat{p}) = \delta(p_0^2 - \hat{p}^2 - m^2)\left[\varphi(\hat{p})\theta(-p_0) + \bar{\varphi}(-\hat{p})\theta(p_0)\right]. \tag{6.80}$$

The additional term $\tilde{\phi}_0(p_0,\hat{p})$ thus is a general solution of the free classical field equation

$$(p_0^2 - \hat{p}^2 - m^2)\tilde{\phi}_0(p_0,\hat{p}) = 0\,,$$

in some specific parametrization. It is non-vanishing only on the mass hyperboloid $p^2 = m^2$. The parametrization (6.80) 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 6.4.2 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 functional integral

$$\mathcal{Z}_{\mathrm{G}}(J) = \int [\mathrm{d}\phi] \exp\left[i\mathcal{A}_0(\phi) - i\mathcal{A}_0(\phi_0) + i\int \mathrm{d}t\,\mathrm{d}^{d-1}x\, J(t,x)\phi(t,x)\right]. \tag{6.81}$$

The interpretation is the following: the field ϕ in the functional integral (6.76) satisfies general free field boundary conditions, $\phi \to \phi_0$, and its two-point function or propagator is then given by equation (6.79).

General interaction. The general functional representation (6.73) can then be rewritten in a different way. Introducing the form (6.81) of Z_{G} in (6.73) and applying the functional derivatives we find:

$$S(\varphi,\bar{\varphi}) = \mathcal{I}(\varphi,\bar{\varphi}) \int [\mathrm{d}\phi] \exp i\left[\mathcal{A}(\phi) - \mathcal{A}_0(\phi_0)\right], \tag{6.82}$$

where $\mathcal{I}$ is the identity kernel (6.67), and the field ϕ satisfies fixed free field boundary conditions.

The expression (6.82), up to the factor $\mathcal{I}$, is a functional integral in the presence of a background field ϕ_0. This functional 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 6.3.1. We have shown that in quantum mechanics 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).

Remark. Considerations based on asymptotic field boundary conditions , or the more direct considerations of Section 6.6.2 lead to the same perturbative S-matrix. However, as suggested by the discussion given at the beginning of Section 6.3.1, the preceding considerations generalize to the scattering of *solitons*, that is, states obtained by expanding the functional integral around finite energy static solutions of the complete classical field equations

$$\frac{\delta \mathcal{A}(\phi)}{\delta \phi(x)} = 0 . \tag{6.83}$$

In this case the S-matrix of soliton scattering is obtained by expanding the functional integral around classical soliton scattering solutions of the complete field equations.

6.8.2 *S-matrix and correlation functions*

Alternatively, let us consider the following expression for a general interaction

$$\mathcal{Z}(\mathcal{J}) = \exp\left[-iV_{\mathrm{I}}\left(\frac{1}{i}\frac{\delta}{\delta J}\right)\right] \exp\left[-\tfrac{1}{2}(\mathcal{J}+J)\Delta(\mathcal{J}+J)\right]\Big|_{J=0} ,$$

where $\mathcal{J}$ is an additional external source. We then expand

$$\begin{aligned}\frac{1}{2}(\mathcal{J}+J)\Delta(\mathcal{J}+J) = \frac{1}{2}\int \mathrm{d}p\, \tilde{J}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{J}(p) + \int \mathrm{d}p\, \tilde{J}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p) \\ + \frac{1}{2}\int \mathrm{d}p\, \tilde{\mathcal{J}}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p).\end{aligned} \tag{6.84}$$

Comparing this expression with expression (6.71) we see that the main formal difference between expressions (6.78) and (6.73) comes from the term linear in J:

$$\frac{i}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p_0,\hat{p}) \mapsto -i\delta(p^2-m^2)\left[\varphi(\hat{p})\theta(-p_0) + \bar{\varphi}(-\hat{p})\theta(p_0)\right]. \tag{6.85}$$

We, therefore, 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 restricting the external momenta to the mass-shell $p^2 = m^2$. 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 (definition (7.79)).

Remark. The relation between correlation functions and S-matrix elements shows that the matrix elements as defined here have disconnected contributions. The new functional

$$\mathcal{T}(\varphi,\bar{\varphi}) = i\ln S(\varphi,\bar{\varphi}), \tag{6.86}$$

is, instead, the generating functional of connected scattering amplitudes (see Section 7.4).

Crossing symmetry. We see that only a linear combination of φ and $\bar{\varphi}$ appears, with identical coefficient, up to the sign of the energy. The sign of p_0 specifies the incoming and outgoing particles. This has deep consequences, specific to relativistic quantum field theory: in $d > 2$ dimensions, the same analytic functions lead to scattering amplitudes of different physical processes, a property known as crossing symmetry.

as suggested by the discussion given at the beginning of Section 6.3.1, the preceding considerations generalize to the scattering of *solitons*, that is, states obtained by expanding the functional integral around finite energy static solutions of the complete classical field equations

$$\frac{\delta \mathcal{A}(\phi)}{\delta \phi(x)} = 0 . \tag{6.83}$$

In this case the S-matrix of soliton scattering is obtained by expanding the functional integral around classical soliton scattering solutions of the complete field equations.

6.8.2 S-matrix and correlation functions

Alternatively, let us consider the following expression for a general interaction

$$\mathcal{Z}(\mathcal{J}) = \exp\left[-iV_{\mathrm{I}}\left(\frac{1}{i}\frac{\delta}{\delta J}\right)\right] \exp\left[-\tfrac{1}{2}(\mathcal{J}+J)\Delta(\mathcal{J}+J)\right]\Big|_{J=0} ,$$

where $\mathcal{J}$ is an additional external source. We then expand

$$\begin{aligned}\frac{1}{2}(\mathcal{J}+J)\Delta(\mathcal{J}+J) &= \frac{1}{2}\int \mathrm{d}p\, \tilde{J}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{J}(p) + \int \mathrm{d}p\, \tilde{J}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p) \\ &\quad + \frac{1}{2}\int \mathrm{d}p\, \tilde{\mathcal{J}}(-p)\frac{i(2\pi)^d}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p).\end{aligned} \tag{6.84}$$

Comparing this expression with expression (6.71) we see that the main formal difference between expressions (6.78) and (6.73) comes from the term linear in J:

$$\frac{i}{p^2-m^2+i\epsilon}\tilde{\mathcal{J}}(p_0,\hat{p}) \mapsto -i\delta(p^2-m^2)\left[\varphi(\hat{p})\theta(-p_0)+\bar{\varphi}(-\hat{p})\theta(p_0)\right]. \tag{6.85}$$

We, therefore, 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 restricting the external momenta to the mass-shell $p^2 = m^2$. 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 (definition (7.79)).

Remark. The relation between correlation functions and S-matrix elements shows that the matrix elements as defined here have disconnected contributions. The new functional

$$\mathcal{T}(\varphi,\bar{\varphi}) = i\ln S(\varphi,\bar{\varphi}), \tag{6.86}$$

is, instead, the generating functional of connected scattering amplitudes (see Section 7.4).

Crossing symmetry. We see that only a linear combination of φ and $\bar{\varphi}$ appears, with identical coefficient, up to the sign of the energy. The sign of p_0 specifies the incoming and outgoing particles. This has deep consequences, specific to relativistic quantum field theory: in $d > 2$ dimensions, the same analytic functions lead to scattering amplitudes of different physical processes, a property known as crossing symmetry.

6.8.3 The ϕ^3 example

Let us illustrate the analysis by calculating a four-point scattering amplitude in the simple ϕ^3 field theory, in the tree approximation. We consider the action

$$\mathcal{A}(\phi) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\tfrac{1}{2}(\partial_t\phi)^2 - \tfrac{1}{2}(\nabla_x\phi)^2 - \tfrac{1}{2}m^2\phi^2(t,x) - \tfrac{1}{3!}g\phi^3(t,x)\right].$$

We introduce the Fourier components of the field ϕ,

$$\phi(t,x) = \int \mathrm{e}^{-ip_0t+i\hat{p}x}\,\tilde{\phi}(p)\mathrm{d}^d p\,. \tag{6.87}$$

In terms of Fourier components, the field equation takes the form

$$\left(p^2 - m^2\right)\tilde{\phi}(p) - \frac{1}{2}g\int \mathrm{d}^d q\,\tilde{\phi}(q)\tilde{\phi}(p-q) = 0\,. \tag{6.88}$$

The equation can be solved as a series in the coupling constant g starting from the solution (6.80) of the free equation:

$$\tilde{\phi}_0(p) = \delta(p^2 - m^2)\left[\varphi(\hat{p})\theta(-p_0) + \bar{\varphi}(-\hat{p})\theta(p_0)\right].$$

The classical solution at order g is

$$\tilde{\phi}(p) = \tilde{\phi}_0(p) + \frac{1}{2}\frac{g}{p^2 - m^2}\int \mathrm{d}^d q\,\tilde{\phi}_0(q)\tilde{\phi}_0(p-q) + O\left(g^2\right).$$

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 6.8.1).

Using equation (6.88) we can rewrite $\mathcal{T}$ in the tree approximation:

$$\mathcal{T}(\phi_0) = i\ln S(\phi_0) = -\frac{1}{12}g(2\pi)^d\int \mathrm{d}^d p_1\,\mathrm{d}^d p_2\,\mathrm{d}^d p_3\,\delta^{(d)}(p_1+p_2+p_3)\tilde{\phi}(p_1)\tilde{\phi}(p_2)\tilde{\phi}(p_3).$$

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

$$-\frac{1}{8}g^2(2\pi)^d\int \delta^{(d)}(p_1+p_2+p_3+p_4)\frac{1}{(p_1+p_2)^2 - m^2}\prod_{i=1}^{4}\mathrm{d}p_i\,\tilde{\phi}_0(p_i).$$

The connected four-particle scattering amplitude is then obtained by differentiating with respect to $\tilde{\phi}_0(p)$. The result is the product of a factor that contains the momentum conservation,

$$(2\pi)^d\delta^{(d)}(p_1+p_2+p_3+p_4),$$

and an amplitude,

$$-\frac{g^2}{(p_1+p_2)^2 - m^2} - \frac{g^2}{(p_1+p_3)^2 - m^2} - \frac{g^2}{(p_1+p_4)^2 - m^2}\,,$$

where $p_i^2 = m^2$. 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 yields five, six ... particle scattering amplitudes.

6.9 Field Renormalization

To describe precisely how the S-matrix can be calculated in a general theory of scalar particles, we still have to discuss field renormalization. This can more conveniently be done in the *euclidean formulation* (see Chapter 7). We define the field $\phi(t,x)$ in such a way that it has zero expectation value. This can always be achieved by a constant shift $\phi(t,x) \mapsto \phi(t,x) - \langle 0|\phi|0\rangle$. We then consider the two-point correlation function $W^{(2)}$. It follows from the analysis of Section 2.4 that it is equal to the vacuum expectation value of the time-ordered product of two fields

$$\delta^{d-1}(\hat{p}+\hat{p}')W^{(2)}(t,\hat{p}) = (2\pi)^{d-1}\, \langle 0|\,\tilde{\phi}(\hat{p})\,\mathrm{e}^{-|t|(\mathbf{H}-E_0)}\,\tilde{\phi}(\hat{p}')\,|0\rangle\,. \tag{6.89}$$

We now adapt the arguments of Section A2.1 to the kinematic properties of an $O(d)$ invariant field theory. We introduce a complete set of eigenvectors $|\mu\rangle$ of $\mathbf{H}$, with real eigenvalues ε_μ and obtain

$$\delta^{d-1}(\hat{p}+\hat{p}')W^{(2)}(t,\hat{p}) = (2\pi)^{d-1}\int \mathrm{d}\mu\, \langle 0|\,\tilde{\phi}(\hat{p})\,|\mu\rangle\, \mathrm{e}^{-|t|\varepsilon(\mu)}\, \langle \mu|\,\tilde{\phi}(\hat{p}')\,|0\rangle\,.$$

The theory being translation invariant, we know that the answer is proportional to $\delta^{d-1}(\hat{p}+\hat{p}')$. In this limit, since $\tilde{\phi}(-p) = \tilde{\phi}^\dagger(p)$, the l.h.s. is a sum of positive terms (a point we have already discussed in Section A2.1). We thus find

$$W^{(2)}(t,\hat{p}) = (2\pi)^{d-1}\int \mathrm{d}\mu\, \rho(\mu,\hat{p})\, \mathrm{e}^{-|t|\varepsilon(\mu)},$$

where ρ is a positive measure. We now calculate the Fourier transform with respect to time

$$W^{(2)}(p_0,\hat{p}) = (2\pi)^{d-1}\int \mathrm{d}t\ \mathrm{e}^{-ip_0t}\int \mathrm{d}\mu\, \rho(\mu,\hat{p})\, \mathrm{e}^{-|t|\varepsilon(\mu)} = (2\pi)^{d-1}\int \mathrm{d}\mu\, \frac{2\varepsilon(\mu)\rho(\mu,\hat{p})}{p_0^2+\varepsilon^2(\mu)}.$$

The $O(d)$ invariance allows to set $\hat{p} = 0$ and replace p_0^2 by p^2 without changing the answer. We conclude

$$W^{(2)}(p) = \int \mathrm{d}\mu \frac{\rho(\mu)}{p^2+\mu^2}\,, \tag{6.90}$$

where $\rho(\mu)$, proportional to $\rho(\mu,0)$, is a positive measure. This form is called the Källen–Lehmann (KL) representation. Since by definition the physical mass m is the lowest energy eigenstate above the ground state, the domain of integration is $\mu \geq m$. Moreover, at $\hat{p} = 0$ 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):

$$\rho(\mu) = Z\delta(\mu - m) + \rho'(\mu), \quad \rho'(\mu) = 0 \text{ for } \mu < 2m\,. \tag{6.91}$$

Conversely the Fourier transform with respect to the energy variable p_0 now is

$$\begin{aligned} W^{(2)}(t,\hat{p}) &= \frac{1}{2\pi}\int \mathrm{d}p_0\ \mathrm{e}^{ip_0t}\, W^{(2)}(p) = \frac{1}{2\pi}\int \mathrm{d}p_0\ \mathrm{e}^{ip_0t}\, W^{(2)}(p)\int \mathrm{d}\mu \frac{\rho(\mu)}{p^2+\mu^2} \\ &= \int \mathrm{d}\mu \frac{\rho(\mu)}{2\sqrt{\hat{p}^2+\mu^2}}\, \mathrm{e}^{-|t|\sqrt{\hat{p}^2+\mu^2}}\,. \end{aligned} \tag{6.92}$$

Returning to the definition (6.89), assuming $t > 0$, taking the derivative with respect to time and the limit $t = 0$ we find

$$\delta^{d-1}(\hat{p}+\hat{p}')\,\frac{\partial}{\partial t}W^{(2)}(\hat{p},t)\bigg|_{t\to 0_+} = -(2\pi)^{d-1}\,\langle 0|\,\tilde{\phi}(\hat{p})(\mathbf{H}-E_0)\tilde{\phi}(\hat{p}')\,|0\rangle\,. \tag{6.93}$$

The product $(\mathbf{H}-E_0)\tilde{\phi}$ in the r.h.s. can be replaced by the commutator $[\mathbf{H},\tilde{\phi}]$. Then, using the reality of the l.h.s., we can take the hermitian part of the operator. Thus,

$$\delta^{d-1}(\hat{p}+\hat{p}')\,\frac{\partial}{\partial t}W^{(2)}(\hat{p},t)\bigg|_{t\to 0_+} = -\tfrac{1}{2}(2\pi)^{d-1}\,\langle 0|\,[\tilde{\phi}(\hat{p}),[\mathbf{H},\tilde{\phi}(\hat{p}')]]\,|0\rangle\,.$$

If the action has the form (6.72) the commutator is proportional to the conjugated momentum. Then,

$$[\tilde{\phi}(\hat{p}),[\mathbf{H},\tilde{\phi}(\hat{p}')]] = (2\pi)^{1-d}\delta^{d-1}(\hat{p}+\hat{p}').$$

It follows that

$$\frac{\partial}{\partial t}W^{(2)}(t,\hat{p})\bigg|_{t\to 0_+} = -\frac{1}{2}\,.$$

Calculating now the l.h.s. from the representation (6.92), we obtain

$$\int \mathrm{d}\mu\,\rho(\mu) = 1\,.$$

We conclude that, except in a free field theory, Z, the residue of the pole at $p^2 = -m^2$, is strictly smaller than 1:

$$0 < Z < 1\,. \tag{6.94}$$

This result has several implications, one being related to the S-matrix. Let us evaluate $W^{(2)}(t,\hat{p})$ for t large. The state of lowest energy, with momentum $\hat{p}$, gives the leading contribution. From the KL representation (6.92) and the decomposition (6.91) we then learn

$$W^{(2)}(t,\hat{p}) \underset{t\to\infty}{\sim} \frac{Z}{2\sqrt{\hat{p}^2+m^2}}\,\mathrm{e}^{-|t|\sqrt{\hat{p}^2+m^2}}\,.$$

If we compare this result with the contribution of a normalized one-particle eigenstate of the hamiltonian $\mathbf{H}$ (see equation (6.69)),

$$(2\pi)^{d-1}\,\langle 1,\hat{p}|\,\mathrm{e}^{-|t|(\mathbf{H}-E_0)}\,|1,\hat{p}'\rangle = \delta^{d-1}(\hat{p}+\hat{p}')\frac{1}{2\sqrt{\hat{p}^2+m^2}}\,\mathrm{e}^{-|t|\sqrt{\hat{p}^2+m^2}},$$

we observe that the field ϕ has only a component $\sqrt{Z}$ on the one-particle states. Another way to formulate the same answer 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 we find the two-point function:

$$\begin{aligned}\langle 0|\,\mathrm{T}[\tilde{\phi}(t,\hat{p})\tilde{\phi}(\hat{p}',0)]\,|0\rangle &= \langle 0|\,\tilde{\phi}(\hat{p})\,\mathrm{e}^{-i\mathbf{H}|t|}\,\tilde{\phi}(\hat{p}')]\,|0\rangle \\ &= (2\pi)^{1-d}\delta^{d-1}(\hat{p}+\hat{p}')\int \mathrm{d}\mu\,\frac{\rho(\mu)}{2\sqrt{\hat{p}^2+\mu^2}}\,\mathrm{e}^{-i|t|\sqrt{\hat{p}^2+\mu^2}}\,.\end{aligned}$$

Returning to the definition (6.89), assuming $t > 0$, taking the derivative with respect to time and the limit $t = 0$ we find

$$\delta^{d-1}(\hat{p}+\hat{p}')\, \frac{\partial}{\partial t} W^{(2)}(\hat{p},t)\Big|_{t\to 0_+} = -(2\pi)^{d-1} \left\langle 0 \right| \tilde{\phi}(\hat{p})(\mathbf{H}-E_0)\tilde{\phi}(\hat{p}') \left| 0 \right\rangle . \qquad (6.93)$$

The product $(\mathbf{H}-E_0)\tilde{\phi}$ in the r.h.s. can be replaced by the commutator $[\mathbf{H},\tilde{\phi}]$. Then, using the reality of the l.h.s., we can take the hermitian part of the operator. Thus,

$$\delta^{d-1}(\hat{p}+\hat{p}')\, \frac{\partial}{\partial t} W^{(2)}(\hat{p},t)\Big|_{t\to 0_+} = -\tfrac{1}{2}(2\pi)^{d-1} \left\langle 0 \right| [\tilde{\phi}(\hat{p}),[\mathbf{H},\tilde{\phi}(\hat{p}')]] \left| 0 \right\rangle .$$

If the action has the form (6.72) the commutator is proportional to the conjugated momentum. Then,

$$[\tilde{\phi}(\hat{p}),[\mathbf{H},\tilde{\phi}(\hat{p}')]] = (2\pi)^{1-d}\delta^{d-1}(\hat{p}+\hat{p}').$$

It follows that

$$\frac{\partial}{\partial t} W^{(2)}(t,\hat{p})\Big|_{t\to 0_+} = -\frac{1}{2}\,.$$

Calculating now the l.h.s. from the representation (6.92), we obtain

$$\int \mathrm{d}\mu\, \rho(\mu) = 1\,.$$

We conclude that, except in a free field theory, Z, the residue of the pole at $p^2 = -m^2$, is strictly smaller than 1:

$$0 < Z < 1\,. \qquad (6.94)$$

This result has several implications, one being related to the S-matrix. Let us evaluate $W^{(2)}(t,\hat{p})$ for t large. The state of lowest energy, with momentum $\hat{p}$, gives the leading contribution. From the KL representation (6.92) and the decomposition (6.91) we then learn

$$W^{(2)}(t,\hat{p}) \underset{t\to\infty}{\sim} \frac{Z}{2\sqrt{\hat{p}^2+m^2}}\, \mathrm{e}^{-|t|\sqrt{\hat{p}^2+m^2}}\,.$$

If we compare this result with the contribution of a normalized one-particle eigenstate of the hamiltonian $\mathbf{H}$ (see equation (6.69)),

$$(2\pi)^{d-1} \left\langle 1,\hat{p} \right| \mathrm{e}^{-|t|(\mathbf{H}-E_0)} \left| 1,\hat{p}' \right\rangle = \delta^{d-1}(\hat{p}+\hat{p}') \frac{1}{2\sqrt{\hat{p}^2+m^2}}\, \mathrm{e}^{-|t|\sqrt{\hat{p}^2+m^2}},$$

we observe that the field ϕ has only a component $\sqrt{Z}$ on the one-particle states. Another way to formulate the same answer 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 we find the two-point function:

$$\begin{aligned}\left\langle 0 \right| \mathrm{T}[\tilde{\phi}(t,\hat{p})\tilde{\phi}(\hat{p}',0)] \left| 0 \right\rangle &= \left\langle 0 \right| \tilde{\phi}(\hat{p})\, \mathrm{e}^{-i\mathbf{H}|t|}\, \tilde{\phi}(\hat{p}')] \left| 0 \right\rangle \\ &= (2\pi)^{1-d}\delta^{d-1}(\hat{p}+\hat{p}') \int \mathrm{d}\mu\, \frac{\rho(\mu)}{2\sqrt{\hat{p}^2+\mu^2}}\, \mathrm{e}^{-i|t|\sqrt{\hat{p}^2+\mu^2}}\,.\end{aligned}$$

It can be verified that its large time behaviour is related to the leading singularity of the measure ρ. Since $\rho(\mu)$ is the sum of a δ-function and a continuous function (for $d \geq 2$), we obtain

$$\langle 0|\, \tilde{\phi}(\hat{p})\, \mathrm{e}^{-i\mathbf{H}|t|}\, \tilde{\phi}(\hat{p}')]\, |0\rangle \underset{|t|\to\infty}{=} Z(2\pi)^{1-d} \frac{1}{2\omega(\hat{p})} \delta^{d-1}(\hat{p}+\hat{p}')\, \mathrm{e}^{-i\omega(\hat{p})|t|} + O(1/t).$$

Introducing $\varphi, \bar{\varphi}$, the properly normalized creation and annihilation operators of the one-particle states, we conclude that for large time the field $\tilde{\phi}(t,\hat{p})$ tends in a weak sense (not in an operator sense, but in all expectation values) towards

$$\tilde{\phi}(t,\hat{p}) \underset{|t|\to\infty}{\sim} \sqrt{Z} \frac{1}{2\omega(\hat{p})} \left[\varphi(-\hat{p}) + \bar{\varphi}(\hat{p})\right].$$

The constant $\sqrt{Z}$ is the field renormalization constant.

Normalized S-matrix elements. To calculate properly normalized S-matrix elements we can calculate with the action $\mathcal{A}(\phi\sqrt{Z})$. Alternatively, if we keep the initial field we have to renormalize the matrix elements.

6.10 S-matrix and Correlation Functions

This section somewhat anticipates concepts and results that are going to be presented in Chapters 7, 9 and 10, but otherwise contains material which has, more naturally, its place in this chapter.

6.10.1 Connected and 1PI correlation functions

Connected correlation functions. By comparing the general form of the S-matrix with correlation functions, we have seen in Section 6.8.2 that connected S-matrix elements are related to the analytic continuation to real time ($t \mapsto it, p_0 \mapsto ip_0$) of euclidean correlation functions in the the mass-shell limit. More precisely in terms of $\mathcal{W}(J)$, the generating functional of connected correlation functions (equation (7.57)),

$$\mathcal{T}(\phi_0) = i\mathcal{W}(\Delta^{-1}J),$$

where J is deduced from equation (6.85), ϕ_0 is given by equation (6.80), Δ is the propagator with a physical pole. Connected correlation functions can then be expressed in terms of amputated functions (inverting equation (7.79))

$$\tilde{W}^{(n)}(p_1,\ldots,p_n) = \left[\prod_{i=1}^{n} \tilde{W}^{(2)}(p_i)\right] \tilde{W}^{(n)}_{\text{amp.}}(p_1,\ldots,p_n). \tag{6.95}$$

The two-point function $\tilde{W}^{(2)}(p)$ has a pole located at $p^2 = -m^2$ where m is the physical mass. Near the pole

$$\tilde{W}^{(2)}(p) \underset{p^2\to -m^2}{\sim} \frac{Z}{p^2+m^2}, \tag{6.96}$$

where Z is the field renormalization constant (Section 6.9). We conclude that the coefficient $\mathcal{T}_{\mathrm{r}}^{(n)}$ of $\mathcal{T}(\phi_0/\sqrt{Z})$ in the expansion in powers of ϕ_0 is given by

$$\mathcal{T}_{\mathrm{r}}^{(n)}(p_1,\ldots,p_n) = iZ^{n/2}\, \tilde{W}^{(n)}_{\text{amp.}}(p_1,\ldots,p_n)\Big|_{p_i^2=-m^2}. \tag{6.97}$$

It can be verified that its large time behaviour is related to the leading singularity of the measure ρ. Since $\rho(\mu)$ is the sum of a δ-function and a continuous function (for $d \geq 2$), we obtain

$$\langle 0|\, \tilde{\phi}(\hat{p})\, \mathrm{e}^{-i\mathbf{H}|t|}\, \tilde{\phi}(\hat{p}')]\, |0\rangle \underset{|t|\to\infty}{=} Z(2\pi)^{1-d} \frac{1}{2\omega(\hat{p})} \delta^{d-1}(\hat{p}+\hat{p}')\, \mathrm{e}^{-i\omega(\hat{p})|t|} + O(1/t).$$

Introducing $\varphi, \bar{\varphi}$, the properly normalized creation and annihilation operators of the one-particle states, we conclude that for large time the field $\tilde{\phi}(t,\hat{p})$ tends in a weak sense (not in an operator sense, but in all expectation values) towards

$$\tilde{\phi}(t,\hat{p}) \underset{|t|\to\infty}{\sim} \sqrt{Z} \frac{1}{2\omega(\hat{p})} \left[\varphi(-\hat{p}) + \bar{\varphi}(\hat{p})\right].$$

The constant $\sqrt{Z}$ is the field renormalization constant.

Normalized S-matrix elements. To calculate properly normalized S-matrix elements we can calculate with the action $\mathcal{A}(\phi\sqrt{Z})$. Alternatively, if we keep the initial field we have to renormalize the matrix elements.

6.10 S-matrix and Correlation Functions

This section somewhat anticipates concepts and results that are going to be presented in Chapters 7, 9 and 10, but otherwise contains material which has, more naturally, its place in this chapter.

6.10.1 Connected and 1PI correlation functions

Connected correlation functions. By comparing the general form of the S-matrix with correlation functions, we have seen in Section 6.8.2 that connected S-matrix elements are related to the analytic continuation to real time ($t \mapsto it, p_0 \mapsto ip_0$) of euclidean correlation functions in the the mass-shell limit. More precisely in terms of $\mathcal{W}(J)$, the generating functional of connected correlation functions (equation (7.57)),

$$\mathcal{T}(\phi_0) = i\mathcal{W}(\Delta^{-1}J),$$

where J is deduced from equation (6.85), ϕ_0 is given by equation (6.80), Δ is the propagator with a physical pole. Connected correlation functions can then be expressed in terms of amputated functions (inverting equation (7.79))

$$\tilde{W}^{(n)}(p_1,\ldots,p_n) = \left[\prod_{i=1}^{n} \tilde{W}^{(2)}(p_i)\right] \tilde{W}^{(n)}_{\text{amp.}}(p_1,\ldots,p_n). \tag{6.95}$$

The two-point function $\tilde{W}^{(2)}(p)$ has a pole located at $p^2 = -m^2$ where m is the physical mass. Near the pole

$$\tilde{W}^{(2)}(p) \underset{p^2\to -m^2}{\sim} \frac{Z}{p^2+m^2}, \tag{6.96}$$

where Z is the field renormalization constant (Section 6.9). We conclude that the coefficient $\mathcal{T}_{\mathrm{r}}^{(n)}$ of $\mathcal{T}(\phi_0/\sqrt{Z})$ in the expansion in powers of ϕ_0 is given by

$$\mathcal{T}_{\mathrm{r}}^{(n)}(p_1,\ldots,p_n) = iZ^{n/2}\, \tilde{W}^{(n)}_{\text{amp.}}(p_1,\ldots,p_n)\Big|_{p_i^2=-m^2}. \tag{6.97}$$

The factor Z in the equation corresponds to a finite renormalization of the field such that the residue of the two-point function (equation (6.96)) on the physical pole $p^2 = -m^2$ of the renormalized field is 1. It ensures that the matrix elements $\mathcal{T}_{\mathrm{r}}^{(n)}$ satisfy the unitarity relations with the proper normalization.

The generalization to several particles is straightforward.

S-matrix and 1PI generating functional. We now also exhibit a direct relation with the analytic continuation of the generating functional of proper vertices $\Gamma(\chi)$. We start from the euclidean form

$$\mathrm{e}^{\mathcal{W}(J)} = \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{A}(\phi)+J\phi} . \tag{6.98}$$

We then substitute $\Gamma(\chi)$, the Legendre transform of $\mathcal{W}(J)$:

$$\begin{aligned} \mathcal{W}(J) + \Gamma(\chi) &= \int \mathrm{d}t\, \mathrm{d}^d x\, J(t,x)\chi(t,x), \quad \chi(t,x) = \frac{\delta \mathcal{W}}{\delta J(t,x)}; \\ \mathrm{e}^{-\Gamma(\chi)+J\chi} &= \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{A}(\phi)+J\phi} . \end{aligned} \tag{6.99}$$

Using

$$J(t,x) = \frac{\delta \Gamma}{\delta \chi(t,x)},$$

we can write equation (6.99) as

$$\mathrm{e}^{-\Gamma(\chi)} = \int [\mathrm{d}\phi] \exp\left[-\mathcal{A}(\phi) + \int \mathrm{d}t\, \mathrm{d}^d x\, (\phi - \chi) \frac{\delta \Gamma}{\delta \chi(t,x)}\right], \tag{6.100}$$

or equivalently translating $\phi(t,x)$,

$$\mathrm{e}^{-\Gamma(\chi)} = \int [\mathrm{d}\phi] \exp\left[-\mathcal{A}(\phi+\chi) + \int \mathrm{d}t\, \mathrm{d}^d x\, \phi(t,x) \frac{\delta \Gamma}{\delta \chi(t,x)}\right]. \tag{6.101}$$

We now take the limit of a vanishing source J:

$$\frac{\delta \Gamma}{\delta \chi(t,x)} = 0 . \tag{6.102}$$

Of course, this 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 (6.80).

6.10.2 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 in local field transformations. Field correlation functions thus contain more information than the scattering matrix. This leads to problems in the point of view where 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 are calculated from connected correlation functions by taking the residues of the poles of the external propagator (see Sections 6.8.2, 6.10). We consider, for simplicity, the case of only one species of field ϕ, which we have defined 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 (equation (6.97)):

$$\mathcal{S}^{(n)}(p_1,\ldots,p_n) = Z^{n/2}\, W^{(n)}_{\text{amp.}}(p_1,\ldots,p_n)\left.\right|_{p^2=-m^2}\ , \tag{6.103}$$

where Z is the field renormalization constant.

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

$$\phi'(x) = C_1\phi(x) + \sum_2^\infty \frac{C_k}{k!}\phi^k(x), \qquad C_1 \neq 0\,. \tag{6.104}$$

We assume, when necessary, that the theory has been regularized in such a way that the new correlation functions exist (see Chapter 9).

Using relation (6.104) 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 position (see figure 6.1).

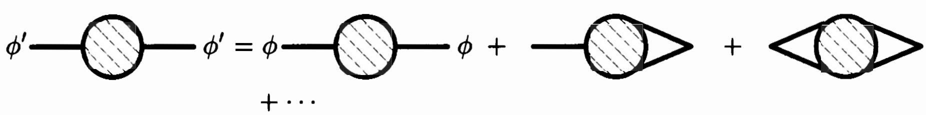


Fig. 6.1

The contributions to n-point functions which have poles on the external lines then have the form shown in figure 6.2.

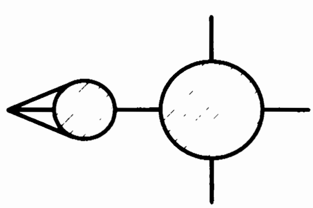


Fig. 6.2

In the mass shell limit ($p_i^2 = -m^2$), the ϕ' and ϕ correlation functions become proportional. The S-matrix elements are identical. Here, again, all fields related by transformation (6.96) are equivalent.

In such situations not all parameters of the theory are physical. For example, the field amplitude renormalization is obviously unphysical. 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 unphysical quantities (however, to prove renormalizability the study of correlation functions cannot be avoided, Chapter 10).

6.11 The Non-Relativistic Limit

Since a general discussion of the non-relativistic limit of quantum field theory would be somewhat involved, we consider here only one example that illustrates the main point. The low-energy limit of relativistic quantum field theory for massive particles is many-body quantum mechanics, and leads to a formalism naturally adapted to the statistical physics of non-relativistic quantum particles, as constructed in Sections 5.5, and 5.6.

The ϕ^4 interaction. We consider a massive scalar field theory with a ϕ^4 type interaction. To discuss the non-relativistic limit, it is convenient to employ the real time formalism. The real-time evolution operator is given by a functional integral of the form

$$U = \int [\mathrm{d}\phi] \exp i\mathcal{A}(\phi),$$
$$\mathcal{A}(\phi) = \int \mathrm{d}t\, \mathrm{d}x \left[\tfrac{1}{2}(\partial_t\phi)^2 - \tfrac{1}{2}\left(\nabla_x\phi\right)^2 - \tfrac{1}{2}m^2\phi^2 - \tfrac{1}{4!}g\phi^4\right]. \tag{6.105}$$

At least for a coupling weak enough, the integral is dominated by fields satisfying the free field equation

$$\left(\partial_t^2 - \nabla_x^2 + m^2\right)\phi(t,x) = 0\,.$$

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 field equation reduce simply to $\phi(t,x) \propto \mathrm{e}^{\pm imt}$. It is thus natural to introduce the holomorphic representation of fields, taking as the unperturbed harmonic oscillator $\mathcal{A}_0(\phi)$:

$$\mathcal{A}_0(\phi) = \int \mathrm{d}t\, \mathrm{d}x\, \left[\tfrac{1}{2}(\partial_t\phi)^2(t,x) - \tfrac{1}{2}m^2\phi^2(t,x)\right].$$

Denoting by $\varphi(t,x), \bar\varphi(t,x)$ the complex fields, in terms of which the field $\phi(t,x)$ reads

$$\phi(t,x) = (2m)^{-1/2}\left(\varphi(t,x) + \bar\varphi(t,x)\right),$$

we find the action

$$\mathcal{A}(\varphi,\bar\varphi) = \int \mathrm{d}t\, \mathrm{d}x \left[-i\bar\varphi\partial_t\varphi - m\varphi\bar\varphi - \frac{1}{4m}\left(\nabla_x\left(\varphi+\bar\varphi\right)\right)^2 - \frac{g}{96m^2}\left(\varphi+\bar\varphi\right)^4\right].$$

To separate the fast time frequencies, we then take new field variables

$$\varphi(t,x) \mapsto \mathrm{e}^{imt}\,\varphi(t,x), \quad \bar\varphi(t,x) \mapsto \mathrm{e}^{-imt}\,\bar\varphi(t,x),$$

where the new fields $\bar\varphi, \varphi$ have slow time variation compared to the factors e^{imt}. After this transformation the monomials of the form $\bar\varphi^r\varphi^s$ are multiplied by a factor $\mathrm{e}^{im(s-r)t}$. 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 5.6.2). Hence, at leading order, the only surviving terms are those that have an equal number of $\bar\varphi$ and φ factors. The non-relativistic action is then

$$\mathcal{A}(\bar\varphi,\varphi) = \int \mathrm{d}t\, \mathrm{d}x \left(-i\varphi\partial_t\bar\varphi - \frac{1}{2m}\nabla_x\varphi\nabla_x\bar\varphi - \frac{g}{16m^2}\varphi\varphi\bar\varphi\bar\varphi\right). \tag{6.106}$$

We recognize a real-time action written in terms of complex fields of the form (5.105). Therefore, the hamiltonian in the non-relativistic limit commutes with the particle number. This property, in general, is shared in relativistic quantum field theory only by free hamiltonians. In the non-relativistic limit of a massive theory, instead, 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

$$H_n = -\frac{1}{2m}\sum_{i=1}^{n} \nabla_{x_i}^2 + \frac{g}{8m^2}\sum_{i<j\le n} \delta(x_i - x_j).$$

This is an n-particle hamiltonian with two-body δ-function repulsive forces.

Finally, note that if we had expanded the fields with respect to the true physical mass, which is equal to m only for $g = 0$, we would have generated an additional chemical potential.

This analysis shows that the low-energy, non-relativistic limit of relativistic quantum field theory is many-body quantum mechanics.

Bibliographical Notes

Two important references for this chapter are

F.A. Berezin, *The Method of Second Quantization* (Academic Press, New York 1966);

L. D. Faddeev in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

Besides the work of C. Itzykson and J.B. Zuber, quoted in the introduction, another general reference here is

J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, (McGraw-Hill, New York 1964).

For a discussion of the eikonal approximation in field theory see, for example,

H.D.I. Abarbanel, *Cargèse Lectures in Physics* vol. 5, D. Bessis ed. (Gordon and Breach, New York 1972).

The non-relativistic limit is discussed in

L.S. Brown, *Quantum Field Theory*, (Cambridge University Press, Cambridge 1992).

The calculation of the S-matrix in terms of functional integrals with prescribed field asymptotic behaviour was proposed in

R.P. Feynman, *Acta Phys. Polon.* 24 (1963) 697; *Magic without Magic: J.A. Wheeler*, J. Klauder ed. (Freeman, San Francisco 1972).

The generating functional for the S-matrix was also discussed in

B. Zumino in *Theory and Phenomenology in Particle Physics*, Erice 1968, A. Zichichi ed. (Academic Press, New York 1968).

The interaction representation is discussed in

F.J. Dyson, *Phys. Rev.* 75 (1949) 486.

The perturbative equivalence of the operator formalism and gaussian integration is shown in

N.N. Bogoliubov and D.V. Shirkov, *Introduction to Theory of Quantized Fields* (Interscience, New York, London, 1959) chap. 7.

We recognize a real-time action written in terms of complex fields of the form (5.105). Therefore, the hamiltonian in the non-relativistic limit commutes with the particle number. This property, in general, is shared in relativistic quantum field theory only by free hamiltonians. In the non-relativistic limit of a massive theory, instead, 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

$$H_n = -\frac{1}{2m}\sum_{i=1}^{n}\nabla^2_{x_i} + \frac{g}{8m^2}\sum_{i<j\le n}\delta(x_i - x_j).$$

This is an n-particle hamiltonian with two-body δ-function repulsive forces.

Finally, note that if we had expanded the fields with respect to the true physical mass, which is equal to m only for $g = 0$, we would have generated an additional chemical potential.

This analysis shows that the low-energy, non-relativistic limit of relativistic quantum field theory is many-body quantum mechanics.

Bibliographical Notes

Two important references for this chapter are

F.A. Berezin, *The Method of Second Quantization* (Academic Press, New York 1966);
L. D. Faddeev in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

Besides the work of C. Itzykson and J.B. Zuber, quoted in the introduction, another general reference here is

J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, (McGraw-Hill, New York 1964).

For a discussion of the eikonal approximation in field theory see, for example,

H.D.I. Abarbanel, *Cargèse Lectures in Physics* vol. 5, D. Bessis ed. (Gordon and Breach, New York 1972).

The non-relativistic limit is discussed in

L.S. Brown, *Quantum Field Theory*, (Cambridge University Press, Cambridge 1992).

The calculation of the S-matrix in terms of functional integrals with prescribed field asymptotic behaviour was proposed in

R.P. Feynman, *Acta Phys. Polon.* 24 (1963) 697; *Magic without Magic: J.A. Wheeler*, J. Klauder ed. (Freeman, San Francisco 1972).

The generating functional for the S-matrix was also discussed in

B. Zumino in *Theory and Phenomenology in Particle Physics*, Erice 1968, A. Zichichi ed. (Academic Press, New York 1968).

The interaction representation is discussed in

F.J. Dyson, *Phys. Rev.* 75 (1949) 486.

The perturbative equivalence of the operator formalism and gaussian integration is shown in

N.N. Bogoliubov and D.V. Shirkov, *Introduction to Theory of Quantized Fields* (Interscience, New York, London, 1959) chap. 7.

APPENDIX A6

A6.1 Time-Ordered Products of Operators

We discuss this problem in the notation of simple quantum mechanics because the generalization to quantum field theory is just a matter of notation. To be consistent with the preceding sections we rewrite all expressions in the real-time formalism. We start from expression (2.39) after continuation to real time and want to directly recover a form analogous to the operator form (2.38). We also take immediately the thermodynamic limit. We define the n-point function $Z^{(n)}(t_1, t_2, \ldots, t_n)$ as

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \int [\mathrm{d}q]\, q(t_1) \ldots q(t_n)\, \mathrm{e}^{iA(q)/\hbar}\,. \tag{A6.1}$$

We then order times:

$$t_1 \le t_2 \le \cdots \le t_n\,. \tag{A6.2}$$

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

$$\mathcal{A}(q) = \sum_{i=1}^{n+1} \int_{t_{i-1}}^{t_i} \left[\tfrac{1}{2} m\dot{q}^2 - V(q)\right] \mathrm{d}t \quad \text{with} \quad t_0 = -\infty\,, \quad t_{n+1} = +\infty\,. \tag{A6.3}$$

We rewrite the path integral ($A6.1$) with the help of the identity

$$\prod_{i=1}^{n} q(t_i) = \int \prod_{i=1}^{n} \mathrm{d}q_i\, \delta\left[q(t_i) - q_i\right] q_i\,.$$

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 (2.18,2.19)), we see that the numerator in expression ($A6.1$) is exactly (recalling the ordering (2.4))

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

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

$$Q(t) = \mathrm{e}^{itH}\, \hat{q}\, \mathrm{e}^{-itH}, \tag{A6.4}$$

we can write

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \langle 0|\, Q(t_n) \ldots Q(t_1)\, |0\rangle\,. \tag{A6.5}$$

The order of the operators in the r.h.s. reflects the time-ordering ($A6.2$).

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

$$\mathrm{T}\left[A_1(t_1) A_2(t_2)\right] = A_1(t_1) A_2(t_2) \theta(t_1 - t_2) + A_2(t_2) A_1(t_1) \theta(t_2 - t_1).$$

We can then rewrite expression ($A6.5$), irrespective now of the order between the times $t_1, \ldots, t_n$,

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \langle 0| \mathrm{T}\left[Q(t_1)Q(t_2)\ldots Q(t_n)\right]|0\rangle . \tag{A6.6}$$

We have expressed the path integral in terms of 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, for instance, calculate scattering amplitudes (see Section 6.8.2).

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

$$Z^{(n)}(t_1, t_2, \ldots, t_n) = \mathcal{Z}^{-1}(\beta)\,\mathrm{tr}\left\{\mathrm{e}^{-\beta H}\,\mathrm{T}\left[Q(t_1)Q(t_2)\ldots Q(t_n)\right]\right\} . \tag{A6.7}$$

A6.2 Perturbation Theory in the Operator Formalism

For completeness and to illustrate the differences with the path integral formulation, let us here recall the basis of perturbation theory in the operator formalism of quantum mechanics. To calculate the S-matrix, for example, we need an expression for the operator $\Omega(t)$ (see equation (6.3)):

$$\Omega(t) = \mathrm{e}^{iH_0 t}\,\mathrm{e}^{-iHt}, \tag{A6.8}$$

in which H_0 is the unperturbed hamiltonian and

$$V = H - H_0$$

the perturbation.

The operator $\Omega(t)$ satisfies the equation

$$\dot{\Omega}(t) = -iV_{\mathrm{I}}(t)\Omega(t), \tag{A6.9}$$

with the boundary condition

$$\Omega(0) = \mathbf{1} .$$

The operator $V_{\mathrm{I}}(t)$ is the perturbation in the interaction representation:

$$V_{\mathrm{I}}(t) = \mathrm{e}^{iH_0 t}\,V\,\mathrm{e}^{-iH_0 t} . \tag{A6.10}$$

One verifies that the solution of equation ($A6.9$) can be formally written as

$$\Omega(t) = \sum_{n=0}^{\infty}(-i)^n \int \mathrm{d}t_1\,\mathrm{d}t_2 \ldots \mathrm{d}t_n V_{\mathrm{I}}(t_n)V_{\mathrm{I}}(t_{n-1})\ldots V_{\mathrm{I}}(t_2)V_{\mathrm{I}}(t_1), \tag{A6.11}$$

the domain of integration in the r.h.s. being

$$0 \le t_1 \le t_2 \le \cdots \le t_n \le t .$$

If we now replace the product of perturbation terms which appears in the l.h.s. by the time-ordered product, as defined in Section A6.1, the product becomes symmetric in the time arguments. We can, therefore, symmetrize the domain of integration provided we divide by a symmetry factor $n!$:

$$\Omega(t) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{\substack{0 \le t_i \le t \\ 1 \le i < n}} \mathrm{d}t_1 \, \mathrm{d}t_2 \ldots \mathrm{d}t_n \mathrm{T} \left[V_{\mathrm{I}}(t_1) V_{\mathrm{I}}(t_2) \ldots V_{\mathrm{I}}(t_n) \right]. \qquad (A6.12)$$

This expression can be formally rewritten as

$$\Omega(t) = \mathrm{T} \left[\exp \left(-i \int_0^t \mathrm{d}t' \, V_{\mathrm{I}}(t') \right) \right]. \qquad (A6.13)$$

We can, in particular, 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 ($A6.12$), we recover the relation between correlation functions and T-products established in Section 2.4.

QUANTUM FIELD THEORY: FUNCTIONAL METHODS AND PERTURBATION THEORY

In this chapter, we proceed with our study of *local, relativistic* quantum field theory.

We first discuss the neutral self-coupled scalar field $\phi(x)$ introduced in Section 6.6.1. An important example is provided by the so-called ϕ^4 theory, which as we argue in Section 6.11, has the theory of spinless bosons interacting through pair potentials, described in Section 5.5.4, as a non-relativistic limit.

We construct a representation of the quantum statistical operator in the form of a functional integral, relativistic extension of the functional integrals introduced in Section 5.5.2. We then mainly investigate the simpler zero temperature limit where all d coordinates, euclidean time and space, can be treated symmetrically.

The functional integral defines a functional measure to which correspond field correlation functions. These functions are analytic continuations to imaginary (euclidean) time of the vacuum expectation values of field operators. As we have already explained in Sections 2.4.1, 2.4.2 in the case of the quantum particle, they have a more direct interpretation as correlation functions of models of classical statistical physics, in some continuum limit, 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.

As in quantum mechanics, 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 functional 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. This perturbative expansion provides an alternative algebraic definition of the functional integral. We show, because this is no longer obvious, that this definition is consistent with the usual manipulations performed on integrals like integration by parts and change of variables. We exhibit several algebraic properties of functional integrals and derive the Dyson–Schwinger field equations, which are the quantum field equations expressed in terms of correlation functions. We also define the functional δ-function.

We then introduce and discuss two new generating functionals, $\mathcal{W}(J)$ the generating functional of connected correlation functions (analogous to the free energy of statistical physics), and its Legendre transform $\Gamma(\varphi)$, the generating functional of proper vertices (also sometimes called effective potential and analogous to the thermodynamic potential of statistical physics). Connected correlation functions, as a consequence of locality, have cluster properties. To proper vertices contribute only one-line irreducible Feynman diagrams (diagrams which cannot be disconnected by cutting only one line). These diagrams are also one-particle irreducible (1PI) in the framework of Particle Physics. We, therefore, often call proper vertices 1PI correlation functions.

We explain how to calculate these various functionals in a reorganized perturbative expansion, called loop expansion.

The functional $\Gamma(\varphi)$ plays an important role in the renormalization of local quantum field theory. In Section 7.10 we give a quantum interpretation for $\Gamma(\varphi)$. We also relate it to the partition function at fixed field time average. This relation explains that $\Gamma(\varphi)$

QUANTUM FIELD THEORY: FUNCTIONAL METHODS AND PERTURBATION THEORY

In this chapter, we proceed with our study of *local*, *relativistic* quantum field theory.

We first discuss the neutral self-coupled scalar field $\phi(x)$ introduced in Section 6.6.1. An important example is provided by the so-called ϕ^4 theory, which as we argue in Section 6.11, has the theory of spinless bosons interacting through pair potentials, described in Section 5.5.4, as a non-relativistic limit.

We construct a representation of the quantum statistical operator in the form of a functional integral, relativistic extension of the functional integrals introduced in Section 5.5.2. We then mainly investigate the simpler zero temperature limit where all d coordinates, euclidean time and space, can be treated symmetrically.

The functional integral defines a functional measure to which correspond field correlation functions. These functions are analytic continuations to imaginary (euclidean) time of the vacuum expectation values of field operators. As we have already explained in Sections 2.4.1, 2.4.2 in the case of the quantum particle, they have a more direct interpretation as correlation functions of models of classical statistical physics, in some continuum limit, 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.

As in quantum mechanics, 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 functional 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. This perturbative expansion provides an alternative algebraic definition of the functional integral. We show, because this is no longer obvious, that this definition is consistent with the usual manipulations performed on integrals like integration by parts and change of variables. We exhibit several algebraic properties of functional integrals and derive the Dyson–Schwinger field equations, which are the quantum field equations expressed in terms of correlation functions. We also define the functional δ-function.

We then introduce and discuss two new generating functionals, $\mathcal{W}(J)$ the generating functional of connected correlation functions (analogous to the free energy of statistical physics), and its Legendre transform $\Gamma(\varphi)$, the generating functional of proper vertices (also sometimes called effective potential and analogous to the thermodynamic potential of statistical physics). Connected correlation functions, as a consequence of locality, have cluster properties. To proper vertices contribute only one-line irreducible Feynman diagrams (diagrams which cannot be disconnected by cutting only one line). These diagrams are also one-particle irreducible (1PI) in the framework of Particle Physics. We, therefore, often call proper vertices 1PI correlation functions.

We explain how to calculate these various functionals in a reorganized perturbative expansion, called loop expansion.

The functional $\Gamma(\varphi)$ plays an important role in the renormalization of local quantum field theory. In Section 7.10 we give a quantum interpretation for $\Gamma(\varphi)$. We also relate it to the partition function at fixed field time average. This relation explains that $\Gamma(\varphi)$

also appears in the discussion of symmetry breaking.

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.

7.1 Functional Integrals. Correlation Functions

The neutral scalar field has been introduced in Section 6.6.1. We again consider the classical lagrangian density (6.51) (everywhere the speed of light is set to one):

$$\mathcal{L}(\phi) = \tfrac{1}{2}\dot{\phi}^2(t,x) - \tfrac{1}{2}\left(\nabla\phi(t,x)\right)^2 - V\left(\phi(t,x)\right), \tag{7.1}$$

for a neutral scalar field $\phi(t,x)$, where $V(\phi)$ is a function expandable in powers of ϕ, a polynomial in the simplest examples like

$$V(\phi) = \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{4!}g\phi^4 .$$

The lagrangian (7.1) is *local* because it depends only on the field $\phi(t,x)$ and its partial derivatives, invariant under space and time translations since space and time do not appear explicitly, relativistic invariant, that is, invariant under the pseudo-orthogonal group $O(1,d-1)$ acting linearly on t and $\mathbf{x}$.

The corresponding hamiltonian density is then

$$\mathcal{H}(\pi,\phi) = \tfrac{1}{2}\pi^2(x) + \tfrac{1}{2}\left[\nabla\phi(x)\right]^2 + V\left(\phi(x)\right), \tag{7.2}$$

and the total hamiltonian

$$\mathbf{H} = \int \mathrm{d}^{d-1}x\, \mathcal{H}\left[\pi(x),\phi(x)\right]. \tag{7.3}$$

The coordinates q_i of quantum mechanics are here replaced by the field $\phi(x)$. The transition from quantum mechanics to field theory can be understood in much the same way as the transition between the discretized action (2.17) and the continuum time limit (2.18).

Since this hamiltonian is quadratic in the momentum $\pi(x)$ it is a continuum generalization of hamiltonians of the form $\sum_i p_i^2 + V(q)$. Following the method explained in Section 2.2, we thus immediately obtain a functional integral representation for the matrix elements of the statistical operator $U(t_2,t_1) = \mathrm{e}^{-(t_2-t_1)\mathbf{H}/\hbar}$:

$$\langle\phi_2|\,U(t_2,t_1)\,|\phi_1\rangle = \int \left[\mathrm{d}\phi(t,x)\right]\exp\left[-\mathcal{S}(\phi)/\hbar\right] \tag{7.4}$$

with the boundary conditions $\phi(t_1,\mathbf{x}) = \phi_1(\mathbf{x})$, $\phi(t_2,\mathbf{x}) = \phi_2(\mathbf{x})$.

As we know from the analysis of Section 2.2, $\mathcal{S}(\phi)$ is the classical euclidean action, obtained by analytic continuation to euclidean (imaginary) time from the action in real time (6.58),

$$\mathcal{S}(\phi) = \int_{t_1}^{t_2}\mathrm{d}t \int \mathrm{d}^{d-1}x \left\{\tfrac{1}{2}\left[\dot{\phi}^2(t,x) + \left(\nabla_x\phi(t,x)\right)^2\right] + V\left(\phi(t,x)\right)\right\}. \tag{7.5}$$

One advantage of the euclidean time formulation, obvious in expression (7.5), 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.

7.1.1 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 integration over all fields with periodic boundary conditions:

$$\mathcal{Z}(\beta) = \operatorname{tr} \mathrm{e}^{-\beta \mathbf{H}} = \int_{\phi(\beta,\mathbf{x})=\phi(0,\mathbf{x})} [\mathrm{d}\phi(t,x)] \exp\left[-\mathcal{S}(\phi)\right] \tag{7.6}$$

with

$$\mathcal{S}(\phi) = \int_0^\beta \mathrm{d}t \int \mathrm{d}^{d-1}x \left\{ \tfrac{1}{2}\left[\dot{\phi}^2(t,x)/\hbar^2 + \left(\nabla_x \phi(t,x)\right)^2\right] + V\left(\phi(t,x)\right)\right\}, \tag{7.7}$$

an expression that one can compare with the non-relativistic form (5.94).

From now on we will in general set $\hbar = 1$.

Correlation functions. The integrand of the functional integral (7.6) defines a positive measure for a classical statistical field theory (see Section 2.4). To this measure correspond correlation functions:

$$\langle \phi(t_1,x_1)\ldots\phi(t_n,x_n)\rangle = \mathcal{Z}^{-1}(\beta)\int [\mathrm{d}\phi]\phi(t_1,x_1)\ldots\phi(t_n,x_n)\,\mathrm{e}^{-\mathcal{S}(\phi)}, \tag{7.8}$$

in which $\mathcal{Z}(\beta)$ is the quantum partition function (7.6). Equal-time correlation functions are also the static correlation functions of the finite temperature quantum field theory (see Chapter 38). 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 large β limit these finite temperature correlation functions become the ground state (usually called *vacuum* in quantum field theory) expectation values of time-ordered products of quantum field operators (see Appendix A6.1). After analytic continuation towards real time, they yield the Green's functions from which elements of the scattering S-matrix can be calculated (Section 6.8.2).

Using the property

$$\frac{\delta}{\delta J(t_1,x_1)} \exp\left[\int \mathrm{d}t\,\mathrm{d}^{d-1}x\, J(t,x)\phi(t,x)\right] = \phi(t_1,x_1)\exp\left[\int \mathrm{d}t\,\mathrm{d}^{d-1}x\, J(t,x)\phi(t,x)\right], \tag{7.9}$$

where $J(t,x)$ is an external field or source, we verify that ϕ-field correlation functions can be obtained from a *generating functional* $\mathcal{Z}(J)$,

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}t\,\mathrm{d}^{d-1}x\, J(t,x)\phi(t,x)\right], \tag{7.10}$$

by functional differentiation:

$$\langle \phi(t_1,x_1)\ldots\phi(t_n,x_n)\rangle = \mathcal{Z}^{-1}(J=0)\left[\frac{\delta}{\delta J(t_1,x_1)}\cdots\frac{\delta}{\delta J(t_n,x_n)}\mathcal{Z}(J)\right]\Bigg|_{J\equiv 0}. \tag{7.11}$$

Zero temperature. Since finite temperature quantum field theory will be discussed extensively in Chapter 38, we work for simplicity in the remaining part of this chapter directly at zero temperature in infinite $\mathbb{R}^{d-1}$ space, and, therefore, in infinite $\mathbb{R}^d$ euclidean space: correlation functions, therefore, correspond to quantum vacuum expectation values. The euclidean action is then fully $O(d)$ invariant and, therefore, we no longer distinguish space and time; $\mathbf{x}$ belongs to $\mathbb{R}^d$.

7.1.1 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 integration over all fields with periodic boundary conditions:

$$\mathcal{Z}(\beta) = \operatorname{tr} e^{-\beta \mathbf{H}} = \int_{\phi(\beta,\mathbf{x})=\phi(0,\mathbf{x})} [d\phi(t,x)] \exp[-\mathcal{S}(\phi)] \tag{7.6}$$

with

$$\mathcal{S}(\phi) = \int_0^\beta dt \int d^{d-1}x \left\{ \tfrac{1}{2} \left[\dot{\phi}^2(t,x)/\hbar^2 + \left(\nabla_x \phi(t,x)\right)^2 \right] + V\left(\phi(t,x)\right) \right\}, \tag{7.7}$$

an expression that one can compare with the non-relativistic form (5.94).

From now on we will in general set $\hbar = 1$.

Correlation functions. The integrand of the functional integral (7.6) defines a positive measure for a classical statistical field theory (see Section 2.4). To this measure correspond correlation functions:

$$\langle \phi(t_1,x_1) \ldots \phi(t_n,x_n) \rangle = \mathcal{Z}^{-1}(\beta) \int [d\phi] \phi(t_1,x_1) \ldots \phi(t_n,x_n)\, e^{-\mathcal{S}(\phi)}, \tag{7.8}$$

in which $\mathcal{Z}(\beta)$ is the quantum partition function (7.6). Equal-time correlation functions are also the static correlation functions of the finite temperature quantum field theory (see Chapter 38). 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 large β limit these finite temperature correlation functions become the ground state (usually called *vacuum* in quantum field theory) expectation values of time-ordered products of quantum field operators (see Appendix A6.1). After analytic continuation towards real time, they yield the Green's functions from which elements of the scattering S-matrix can be calculated (Section 6.8.2).

Using the property

$$\frac{\delta}{\delta J(t_1,x_1)} \exp\left[\int dt\, d^{d-1}x\, J(t,x)\phi(t,x)\right] = \phi(t_1,x_1) \exp\left[\int dt\, d^{d-1}x\, J(t,x)\phi(t,x)\right], \tag{7.9}$$

where $J(t,x)$ is an external field or source, we verify that ϕ-field correlation functions can be obtained from a *generating functional* $\mathcal{Z}(J)$,

$$\mathcal{Z}(J) = \int [d\phi] \exp\left[-\mathcal{S}(\phi) + \int dt\, d^{d-1}x\, J(t,x)\phi(t,x)\right], \tag{7.10}$$

by functional differentiation:

$$\langle \phi(t_1,x_1) \ldots \phi(t_n,x_n) \rangle = \mathcal{Z}^{-1}(J=0) \left[\frac{\delta}{\delta J(t_1,x_1)} \cdots \frac{\delta}{\delta J(t_n,x_n)} \mathcal{Z}(J) \right] \bigg|_{J\equiv 0}. \tag{7.11}$$

Zero temperature. Since finite temperature quantum field theory will be discussed extensively in Chapter 38, we work for simplicity in the remaining part of this chapter directly at zero temperature in infinite $\mathbb{R}^{d-1}$ space, and, therefore, in infinite $\mathbb{R}^d$ euclidean space: correlation functions, therefore, correspond to quantum vacuum expectation values. The euclidean action is then fully $O(d)$ invariant and, therefore, we no longer distinguish space and time; $\mathbf{x}$ belongs to $\mathbb{R}^d$.

7.1.2 Lattice regularization. Classical statistical physics

We have already seen that path integrals reduced to their formal definition in continuous time are sometimes ill defined and that in such cases it is necessary to complete the formal definition with a limiting procedure based on a time discretized version. We shall eventually find out that this problem is even more severe in field theories. This will be the subject of many chapters. Let us here note only that we may define the functional integral (7.4) as the formal limit of an integral in which both time and space are discretized (see Section 9.6 for details). We introduce a d-dimensional hypercubic lattice, with lattice spacing a, and use as dynamical variables the values of the field on all lattice sites. We replace the field derivatives $\partial_\mu \phi$ by

$$\partial_\mu \phi \mapsto \Delta_\mu \phi = \frac{1}{a}\left[\phi(\mathbf{r}+a\mathbf{n}_\mu) - \phi(\mathbf{r})\right], \tag{7.12}$$

in which μ and $\mathbf{r}$ refer *both to space and time* and $\mathbf{n}_\mu$ is the unit vector in the μ direction. The regularized euclidean action $\mathcal{S}_a(\phi)$ is then:

$$\mathcal{S}_a(\phi) = a^d \sum_{\mathbf{r}} \left\{ \tfrac{1}{2}\left[\Delta_\mu \phi(\mathbf{r})\right]^2 + V\left(\phi(\mathbf{r})\right)\right\}. \tag{7.13}$$

The discretized action can be considered as the configuration energy of a lattice model in classical statistical physics. The quantum partition function (7.6) corresponds to a lattice of finite size β in the euclidean time direction, and periodic boundary conditions. The infinite $\beta = t_2 - t_1$ (zero temperature) limit corresponds to the thermodynamic (infinite volume) of the classical model.

Correlation functions of lattice variables have also as a formal limit the continuum ϕ-field correlation functions.

In expression (7.13) it is clear again that the potential and the gradient squared (kinetic) term play different roles: the potential term $V(\phi)$ is ultra-local, it weights the functional integral according to the values of the field at each point independently; the gradient squared term instead suppresses the fields that are too singular in the small lattice spacing limit, $a \to 0$, those for which the quantity

$$\left|\phi(\mathbf{r}+a\mathbf{n}_\mu) - \phi(\mathbf{r})\right|^2 a^{d-2}$$

diverges.

Note that this condition becomes weaker when the dimension increases. It implies continuity only for $d < 2$, that is, in non-relativistic quantum mechanics. This feature has deep consequences. As we shall discuss later, for $d \geq 2$ the continuum limit does not exist in general. One situation, which we study extensively and which leads to a continuum theory, is the following: the regularized theory, considered as a statistical model, has a continuous phase transition. In the limit in which one parameter of the theory, such that m^2 the second derivative of $V(\phi)$ at its minimum, approaches some special value, corresponding to the critical temperature of the statistical model, a continuum limit can be defined.

The problem of the continuum limit will lead us to establish relations between renormalization theory explained in Chapters 9–10 and the statistical theory of continuous phase transitions which will be discussed in Chapters 23–37 devoted to critical phenomena.

7.1.2 Lattice regularization. Classical statistical physics

We have already seen that path integrals reduced to their formal definition in continuous time are sometimes ill defined and that in such cases it is necessary to complete the formal definition with a limiting procedure based on a time discretized version. We shall eventually find out that this problem is even more severe in field theories. This will be the subject of many chapters. Let us here note only that we may define the functional integral (7.4) as the formal limit of an integral in which both time and space are discretized (see Section 9.6 for details). We introduce a d-dimensional hypercubic lattice, with lattice spacing a, and use as dynamical variables the values of the field on all lattice sites. We replace the field derivatives $\partial_\mu \phi$ by

$$\partial_\mu \phi \mapsto \Delta_\mu \phi = \frac{1}{a} \left[\phi(\mathbf{r} + a\mathbf{n}_\mu) - \phi(\mathbf{r})\right], \tag{7.12}$$

in which μ and $\mathbf{r}$ refer *both to space and time* and $\mathbf{n}_\mu$ is the unit vector in the μ direction. The regularized euclidean action $\mathcal{S}_a(\phi)$ is then:

$$\mathcal{S}_a(\phi) = a^d \sum_{\mathbf{r}} \left\{ \tfrac{1}{2} \left[\Delta_\mu \phi(\mathbf{r})\right]^2 + V\left(\phi(\mathbf{r})\right) \right\}. \tag{7.13}$$

The discretized action can be considered as the configuration energy of a lattice model in classical statistical physics. The quantum partition function (7.6) corresponds to a lattice of finite size β in the euclidean time direction, and periodic boundary conditions. The infinite $\beta = t_2 - t_1$ (zero temperature) limit corresponds to the thermodynamic (infinite volume) of the classical model.

Correlation functions of lattice variables have also as a formal limit the continuum ϕ-field correlation functions.

In expression (7.13) it is clear again that the potential and the gradient squared (kinetic) term play different roles: the potential term $V(\phi)$ is ultra-local, it weights the functional integral according to the values of the field at each point independently; the gradient squared term instead suppresses the fields that are too singular in the small lattice spacing limit, $a \to 0$, those for which the quantity

$$\left|\phi(\mathbf{r} + a\mathbf{n}_\mu) - \phi(\mathbf{r})\right|^2 a^{d-2}$$

diverges.

Note that this condition becomes weaker when the dimension increases. It implies continuity only for $d < 2$, that is, in non-relativistic quantum mechanics. This feature has deep consequences. As we shall discuss later, for $d \geq 2$ the continuum limit does not exist in general. One situation, which we study extensively and which leads to a continuum theory, is the following: the regularized theory, considered as a statistical model, has a continuous phase transition. In the limit in which one parameter of the theory, such that m^2 the second derivative of $V(\phi)$ at its minimum, approaches some special value, corresponding to the critical temperature of the statistical model, a continuum limit can be defined.

The problem of the continuum limit will lead us to establish relations between renormalization theory explained in Chapters 9–10 and the statistical theory of continuous phase transitions which will be discussed in Chapters 23–37 devoted to critical phenomena.

7.2 Perturbative Expansion. Wick's Theorem and Feynman Diagrams

In Section 2.5.1 we have shown how to calculate a path integral for a hamiltonian of the form $\frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2 + V_{\rm I}(q)$ as a series expansion in powers of $V_{\rm I}(q)$, for any function $V_{\rm I}(q)$ expandable in powers of q. The result was based on the calculation of a reference gaussian integral (in Chapter 2, the harmonic oscillator). Here, we apply exactly the same strategy to functional integrals.

Furthermore, although most results derived in this chapter will be illustrated with examples corresponding to an action of the form (7.5), the results are more general, applying to any bosonic field theory, and this explains the abstract notation used below.

7.2.1 The gaussian integral

We consider a general gaussian functional integral:

$$\mathcal{Z}_{\rm G}(J) = \int [{\rm d}\phi] \exp\left[-\tfrac{1}{2}\phi K \phi + J \cdot \phi\right], \tag{7.14}$$

where we have used the symbolic notation:

$$\phi K \phi \equiv \int {\rm d}^d x\, {\rm d}^d y\, \phi(x) K(x,y) \phi(y), \qquad J \cdot \phi \equiv \int {\rm d}^d x\, J(x)\phi(x).$$

We assume that the kernel K is symmetric and positive. In expression (7.14), a normalization is implied; we choose $\mathcal{Z}_{\rm G}(0) = 1$.

We denote the inverse of K by Δ:

$$\int {\rm d}^d z\, \Delta(x,z) K(z,y) = \delta^d(x-y). \tag{7.15}$$

To calculate the functional integral (7.14), we simply shift ϕ by ΔJ and find after integration

$$\mathcal{Z}_{\rm G}(J) = \exp\left(\tfrac{1}{2} J \Delta J\right), \tag{7.16}$$

again with the convention

$$J \Delta J = \int {\rm d}^d x\, {\rm d}^d y\, J(x) \Delta(x,y) J(y).$$

The identities derived in Section 7.1.1 show that $\mathcal{Z}_{\rm G}(J)$ is also the generating functional of correlation functions corresponding to a general action quadratic in the field ϕ. The inverse kernel Δ in equation (7.15) is thus the two-point function of the gaussian or *free field theory*:

$$\langle \phi(x_1)\phi(x_2)\rangle_0 = \left.\frac{\delta^2 \mathcal{Z}_{\rm G}(J)}{\delta J(x_1)\delta J(x_2)}\right|_{J\equiv 0} = \Delta(x_1, x_2), \tag{7.17}$$

where $\langle \bullet \rangle_0$ means gaussian or free field expectation value. In the example of a free massive theory with mass m, which corresponds in expression (7.5) to the choice

$$V(\phi) = \tfrac{1}{2} m^2 \phi^2,$$

the kernel K is a local operator:

$$K(\mathbf{x},\mathbf{y}) = \left(-\nabla^2 + m^2\right) \delta^d(\mathbf{x}-\mathbf{y}), \tag{7.18}$$

where we denote by ∇ the gradient operator in d dimensions. The kernel Δ then takes the form

$$\Delta(\mathbf{x},\mathbf{y}) = \frac{1}{(2\pi)^d} \int {\rm d}^d p\, \frac{{\rm e}^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{p^2 + m^2}. \tag{7.19}$$

7.2.2 Perturbation theory. Wick's theorem

We now consider a more general euclidean action of the form

$$\mathcal{S}(\phi) = \tfrac{1}{2}\phi K\phi + \mathcal{V}_{\mathrm{I}}(\phi). \tag{7.20}$$

Using the property (7.9),

$$\frac{\delta}{\delta J(x)}\,\mathrm{e}^{J\cdot\phi} = \phi(x)\,\mathrm{e}^{J\cdot\phi},$$

we can express the functional integral

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi]\exp\left[-\mathcal{S}(\phi) + J\cdot\phi\right], \tag{7.21}$$

in terms of $\mathcal{Z}_{\mathrm{G}}(J)$ as

$$\mathcal{Z}(J) = \exp\left[-\mathcal{V}_{\mathrm{I}}\left(\frac{\delta}{\delta J}\right)\right]\mathcal{Z}_{\mathrm{G}}(J) = \exp\left[-\mathcal{V}_{\mathrm{I}}\left(\frac{\delta}{\delta J}\right)\right]\exp\left(\tfrac{1}{2}J\Delta J\right), \tag{7.22}$$

an expression analogous to (2.62). Note that in this chapter we use the convention that a differential operator like $\delta/\delta J$ acts on all factors placed to its right.

The identities (7.22,7.11) can then be combined to calculate all ϕ-field correlation functions as a formal series in powers of the interaction potential $\mathcal{V}_{\mathrm{I}}(\phi)$.

Wick's theorem. Perturbation theory, that is, an expansion in powers of $\mathcal{V}_{\mathrm{I}}$, reduces all calculations to gaussian averages of product of fields. From the expression (7.22), and using the arguments of Section 1.1, we obtain a straightforward generalization of equations (1.9–1.14) or (2.50), which expresses Wick's theorem in field theory:

$$\begin{aligned}\left\langle\prod_1^{2s}\phi(z_i)\right\rangle_0 &= \left[\prod_{i=1}^{2s}\frac{\delta}{\delta J(z_i)}\exp\left(\tfrac{1}{2}J\Delta J\right)\right]\Bigg|_{J\equiv 0}\\ &= \sum_{\substack{\text{all possible pairings}\\ \text{of }\{1,2,\ldots,2s\}}}\Delta(z_{i_1}z_{i_2})\ldots\Delta(z_{i_{2s-1}},z_{i_{2s}}),\end{aligned} \tag{7.23}$$

where $\langle\bullet\rangle_0$ means gaussian or free field average. Perturbation theory involves a basic ingredient: the two-point function Δ of the gaussian theory (equation (7.17)) which we call the *propagator*.

Graphically, each term in the sum can be represented by a set of contractions corresponding to the particular pairing chosen. For example, for $s=2$ one finds

$$\begin{aligned}\langle\phi(z_1)\phi(z_2)\phi(z_3)\phi(z_4)\rangle_0 &= \overbrace{\phi(z_1)\phi(z_2)}\overbrace{\phi(z_3)\phi(z_4)} + \ 2 \text{ terms}\\ &= \Delta(z_1,z_2)\Delta(z_3,z_4) + \Delta(z_1,z_3)\Delta(z_2,z_4)\\ &\quad + \Delta(z_1,z_4)\Delta(z_2,z_3).\end{aligned}$$

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

$$\left\langle\phi(x_1)\cdots\phi(x_n)\int\mathrm{d}^d y_1\,\phi^{p_1}(y_1)\int\mathrm{d}^d y_2\,\phi^{p_2}(y_2)\cdots\int\mathrm{d}^d y_k\,\phi^{p_k}(y_k)\right\rangle_0 .$$

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: 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.

Remark. It will be verified in the coming chapters that interacting theories with a propagator of the form (7.19) have large momentum or short distance divergences. Therefore, in what follows we assume either that the field theory has been regularized by replacing continuum space by a lattice as in Section 7.1.2, or the propagator Δ has been replaced by a more complicated function that ensures the convergence of all terms in the perturbative expansion (see Sections 9.5–9.6).

7.2.3 The ϕ^4 example

We now illustrate the previous discussion with the important quartic example,

$$\mathcal{V}_{\mathrm{I}}(\phi) = \frac{g}{4!}\int \mathrm{d}^d x\, \phi^4(x).$$

The two-point function. The two-point function to order g^2 has the expansion

$$\langle \phi(x_1)\phi(x_2)\rangle = (\mathrm{a}) - \frac{1}{2}g\,(\mathrm{b}) + \frac{g^2}{4}\,(\mathrm{c}) + \frac{g^2}{4}\,(\mathrm{d}) + \frac{g^2}{6}\,(\mathrm{e}) + O\left(g^3\right).$$

Note that three additional contributions which factorize into

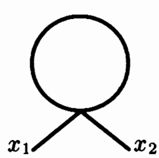
$$\langle \phi(x_1)\phi(x_2)\rangle_0 \left\langle \phi^4(y)\right\rangle_0 , \quad \langle \phi(x_1)\phi(x_2)\phi^4(y_1)\rangle_0 \left\langle \phi^4(y_2)\right\rangle_0 \text{ and}$$
$$\langle \phi(x_1)\phi(x_2)\rangle_0 \left\langle \phi^4(y_1)\phi^4(y_2)\right\rangle_0 ,$$

cancel in the division by $\mathcal{Z}(J=0)$.

Then (a) is the propagator:

(a): x_1 x_2

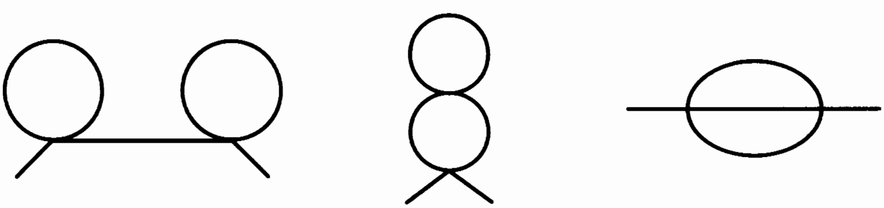
(b), the Feynman diagram which appears at order g, is displayed in figure 7.1:



(b)

Fig. 7.1 Two-point function at order g.

and the diagrams (c),(d),(e) are displayed in figure 7.2:



(c) (d) (e)

Fig. 7.2 Contributions of order g^2 to the two-point function.

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

$$\frac{g^2}{2!\,(4!)^2}\int \mathrm{d}^d y_1 \int \mathrm{d}^d y_2 \left\langle \phi(x_1)\phi(x_2)\phi^4(y_1)\phi^4(y_2)\right\rangle_0 ,$$

and 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. The three remaining fields of the first interaction term can finally 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

$$\mathcal{S}(\phi) = \frac{1}{2}\phi^2 + \frac{g}{4!}\phi^4 ,$$

and the two-point function is given by

$$Z^{(2)} = \frac{\int \mathrm{d}\phi\, \phi^2 \exp\left[-\mathcal{S}(\phi)\right]}{\int \mathrm{d}\phi\, \exp\left[-\mathcal{S}(\phi)\right]} = 1 - \frac{g}{2} + \frac{2}{3}g^2 + O\left(g^3\right),$$

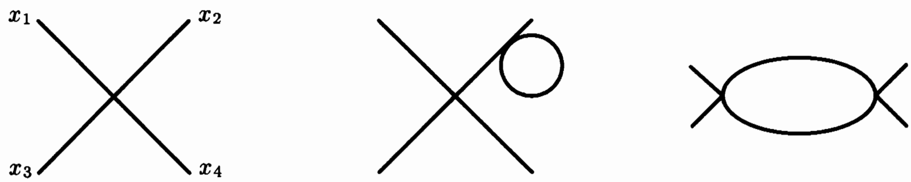
in which the expressions correspond to ordinary one-variable integrals.

The sum rules are satisfied. For example, at order g^2

$$\frac{2}{3} = \frac{1}{4} + \frac{1}{4} + \frac{1}{6}\,.$$

The four-point function. The expansion of the four-point function to order g^2 is

$$\begin{aligned}\langle\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle &= \left[\,(\mathrm{a})_{12}\ (\mathrm{a})_{34}\ + 2\,\mathrm{terms}\right] - \frac{g}{2}\left[\,(\mathrm{a})_{12}\ (\mathrm{b})_{34}\right.\\ &\left. +5\,\mathrm{terms}\right] - g\,(\mathrm{f}) + g^2\left\{(\mathrm{a})_{12}\left[\frac{1}{4}\left((\mathrm{c})_{34} + (\mathrm{d})_{34}\right) + \frac{1}{6}(\mathrm{e})_{34}\right] + 5\ \mathrm{terms}\right\}\\ &+ \frac{g^2}{4}\left[\,(\mathrm{b})_{12}\ (\mathrm{b})_{34}\ + 2\,\mathrm{terms}\right] + \frac{g^2}{2}\left[(\mathrm{g}) + 3\ \mathrm{terms}\right] + \frac{g^2}{2}\left[(\mathrm{h}) + 2\ \mathrm{terms}\right] + O\left(g^3\right).\end{aligned}$$



(f) (g) (h)

Fig. 7.3 New Feynman diagrams contributing to the four-point function.

The new diagrams (f), (g), (h) are displayed in figure 7.3. The notation for example $(\mathrm{a})_{12}$ means the 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.

Note that the graphs which involve the two-point functions 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. The question will be discussed further starting with Section 7.4.

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 7.4 will confirm, these diagrams are cancelled by the perturbative expansion of $\mathcal{Z}(J=0)$ in expression (7.11). The diagrams contributing to $\mathcal{Z}(0)$ (the partition function) are called vacuum diagrams.

A final remark: local interaction terms may also involve derivatives of the field $\phi(x)$. Then in expression (7.23) derivatives of the propagator appear. The representation in terms of the Feynman diagrams given above is no longer faithful since it does not indicate where the derivatives are. A more faithful representation can be obtained by splitting points at the vertices and putting arrows on lines.

7.3 Algebraic Properties of Functional Integrals. Field Equations

Functional integrals: perturbative definition. We have seen that path integrals sometimes are ill defined, and a limiting procedure has to be supplied to lift ambiguities. We will discover in Chapter 9 that the problem is even more severe in local quantum field theories for functional integrals. We have exhibited in Section 7.1.2 a possible regularization by a space time lattice. However, this regularization complicates perturbative calculations enormously, and in Chapter 9 more practical regularizations will be proposed. This will lead to a consistency problem: some regularization schemes like dimensional regularization have no interpretation beyond perturbation theory. It is then no longer obvious whether identities proven using standard properties of functional integrals remain valid after regularization. To avoid these difficulties, one can take expression (7.22) as a proper, though perturbative, definition of the functional integral. In particular, one may then avoid the reference to a discretized space and time. However, with this new definition, one has to prove that the usual transformations performed on standard integrals lead to identities that are also true with such a perturbative definition.

The proofs rely on simple but powerful techniques which simultaneously allow a derivation of important identities satisfied by correlation functions like Dyson–Schwinger equations, a form of the quantum field equations when expressed in terms of correlation functions. Finally, by repeating the formal manipulations involved in these derivations, the reader should become more familiar with the algebraic properties of the perturbative expansion.

We, therefore, now assume that the functional integral (7.21) with action (7.20),

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\tfrac{1}{2}\phi K \phi - \mathcal{V}_{\mathrm{I}}(\phi) + J\cdot\phi\right], \tag{7.24}$$

is really defined by expression (7.22):

$$\mathcal{Z}(J) = \exp\left[-\mathcal{V}_{\mathrm{I}}\left(\frac{\delta}{\delta J}\right)\right] \exp\left(\tfrac{1}{2}J\Delta J\right).$$

We also assume, of course, that all terms in the expansion of this expression in powers of $\mathcal{V}_{\mathrm{I}}$ exist and are finite (which, as we shall see later, implies some conditions for the kernel K and the interaction term $\mathcal{V}_{\mathrm{I}}(\phi)$. From the functional integral representation (7.21) we now heuristically derive identities satisfied by the generating functional $\mathcal{Z}(J)$ and then prove algebraically that these identities also follow from the definition (7.22).

7.3.1 Integration by parts and quantum field equations

The integral of a total derivative vanishes. This simple property can be used to derive heuristically an equation for $\mathcal{Z}(J)$:

$$\int [\mathrm{d}\phi] \frac{\delta}{\delta\phi(x)} \exp\left(-\mathcal{S} + J\cdot\phi\right) = 0\,, \tag{7.25}$$

and, therefore,

$$\int [\mathrm{d}\phi] \left[J(x) - \frac{\delta\mathcal{S}}{\delta\phi(x)}\right] \exp\left(-\mathcal{S} + J\cdot\phi\right) = 0\,. \tag{7.26}$$

As a consequence of identity (7.9) this equation can be transformed into

$$\left[\frac{\delta\mathcal{S}}{\delta\phi(x)}\left(\frac{\delta}{\delta J}\right) - J(x)\right] \mathcal{Z}(J) = 0\,. \tag{7.27}$$

Equation (7.27) is a compact form of quantum field or Dyson–Schwinger equations. 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 now show, determine correlation functions completely.

Solution of the quantum field equations. Conversely the solution of equation (7.27) is the functional integral (7.21). To prove this statement, we write $\mathcal{Z}(J)$, the solution of equation (7.27), as a generalized Laplace transform:

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\tilde{\mathcal{S}}(\phi) + J\cdot\phi\right]. \tag{7.28}$$

Equation (7.27) implies

$$\int [\mathrm{d}\phi] \exp\left[-\tilde{\mathcal{S}}(\phi) + J\cdot\phi\right] \left[\frac{\delta\mathcal{S}(\phi)}{\delta\phi(x)} - J(x)\right] = 0\,. \tag{7.29}$$

Using equation (7.26) but with $\mathcal{S}$ replaced by $\tilde{\mathcal{S}}$ we transform this equation into

$$\int [\mathrm{d}\phi] \exp\left[-\tilde{\mathcal{S}}(\phi) + J\cdot\phi\right] \frac{\delta}{\delta\phi(x)} \left[\mathcal{S}(\phi) - \tilde{\mathcal{S}}(\phi)\right] = 0\,. \tag{7.30}$$

Therefore,

$$\mathcal{S}(\phi) - \tilde{\mathcal{S}}(\phi) = \text{const.}. \tag{7.31}$$

The constant is fixed by the normalization of $\mathcal{Z}(J)$.

Example. We again consider the euclidean action

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left\{\frac{1}{2}\left[\partial_\mu \phi(x)\right]^2 + \frac{1}{2}m^2\phi^2(x) + \frac{g}{4!}\phi^4(x)\right\}.$$

Equation (7.27), in this example, reads

$$\left[\left(-\nabla_x^2 + m^2\right)\frac{\delta}{\delta J(x)} + \frac{g}{3!}\left(\frac{\delta}{\delta J(x)}\right)^3 - J(x)\right]\mathcal{Z}(J) = 0\,.$$

Let us write the two first equations obtained by expanding in powers of $J(x)$. Differentiating once with respect to $J(y)$ before setting J to zero we find

$$\left(-\nabla_x^2 + m^2\right)\langle\phi(x)\phi(y)\rangle + \frac{g}{3!}\left\langle\phi(x)^3\phi(y)\right\rangle = \delta^d(x-y).$$

Differentiating three times we obtain

$$\left(-\nabla_x^2 + m^2\right)\langle\phi(x)\phi(y_1)\phi(y_2)\phi(y_3)\rangle + \frac{g}{3!}\left\langle\phi(x)^3\phi(y_1)\phi(y_2)\phi(y_3)\right\rangle = \delta^d(x-y_1)$$
$$\times\langle\phi(y_2)\phi(y_3)\rangle + \ 2\ \text{terms}\,.$$

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 try solving these equations in a non-perturbative way, but this requires truncating, in some way, the infinite set.

7.3.2 Direct algebraic proof of the quantum field equations

We now show with purely algebraic transformations that $\mathcal{Z}(J)$ defined by equation (7.22) indeed satisfies equation (7.27) that, for the explicit form (7.20) of $\mathcal{S}(\phi)$, reads

$$\left[\int \mathrm{d}y\, K(x,y)\frac{\delta}{\delta J(y)} + \frac{\delta\mathcal{V}_{\mathrm{I}}}{\delta\phi(x)}\left(\frac{\delta}{\delta J}\right) - J(x)\right]\mathcal{Z}(J) = 0\,. \tag{7.32}$$

It is convenient here to introduce the notation D_J for $\delta/\delta J$ when it appears as an argument:

$$D_J \equiv \delta/\delta J\,.$$

We want thus to prove algebraically

$$\left[\int \mathrm{d}y\, K(x,y)\frac{\delta}{\delta J(y)} + \frac{\delta\mathcal{V}_{\mathrm{I}}(D_J)}{\delta\phi(x)} - J(x)\right]\exp\left[-\mathcal{V}_{\mathrm{I}}(D_J)\right]\exp\left(\tfrac{1}{2}J\Delta J\right) = 0\,. \tag{7.33}$$

We first observe that the identity is true in the gaussian theory since, as a consequence of the definition of Δ,

$$\int \mathrm{d}y\, K(x,y)\Delta(y,z) = \delta(x-z),$$

we find

$$\left[\int \mathrm{d}y\, K(x,y)\frac{\delta}{\delta J(y)} - J(x)\right] \exp\left(\tfrac{1}{2}J\Delta J\right) = 0\,. \tag{7.34}$$

We then act with the operator $\exp(-\mathcal{V}_{\mathrm{I}}(D_J))$ on the left of equation (7.34).

(i) The operator commutes with KD_J.

(ii) We use the commutation relation

$$[F(D_J), J(x)] = \frac{\delta F(D_J)}{\delta\phi(x)}, \tag{7.35}$$

which can easily be verified by expanding F in powers of $\delta/\delta J$. Applied to the functional

$$F(\phi) = \exp\left[-\mathcal{V}_{\mathrm{I}}(\phi)\right],$$

the commutation relation implies

$$\exp\bigl(-\mathcal{V}_{\mathrm{I}}(D_J)\bigr)J(x) = J(x)\exp\bigl(-\mathcal{V}_{\mathrm{I}}(D_J)\bigr) - \frac{\delta\mathcal{V}_{\mathrm{I}}(D_J)}{\delta\phi(x)}\exp\bigl(-\mathcal{V}_{\mathrm{I}}(D_J)\bigr). \tag{7.36}$$

This completes the proof.

7.3.3 The infinitesimal change of variables

Various identities satisfied by correlation functions in the case of field theories possessing symmetries can be proven by the method of infinitesimal change of variables.

We thus change variables $\phi(x) \mapsto \chi(x)$ in a functional integral, setting

$$\phi(x) = \chi(x) + \varepsilon F(x;\chi), \tag{7.37}$$

in which ε is an infinitesimal parameter and $F(x;\chi)$ a general functional of χ:

$$F(x;\chi) = \sum_1^{\infty}\frac{1}{n!}\int \mathrm{d}y_1\ldots\mathrm{d}y_n\,\chi(y_1)\ldots\chi(y_n)F^{(n)}(x;y_1,\ldots,y_n). \tag{7.38}$$

The variation of the action $\mathcal{S}(\phi)$ is

$$\mathcal{S}(\phi) = \mathcal{S}(\chi) + \varepsilon\int \mathrm{d}x\frac{\delta\mathcal{S}}{\delta\chi(x)}F(x;\chi) + O\left(\varepsilon^2\right). \tag{7.39}$$

The change of variables (7.37) in the functional integral generates a jacobian $\mathcal{J}$:

$$\mathcal{J} = \det\frac{\delta\phi(x)}{\delta\chi(y)} = \det\left[\delta(x-y) + \varepsilon\frac{\delta F(x;\chi)}{\delta\chi(y)}\right]. \tag{7.40}$$

As a consequence of identity (1.101),

$$\det\left(1+\varepsilon M\right) = 1 + \varepsilon\,\mathrm{tr}\,M + O\left(\varepsilon^2\right),$$

we obtain

$$\mathcal{J} = 1 + \varepsilon \int \mathrm{d}x \frac{\delta F(x;\chi)}{\delta \chi(x)} + O\left(\varepsilon^2\right). \tag{7.41}$$

It follows that

$$\begin{aligned} \mathcal{Z}(J) = \int [\mathrm{d}\chi] \left(1 + \varepsilon \int \mathrm{d}x \frac{\delta F(x;\chi)}{\delta \chi(x)}\right) \bigg[1 - \varepsilon \int \mathrm{d}x \frac{\delta \mathcal{S}}{\delta \chi(x)} F(x;\chi) \\ + \varepsilon \int \mathrm{d}x\, J(x) F(x;\chi) \bigg] \exp\left[-\mathcal{S}(\chi) + J\cdot\chi\right] + O\left(\varepsilon^2\right). \end{aligned} \tag{7.42}$$

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, we obtain the identity

$$\int \mathrm{d}x \left[F(x;D_J) \frac{\delta \mathcal{S}(D_J)}{\delta \chi(x)} - \frac{\delta F(x;D_J)}{\delta \chi(x)} - J(x) F(x;D_J) \right] \mathcal{Z}(J) = 0\,. \tag{7.43}$$

Algebraic proof. The algebraic proof of the identity relies on acting with the differential operator $\int \mathrm{d}x\, F(x;\delta/\delta J)$ on the field equation (7.27):

$$\left[\frac{\delta \mathcal{S}(D_J)}{\delta \phi(x)} - J(x)\right] \mathcal{Z}(J) = 0\,.$$

Equation (7.43) then follows immediately from the commutation relation (7.35) used in the form

$$\left[F(y;D_J), J(x)\right] = \frac{\delta F(y;D_J)}{\delta \chi(x)}.$$

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.

7.3.4 The choice of the gaussian measure

We now want to show that, provided the sums of some geometric series exist, a part of the quadratic term K can be treated as an interaction.

We thus decompose the kernel K in expression (7.20) into a sum of two terms:

$$K = K_1 + K_2 \qquad \text{with } K_1 > 0\,.$$

We want to prove algebraically that

$$\begin{aligned} &\exp\left[-\mathcal{V}_{\mathrm{I}}(D_J) - \tfrac{1}{2} D_J K_2 D_J\right] \exp\left(\tfrac{1}{2} J K_1^{-1} J\right) \\ &\quad = \mathcal{N}(K_1, K_2) \exp\left[-\mathcal{V}_{\mathrm{I}}(D_J)\right] \exp\left[\tfrac{1}{2} J \left(K_1 + K_2\right)^{-1} J\right], \end{aligned} \tag{7.44}$$

where $\mathcal{N}$ is independent of the source J.

Since the operator $\exp\left[-\mathcal{V}_{\mathrm{I}}(\delta/\delta J)\right]$ can be factorized in both sides of equation (7.44), it is sufficient to prove the identity for $\mathcal{V}_{\mathrm{I}} = 0$, that is, to calculate the functional $\mathcal{Z}_{\mathrm{G}}(J)$:

$$\mathcal{Z}_{\mathrm{G}}(J) = \exp\left(-\tfrac{1}{2} D_J K_2 D_J\right) \exp\left(\tfrac{1}{2} J K_1^{-1} J\right). \tag{7.45}$$

We act with $K_1 D_J$ on $\mathcal{Z}_G(J)$:

$$\int \mathrm{d}y\, K_1(x,y) \frac{\delta}{\delta J(y)} \mathcal{Z}_G(J) = \exp\left(-\tfrac{1}{2} D_J K_2 D_J\right) J(x) \exp\left(\tfrac{1}{2} J K_1^{-1} J\right). \qquad (7.46)$$

In the r.h.s. we then commute J to bring it to the left (equation (7.35)):

$$\exp\left(-\tfrac{1}{2} D_J K_2 D_J\right) J(x) = J(x) \exp\left(-\tfrac{1}{2} D_J K_2 D_J\right) - (K_2 D_J)(x) \exp\left(-\tfrac{1}{2} D_J K_2 D_J\right).$$

We conclude that $\mathcal{Z}_G$ satisfies the equation

$$\left[\int \mathrm{d}y\, (K_1 + K_2)(x,y) \frac{\delta}{\delta J(y)} - J(x)\right] \mathcal{Z}_G(J) = 0. \qquad (7.47)$$

Integrating the equation we thus find

$$\mathcal{Z}_G(J) = \mathcal{N}(K_1, K_2) \exp\left[\tfrac{1}{2} J (K_1 + K_2)^{-1} J\right],$$

proving identity (7.44). 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

$$\exp\left[\int \mathrm{d}x\, L(x) \frac{\delta}{\delta J(x)}\right] F(J) = F(J + L).$$

Of course all these identities make sense only if both sides exist.

7.3.5 The functional Dirac δ-function

We consider, here, a field $\boldsymbol{\phi}(x)$ with N components $\phi_i(x)$, $i = 1, \ldots, N$, satisfying a constraint:

$$F(x, \boldsymbol{\phi}) = 0. \qquad (7.48)$$

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

$$\delta[F] \equiv \int [\mathrm{d}\lambda(x)] \exp\left[\int \mathrm{d}x\, \lambda(x) F(x, \boldsymbol{\phi})\right], \qquad (7.49)$$

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 (7.49) has the properties expected from a δ-function. We consider a functional integral in which the integration is restricted to fields $\boldsymbol{\phi}$ satisfying the constraint $F(x, \boldsymbol{\phi}) = 0$:

$$\mathcal{Z}_F(\mathbf{J}) = \int \prod_{i=1}^{N} [\mathrm{d}\phi_i(x)]\, \delta(F) \exp\left(-\mathcal{S}(\boldsymbol{\phi}) + \mathbf{J} \cdot \boldsymbol{\phi}\right). \qquad (7.50)$$

We use the representation (7.49) to write $\mathcal{Z}_F$ as

$$\mathcal{Z}_F(\mathbf{J}) = \int [\mathrm{d}\lambda] \prod_{i=1}^{N} [\mathrm{d}\phi_i] \exp\left[-\mathcal{S}(\phi) + \lambda \cdot F(\phi) + \mathbf{J} \cdot \phi\right]. \tag{7.51}$$

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 (7.22) to define, algebraically, integral (7.51) (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 $\mathcal{S}(\phi)$,

$$\mathcal{S}(\phi) \mapsto \mathcal{S}(\phi) + \int \mu(x) F(x, \phi) \mathrm{d}x\,, \tag{7.52}$$

we can cancel this change by a 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

$$\int \prod_i [\mathrm{d}\phi_i] [\mathrm{d}\lambda] \frac{\delta}{\delta\lambda(x)} \exp\left[\lambda \cdot F + \mathbf{J} \cdot \phi - \mathcal{S}(\phi)\right] = 0\,,$$

we derive

$$F(x; \frac{\delta}{\delta J}) \mathcal{Z}_F(\mathbf{J}) = 0\,. \tag{7.53}$$

This equation expresses the constraint $F = 0$ on the generating functional of correlation functions $\mathcal{Z}(J)$.

Examples. In Section 14.9 a functional δ-function in the form (7.49) will be 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 18, one covariant gauge will correspond to constraining the gauge field A_μ to satisfy the condition $\partial_\mu A_\mu = 0$. Again the representation (7.49) will be useful. The equation (7.53) then becomes $\partial_\mu^x \delta\mathcal{Z}/\delta J_\mu(x) = 0$, where J_μ is the gauge field source.

7.4 Connected Correlation Functions. Cluster Properties

In Section 7.1.1, we have introduced the generating functional of correlation functions $\mathcal{Z}(J)$, where $J(x)$ is a classical field (also called source). We will now introduce another generating functional, the generating functional of connected correlation functions.

For what follows, it is convenient to change the normalization of $\mathcal{Z}(J)$ such that $\mathcal{Z}(J = 0) = 1$. Then,

$$\mathcal{Z}(J) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \mathrm{d}x_1 \ldots \mathrm{d}x_n\, Z^{(n)}(x_1, \ldots, x_n) J(x_1) \ldots J(x_n), \tag{7.54}$$

$$Z^{(n)}(x_1, \ldots, x_n) \equiv \langle \phi(x_1) \ldots \phi(x_n) \rangle = \left[\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{Z}(J)\right]\Bigg|_{J=0}. \tag{7.55}$$

The functional $\mathcal{Z}(J)$ is proportional to the functional integral (7.10):

$$\mathcal{Z}(J) = \mathcal{N} \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, J(x)\phi(x)\right] \tag{7.56}$$

with

$$\frac{1}{\mathcal{N}} = \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{S}(\phi)} .$$

Perturbative expansion: connected Feynman diagrams. In Section 7.2.3 we have noticed that some Feynman diagrams are disconnected in the sense of graphs. They can then be factorized into a product 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.

We now define a new generating functional $\mathcal{W}(J)$, analogous to the free energy of statistical physics:

$$\mathcal{W}(J) = \ln \mathcal{Z}(J). \tag{7.57}$$

We introduce the notation

$$\mathcal{W}(J) = \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}x_1 \dots \mathrm{d}x_n\, W^{(n)}(x_1, x_2, \dots, x_n) J(x_1) J(x_2) \dots J(x_n). \tag{7.58}$$

Then,

$$W^{(n)}(x_1, \dots, x_n) = \left[\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \mathcal{W}(J)\right]\Bigg|_{J=0} . \tag{7.59}$$

In Section 1.2.1 we have already given a combinatorial proof that only connected Feynman diagrams contribute to $\mathcal{W}(J)$. This implies in particular that vacuum connected diagrams can only contribute to $\mathcal{W}(J = 0)$. They do not contribute to correlation functions.

We give, below, independent derivations which emphasize other related properties.

Let us first write explicitly a few relations, as implied by the definition (7.57), between correlation functions and the functions $W^{(n)}$:

$$\begin{aligned}
Z^{(1)}(x) &= W^{(1)}(x), \\
Z^{(2)}(x_1, x_2) &= W^{(2)}(x_1, x_2) + W^{(1)}(x_1) W^{(1)}(x_2), \\
Z^{(3)}(x_1, x_2, x_3) &= W^{(3)}(x_1, x_2, x_3) + W^{(1)}(x_1) W^{(2)}(x_2, x_3) + W^{(1)}(x_2) W^{(2)}(x_3, x_1) \\
&\quad + W^{(1)}(x_3) W^{(2)}(x_1, x_2) + W^{(1)}(x_1) W^{(1)}(x_2) W^{(1)}(x_3), \\
&\dots
\end{aligned}$$

Note that the r.h.s. involves all possible products of connected functions with coefficient one.

7.4.1 A proof of connectivity

We now give an algebraic proof of connectivity that relies, in particular, on linearity: a linear combination of connected functions is still connected.

In Section 7.2 we have calculated the generating functional $\mathcal{Z}_{\mathrm{G}}(J)$ for a gaussian theory. We have found (equation (7.16)) that

$$\mathcal{Z}_{\mathrm{G}}(J) = \exp\left[\frac{1}{2} \int \mathrm{d}^d x\, \mathrm{d}^d y\, J(x) \Delta(x, y) J(y)\right], \tag{7.60}$$

where $\Delta(x,y)$ is the propagator. Expanding in powers of J, we immediately verify that, except the two-point correlation function, all correlation functions are disconnected. The functional $\mathcal{W}_{\mathrm{G}}(J)$ is instead

$$\mathcal{W}_{\mathrm{G}}(J) = \ln \mathcal{Z}_{\mathrm{G}}(J) = \frac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}^d y\, J(x)\Delta(x-y)J(y).$$

All $W^{(n)}$ correlation functions vanish, except the two-point function which is connected.

We now assume that for some action $\mathcal{S}(\phi)$ we have shown that $\mathcal{W}(J)$ generates connected correlation functions and we add a small local perturbation to the action

$$\mathcal{S}_\varepsilon(\phi) = \mathcal{S}(\phi) + \varepsilon \int \mathrm{d}x\, \phi^N(x),$$

(derivative couplings leave the argument unchanged). Then,

$$\exp \mathcal{W}_\varepsilon(J) = \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{S}_\varepsilon(\phi)+J\phi} = \exp\left[-\varepsilon \int \mathrm{d}x \left(\frac{\delta}{\delta J(x)}\right)^N\right] \exp \mathcal{W}(J).$$

It follows that

$$\mathcal{W}_\varepsilon(J) = \mathcal{W}(J) - \varepsilon\, \mathrm{e}^{-\mathcal{W}(J)} \int \mathrm{d}x \left(\frac{\delta}{\delta J(x)}\right)^N \mathrm{e}^{\mathcal{W}(J)} + O\left(\varepsilon^2\right).$$

Examining the contribution of order ε we see that it 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 $N=3$

$$\mathcal{W}_\varepsilon(J) - \mathcal{W}(J) = -\varepsilon \int \mathrm{d}x \left[\frac{\delta^3 \mathcal{W}(J)}{[\delta J(x)]^3} + 3\frac{\delta^2 \mathcal{W}(J)}{[\delta J(x)]^2}\frac{\delta \mathcal{W}(J)}{\delta J(x)} + \left(\frac{\delta \mathcal{W}(J)}{\delta J(x)}\right)^3\right] + O\left(\varepsilon^2\right).$$

The corresponding diagrams are connected diagrams contributing to $\mathcal{W}(J)$ attached to the point x. Therefore, if $\mathcal{W}(J)$ is connected, all additional terms are also connected.

Since $\mathcal{W}(J)$ is connected for a general gaussian theory, it follows, after integration over the corresponding coupling constant, that it remains connected for any interaction.

The functions $W^{(n)}$ are called connected correlation functions. Sometimes we will emphasize this character by using the notation

$$W^{(n)}(x_1,\ldots,x_n) = \langle \phi(x_1)\ldots\phi(x_n)\rangle_{\mathrm{c}}\,, \tag{7.61}$$

where the symbol $< \bullet >_{\mathrm{c}}$ means connected contribution to the corresponding correlation function.

7.4.2 Connected correlation functions and cluster properties

We now consider *local* euclidean actions $\mathcal{S}(\phi)$, for simplicity, functions only of a scalar field $\phi(x)$, and invariant under space translations:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x\, \mathcal{L}(\phi; x). \tag{7.62}$$

This means that 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 (7.1) after continuation to euclidean time).

We have already observed in quantum mechanics that the partition function $\operatorname{tr} e^{-\beta H}$ in the large volume limit, that is, for $\beta \to \infty$, has an exponential behaviour; more precisely (equation (2.2)),

$$\lim_{\beta\to\infty} -\frac{1}{\beta} \ln \operatorname{tr} e^{-\beta H} = E_0 ,$$

in which E_0 is the ground state energy. Moreover, the convergence towards the limit is exponential when the ground state is isolated (equation (2.3)).

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. Peculiarities found in massless theories or due to ground state degeneracy will be ignored here and discussed only later.

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:

$$J(x) = J_1(x) + J_2(x), \quad \begin{cases} J_1(x) = 0 & \text{for } x \notin \Omega_1 , \\ J_2(x) = 0 & \text{for } x \notin \Omega_2 , \end{cases} \quad \Omega_1 \cap \Omega_2 = \emptyset . \tag{7.63}$$

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 that we can write

$$\begin{aligned} \mathcal{S}(\phi) - \int \mathrm{d}x\, J\phi = {} & \int_{x\in\Omega_1} \mathrm{d}x \left[\mathcal{L}(\phi;x) - J_1(x)\phi(x)\right] + \int_{x\in\Omega_2} \mathrm{d}x \left[\mathcal{L}(\phi;x) - J_2(x)\phi(x)\right] \\ & + \int_{x\notin\Omega_1\cup\Omega_2} \mathrm{d}x\, \mathcal{L}(\phi;x) + \text{contributions from boundaries}. \end{aligned} \tag{7.64}$$

We then write $\mathcal{Z}(J)$ as

$$\mathcal{Z}(J) = \mathcal{Z}_1(J_1)\mathcal{Z}_2(J_2)\mathcal{Z}_{12}(J_1,J_2) \tag{7.65}$$

with the definitions

$$\begin{aligned} \mathcal{Z}_1(J_1) &= \int_{x\in\Omega_1} [\mathrm{d}\phi(x)] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, J_1(x)\phi(x)\right], \\ \mathcal{Z}_2(J_2) &= \int_{x\in\Omega_2} [\mathrm{d}\phi(x)] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, J_2(x)\phi(x)\right]. \end{aligned}$$

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 (7.65). Its dependence in J_1 and J_2 comes entirely from the existence of boundary terms in equation (7.64). When we scale up Ω_1 and Ω_2, these boundary terms grow like surfaces while the first two terms in equation (7.64) 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 $V_2 \to \infty$.

To express this property, it is natural to introduce the functional $\mathcal{W}(J) = \ln \mathcal{Z}(J)$ (equation (7.57)) which then satisfies

$$\mathcal{W}(J_1 + J_2) = \mathcal{W}_1(J_1) + \mathcal{W}_2(J_2) + \text{negligible}. \tag{7.66}$$

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 (2.2).

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:

$$\mathcal{W}(J_1 + J_2) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{p=0}^{n} \binom{n}{p} \int \mathrm{d}x_1 \ldots \mathrm{d}x_p \, \mathrm{d}y_{p+1} \ldots \mathrm{d}y_n W^{(n)}(x_1, \ldots, x_p, y_{p+1}, \ldots, y_n) \\ \times J_1(x_1) \ldots J_1(x_p) J_2(y_{p+1}) \ldots J_2(y_n), \tag{7.67}$$

with $x_i \in \Omega_1$, $y_j \in \Omega_2$.

Property (7.66) implies that all terms with $p \neq 0$ or $p \neq n$ are negligible for V_1 and V_2 large. Considering expression (7.67) we see that this implies that the functions $W^{(n)}$ must decrease rapidly enough when the two non-empty sets of points $\{x_1 \ldots x_p\}$ and $\{y_{p+1} \ldots y_n\}$ have large separations:

$$W^{(n)}(x_1 \ldots x_p, y_{p+1} \ldots y_n) \to 0 \quad \text{when} \quad \min_{\substack{i=1\cdots p \\ j=p+1\cdots n}} |x_i - y_j| \to \infty \,. \tag{7.68}$$

This property, which we have here described only qualitatively, is called the *cluster property,* and is a characteristic property of the connected correlation functions generated by the functional $\mathcal{W}(J)$.

Feynman diagrams. We have seen that a Feynman diagram which is disconnected in the sense of graphs can be factorized into a product of the form

$$F_1(x_1, \ldots, x_p) F_2(y_1, \ldots, y_q).$$

It is clear that we can separate the set $\{x_i\}$ and the set $\{y_i\}$ in a translation invariant theory, 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 that is the inverse of the smallest mass in the theory. This property is a consequence of the exponential decrease of the propagator (see Appendix A7.3 for details).

Remark. The relation between action and generating functional, which has the form of a Laplace transformation, can formally be inverted:

$$\mathrm{e}^{-\mathcal{S}(\phi)} = \int [\mathrm{d}J] \exp\left[\mathcal{W}(J) - \int \mathrm{d}x \, J(x)\phi(x)\right], \tag{7.69}$$

where one integrates over imaginary sources $J(x)$. A truncated loopwise expansion (see Section 7.7) of the functional integral then yields approximate non-linear equations for correlation functions. It is actually convenient to introduce in the r.h.s. the generating functional of proper vertices defined in the next section.

7.5 Legendre Transformation. Proper Vertices

We now introduce a new generating functional, the Legendre transform of the generating functional of connected correlation functions $\mathcal{W}(J)$, called the generating functional of proper vertices.

Legendre transformation. We shall see later that in the context of statistical physics and phase transitions, it is natural to consider the thermodynamic potential, Legendre transform of the *free energy* $\mathcal{W}(J)$. Here, proper vertices are introduced because they have special properties from the point of view of perturbation theory, as we demonstrate in Sections 7.7 and 7.8. The generating functional of proper vertices $\Gamma(\varphi)$, in which $\varphi(x)$ is a classical field argument of Γ, is obtained from $\mathcal{W}(J)$ by

$$\Gamma(\varphi) + \mathcal{W}(J) - \int \mathrm{d}x\, J(x)\varphi(x) = 0\,, \tag{7.70}$$

$\varphi(x)$ being related to $J(x)$ by

$$\varphi(x) = \frac{\delta \mathcal{W}}{\delta J(x)}. \tag{7.71}$$

We have discussed the Legendre transformation in Section 1.8 and shown that it is involutive. In particular,

$$J(x) = \frac{\delta \Gamma}{\delta \varphi(x)}. \tag{7.72}$$

Moreover, if $\mathcal{W}(J)$ depends on some parameter v, then,

$$\frac{\partial \mathcal{W}}{\partial v} + \frac{\partial \Gamma}{\partial v} = 0\,, \tag{7.73}$$

a result that will be quite useful. Note that we have derived this result for one external variable but it obviously applies also for an external field or source.

Remark. If we set $J = 0$ in equation (7.71) 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 (7.72) then implies that the expectation value of $\phi(x)$ is an extremum of $\Gamma(\varphi)$.

Expansion of $\Gamma(\varphi)$. Inverting the relation (7.71) we can expand the source $J(x)$ as a series of powers of $\xi(x)$:

$$\xi(x) = \varphi(x) - W^{(1)}(x) = \varphi(x) - \langle \phi(x) \rangle\,. \tag{7.74}$$

We see that $\xi(x)$ is related to the correlation functions of the field:

$$\Xi(x) = \phi(x) - \langle \phi(x) \rangle\,, \tag{7.75}$$

which has vanishing expectation value, $\langle \Xi(x) \rangle = 0$.

The first terms of the expansion of equation (7.71) then are

$$\xi(x) = \int \mathrm{d}x_1\, W^{(2)}(x, x_1) J(x_1) + \frac{1}{2!} \int \mathrm{d}x_1\, \mathrm{d}x_2\, W^{(3)}(x, x_1, x_2) J(x_1) J(x_2) + \cdots\,. \tag{7.76}$$

We introduce the inverse $S(x,y)$ of the connected two-point function:

$$\int \mathrm{d}z\, S(x,z) W^{(2)}(z,y) = \delta(x-y). \tag{7.77}$$

We can then write the expansion of $J(x)$:

$$J(x) = \int \mathrm{d}x_1\, S(x,x_1)\xi(x_1) - \frac{1}{2!}\int \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}y_3\, \mathrm{d}x_1\, \mathrm{d}x_2\, S(x,y_3) \\ \times S(x_1,y_1) S(x_2,y_2) W^{(3)}(y_1,y_2,y_3)\xi(x_1)\xi(x_2) + \cdots\,. \tag{7.78}$$

Note that this expansion can be naturally expressed in terms of so-called amputated correlation functions $W^{(n)}_{\text{amp.}}$:

$$W^{(n)}_{\text{amp.}}(x_1,\ldots,x_n) = \int \left[\prod_{i=1}^{n} \mathrm{d}y_i\, S(x_i,y_i)\right] W^{(n)}(y_1,\ldots,y_n). \tag{7.79}$$

In terms of Feynman diagrams this means that propagators and contributions to the two-point function on external lines are omitted.

We can now use equation (7.72) to calculate $\Gamma(\varphi)$. Setting

$$\Gamma(\varphi) = \sum_{1}^{\infty} \frac{1}{n!} \int \mathrm{d}x_1 \ldots \mathrm{d}x_n\, \Gamma^{(n)}(x_1,\ldots,x_n)\xi(x_1)\ldots\xi(x_n), \tag{7.80}$$

where the $\Gamma^{(n)}$ are associated with the field Ξ (defined by (7.75)), we find

$$\begin{aligned}
\Gamma^{(1)}(x) &= 0\,,\\
\Gamma^{(2)}(x_1,x_2) &= S(x_1,x_2) = [W^{(2)}]^{-1}(x_1,x_2),\\
\Gamma^{(3)}(x_1,x_2,x_3) &= -W^{(3)}_{\text{amp.}}(x_1,x_2,x_3),\\
\Gamma^{(4)}(x_1,x_2,x_3,x_4) &= -W^{(4)}_{\text{amp.}}(x_1,x_2,x_3,x_4)\\
&\quad + \int \mathrm{d}y\, \mathrm{d}z\, W^{(3)}_{\text{amp.}}(x_1,x_2,y) W^{(2)}(y,z) W^{(3)}_{\text{amp.}}(z,x_3,x_4) + 2 \text{ terms}\\
&\ldots
\end{aligned}$$

The inverse relations are even more useful, because, as we show in Section 7.8, $\Gamma(\varphi)$ has simpler properties than $\mathcal{W}(J)$

$$\begin{aligned}
W^{(2)}(x_1,x_2) &= [\Gamma^{(2)}]^{-1}(x_1,x_2),\\
W^{(3)}_{\text{amp.}}(x_1,x_2,x_3) &= -\Gamma^{(3)}(x_1,x_2,x_3),\\
W^{(4)}_{\text{amp.}}(x_1,x_2,x_3,x_4) &= -\Gamma^{(4)}(x_1,x_2,x_3,x_4)\\
&\quad + \int \mathrm{d}y\, \mathrm{d}z\, \Gamma^{(3)}(x_1,x_2,y) W^{(2)}(y,z) \Gamma^{(3)}(z,x_3,x_4) + 2 \text{ terms}\\
&\ldots
\end{aligned}$$

Let us give a graphical representation of the first equations. We define

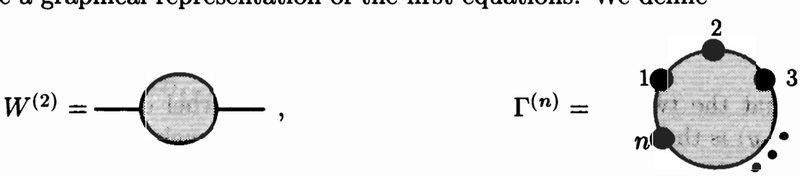


Fig. 7.4

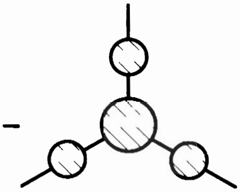


Fig. 7.5 The connected three-point function $W^{(3)}$.

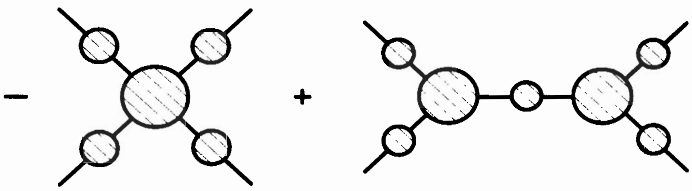


Fig. 7.6 The connected four-point function $W^{(4)}$.

The correlation functions $W^{(3)}$ and $W^{(4)}$ can then be represented as shown in figures 7.5 and 7.6, respectively.

Mass operator. It follows from the set of relations between connected correlation functions and proper vertices that Feynman diagrams that contribute to proper vertices 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 (7.20))

$$\Gamma^{(2)}(x,y) = K(x,y) + \Sigma(x,y),$$

where K is the two-point function $\Gamma^{(2)}$ in the free (gaussian) theory, then $\Sigma(x,y)$, called the mass operator, contains all perturbative corrections. The connected two-point function $W^{(2)}(x,y)$ then takes the form of a geometric series:

$$W^{(2)}(x,y) = \Delta(x,y) - \int \mathrm{d}z_1\, \mathrm{d}z_2\, \Delta(x,z_1)\Sigma(z_1,z_2)\Delta(z_2,y) + \cdots\ ,$$

where Δ is the propagator (7.15). 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

$$J\cdot\Xi \equiv \int \mathrm{d}x\, J(x)\Xi(x),$$

we immediately obtain

$$\int \mathrm{d}x\,\mathrm{d}y\, J(x)J(y)W^{(2)}(x,y) = \left\langle(J\cdot\Xi)^2\right\rangle \geq 0\,.$$

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,

$$\int \mathrm{d}x\,\mathrm{d}y\,\varphi(x)\Gamma^{(2)}(x,y)\varphi(y) \geq 0\,. \tag{7.81}$$

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 $\Gamma(\varphi)$ which are not obviously convex (see Section 7.10).

7.6 Momentum Representation

In this work we shall mainly study theories invariant under space translations. Correlation functions then depend only on differences of space arguments. It is thus natural to introduce the Fourier transforms of correlation functions. To establish a consistent set of conventions we start from the generating functional of proper vertices (in d space dimensions)

$$\Gamma(\varphi)=\sum_n \frac{1}{n!}\int\prod_{i=1}^{n}\mathrm{d}^d x_i\,\varphi(x_i)\Gamma^{(n)}(x_1,\ldots,x_n).$$

We introduce the Fourier components of the field $\varphi(x)$:

$$\varphi(x)=\int \mathrm{e}^{ipx}\,\widetilde{\varphi}(p)\mathrm{d}^d p\,. \tag{7.82}$$

We define the proper vertex $\widetilde{\Gamma}^{(n)}(p_1,\ldots,p_n)$ in the momentum representation in terms of the coefficient of $\widetilde{\varphi}(p_1)\ldots\widetilde{\varphi}(p_n)$ in $\Gamma(\varphi)$, taking into account translation invariance which implies momentum conservation,

$$(2\pi)^d\delta\left(\sum p_k\right)\widetilde{\Gamma}^{(n)}(p_1,\ldots,p_n)=\int\left(\prod_k \mathrm{d}^d x_k\,\mathrm{e}^{ix_k p_k}\right)\Gamma^{(n)}(x_1,\ldots,x_n). \tag{7.83}$$

In the same way we introduce the Fourier components $\tilde{J}(p)$ of the source $J(x)$:

$$J(x)=\int \mathrm{e}^{ipx}\,\tilde{J}(p)\mathrm{d}^d p\,,$$

which generate the correlation functions of $\int \mathrm{d}x\,\mathrm{e}^{ipx}\,\phi(x)$. Inserting this representation into the generating functional $\mathcal{W}(J)$ of connected correlation functions we define $\widetilde{W}^{(n)}(p_1,\ldots,p_n)$ by

$$(2\pi)^d\delta\left(\sum p_k\right)\widetilde{W}^{(n)}(p_1,\ldots,p_n)=\int\left(\prod_k \mathrm{d}^d x_k\,\mathrm{e}^{ix_k p_k}\right)W^{(n)}(x_1,\ldots,x_n). \tag{7.84}$$

Inverting the Fourier transformation, we find for the various two-point functions

$$\Gamma^{(2)}(x,y)=\frac{1}{(2\pi)^d}\int \mathrm{d}^d p\;\mathrm{e}^{ip(x-y)}\,\widetilde{\Gamma}^{(2)}(-p,p),$$
$$W^{(2)}(x,y)=\frac{1}{(2\pi)^d}\int \mathrm{d}^d p\;\mathrm{e}^{ip(x-y)}\,\widetilde{W}^{(2)}(-p,p).$$

The Legendre transformation then takes a simple form

$$\widetilde{\Gamma}^{(2)}(-p,p)\widetilde{W}^{(2)}(-p,p)=1\,. \tag{7.85}$$

In what follows, we denote simply by $\widetilde{\Gamma}^{(2)}(p)$ and $\widetilde{W}^{(2)}(p)$, the functions

$$\widetilde{\Gamma}^{(2)}(p)=\widetilde{\Gamma}^{(2)}(-p,p)\,,\quad \widetilde{W}^{(2)}(p)=\widetilde{W}^{(2)}(-p,p).$$

The explicit expressions of Section 7.5 (for example equation (7.79)) then show that amputation and Legendre transformation become in momentum representation purely algebraic operations in the sense that no momentum integration is involved because the two-point function is proportional to $\delta(p_1+p_2)$. For example,

$$\widetilde{W}^{(n)}_{\text{amp.}}(p_1,\ldots,p_n)=\widetilde{W}(p_1,\ldots,p_n)\prod_{i=1}^{n}\left[\widetilde{W}^{(2)}(p_i)\right]^{-1},$$

and also

$$\begin{aligned}\widetilde{\Gamma}^{(3)}(p_1,p_2,p_3)&=-\widetilde{W}^{(3)}_{\text{amp.}}(p_1,p_2,p_3),\\ \widetilde{\Gamma}^{(4)}(p_1,p_2,p_3,p_4)&=-\widetilde{W}^{(4)}_{\text{amp.}}(p_1,p_2,p_3,p_4)+\Big[\widetilde{W}^{(3)}_{\text{amp.}}(p_1,p_2,-p_1-p_2)\widetilde{W}^{(2)}(p_1+p_2)\\ &\quad\times\widetilde{W}^{(3)}_{\text{amp.}}(p_3,p_4,-p_3-p_4)+\text{ cyclic permutation of }\{p_2,p_3,p_4\}\Big],\\ &\ldots\,.\end{aligned}$$

This remark will be useful for the discussion of divergences in perturbation theory.

7.7 Semi-Classical or Loop Expansion

For reasons which will become clear at the end of this section, it is sometimes useful to reorganize perturbation theory by grouping some classes of Feynman diagrams. For this purpose, we perform a semi-classical expansion of the functional integral, that is, a formal expansion in powers of $\hbar$ which we, therefore, again set in front of the classical action and the source term:

$$\mathcal{Z}(J)=\mathcal{N}\int[\mathrm{d}\phi]\exp\left[-\frac{1}{\hbar}(\mathcal{S}(\phi)-J\cdot\phi)\right],\tag{7.86}$$

($\mathcal{Z}(0)=1$), in the symbolic notation of Section 7.2.

7.7.1 Loop expansion at leading order

For $\hbar\to 0$ the functional integral can be calculated by the steepest descent method. The saddle point equation is

$$\frac{\delta\mathcal{S}}{\delta\phi(x)}\left[\phi_c(J)\right]=J(x).\tag{7.87}$$

We assume below 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, we obtain $\mathcal{Z}(J)$ at leading order:

$$\ln\mathcal{Z}(J)\sim\ln\mathcal{Z}_0(J)\equiv\frac{1}{\hbar}\left[-\mathcal{S}(\phi_c)+J\cdot\phi_c\right].\tag{7.88}$$

When $\hbar$ is explicit, it is convenient to define $\mathcal{W}(J)$ by

$$\mathcal{W}(J)=\hbar\ln\mathcal{Z}(J).\tag{7.89}$$

Then, at leading order

$$\mathcal{W}(J) = \mathcal{W}_0(J) \equiv -\mathcal{S}(\phi_c) + J \cdot \phi_c \,. \tag{7.90}$$

Together, the two equations (7.87,7.90) imply that the two functionals $\mathcal{S}(\phi)$ and $\mathcal{W}_0(J)$ are related by 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 (7.20)). Moreover, we assume

$$\mathcal{V}_{\mathrm{I}}(\phi) = O\left(\phi^3\right) \quad \text{for } \phi \to 0\,.$$

Equation (7.87) then takes the symbolic form

$$K\phi_c + \frac{\delta \mathcal{V}_{\mathrm{I}}(\phi_c)}{\delta\phi} = J\,. \tag{7.91}$$

It can be iteratively solved as ($K\Delta = 1$)

$$\phi_c = \Delta J - \Delta \frac{\delta \mathcal{V}_{\mathrm{I}}(\phi_c)}{\delta\phi} = \Delta J - \Delta \frac{\delta \mathcal{V}_{\mathrm{I}}(\Delta J)}{\delta\phi} + \cdots\,. \tag{7.92}$$

If, for example,

$$\mathcal{V}_{\mathrm{I}}(\phi) = \frac{g}{4!} \int \mathrm{d}x\, \phi^4(x),$$

ϕ_c has the Feynman diagram expansion,

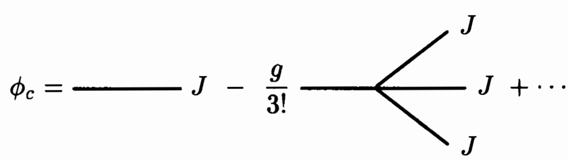


Fig. 7.7

We observe that only tree, that is, diagrams without loops, are generated. Substituting the expansion into equation (7.90) 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.

Proper vertices. 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

$$\Gamma_0(\varphi) = \mathcal{S}(\varphi). \tag{7.93}$$

At leading order, $\Gamma(\varphi)$, the generating functional of proper vertices, is identical to the classical action. From the point of view of Feynman diagrams, it contains only the vertices of the field theory.

7.7.2 Order $\hbar$ corrections

The gaussian integral obtained by expanding around the saddle point generates the order $\hbar$ corrections. We set

$$\phi = \phi_c(J) + \sqrt{\hbar}\chi\,, \tag{7.94}$$

$$S^{(2)}(x_1, x_2; \phi) = \frac{\delta^2 \mathcal{S}}{\delta\phi(x_1)\delta\phi(x_2)}\,. \tag{7.95}$$

Expanding the action in powers of $\hbar$ we find

$$\mathcal{S}(\phi) - J\cdot\phi = \mathcal{S}(\phi_c) - J\cdot\phi_c + \frac{\hbar}{2}\int \mathrm{d}x_1\,\mathrm{d}x_2\, S^{(2)}(x_1, x_2; \phi_c)\chi(x_1)\chi(x_2) + O\left(\hbar^{3/2}\right).$$

The functional integral at this order becomes

$$\mathcal{Z}(J) \sim \mathcal{Z}_0(J)\int [\mathrm{d}\chi]\exp\left[-\frac{1}{2}\int \mathrm{d}x_1\,\mathrm{d}x_2\, S^{(2)}(x_1, x_2; \phi_c)\chi(x_1)\chi(x_2)\right],$$

and, therefore,

$$\mathcal{Z}(J) \propto \mathcal{Z}_0(J)\left[\det S^{(2)}(x_1, x_2; \phi_c)/\det S^{(2)}(x_1, x_2; 0)\right]^{-1/2}, \tag{7.96}$$

where the normalization follows from the conditions $\phi_c(0) = 0$ and $\mathcal{Z}(0) = 1$.

The connected generating functional $\mathcal{W}(J)$ at this order is then ($\ln\det = \operatorname{tr}\ln$)

$$\mathcal{W}(J) = \mathcal{W}_0(J) + \hbar\mathcal{W}_1(J) + O\left(\hbar^2\right), \tag{7.97}$$

with

$$\mathcal{W}_1(J) = -\frac{1}{2}\left[\operatorname{tr}\ln S^{(2)}(x_1, x_2; \phi_c) - \operatorname{tr}\ln S^{(2)}(x_1, x_2; 0)\right]. \tag{7.98}$$

We again illustrate the result with the example of the ϕ^4 interaction:

$$S^{(2)}(x_1, x_2; \phi) = K(x_1, x_2) + \frac{g}{2}\phi^2(x_1)\delta(x_1 - x_2).$$

Then,

$$\operatorname{tr}\ln S^{(2)}(x_1, x_2; \phi_c) - \operatorname{tr}\ln S^{(2)}(x_1, x_2; 0) = \operatorname{tr}\ln\left[\delta(x_1 - x_2) + \frac{g}{2}\Delta(x_1, x_2)\phi_c^2(x_2)\right].$$

The expansion of $\mathcal{W}_1(J)$ in powers of ϕ_c takes the form

$$\begin{aligned}\mathcal{W}_1(J) = -\frac{1}{2}\Bigg[&\frac{g}{2}\int \mathrm{d}x_1\,\Delta(x_1, x_1)\phi_c^2(x_1)\\ &-\frac{g^2}{8}\int \mathrm{d}x_1\,\mathrm{d}x_2\,\Delta(x_1, x_2)\phi_c^2(x_2)\Delta(x_2, x_1)\phi_c^2(x_1) + \cdots\Bigg].\end{aligned}$$

It is important to 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 figure 7.8.

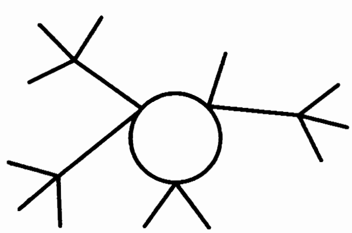


Fig. 7.8

The generating functional $\Gamma(\varphi)$ at order $\hbar$ can be obtained from the relation

$$\frac{\partial \Gamma}{\partial \hbar} + \frac{\partial \mathcal{W}}{\partial \hbar} = 0 ,$$

for $\hbar \to 0$. Thus,

$$\Gamma(\varphi) = \mathcal{S}(\varphi) - \hbar \mathcal{W}_1(J) = \mathcal{S}(\varphi) + \tfrac{1}{2}\hbar \left[\operatorname{tr} \ln S^{(2)}(x_1, x_2; \phi_c) - \operatorname{tr} \ln S^{(2)}(x_1, x_2; 0)\right] . \quad (7.99)$$

At this order we can replace $\phi_c(J)$ by φ. Setting

$$\Gamma(\varphi) = \mathcal{S}(\varphi) + \hbar \Gamma_1(\varphi) + O\left(\hbar^2\right) , \qquad (7.100)$$

we obtain for the order $\hbar$ correction $\Gamma_1(\varphi)$:

$$\Gamma_1(\varphi) = \tfrac{1}{2} \operatorname{tr} \left[\ln S^{(2)}(x_1, x_2; \varphi) - \ln S^{(2)}(x_1, x_2; 0)\right] . \qquad (7.101)$$

If $\mathcal{S}(\phi)$ has the decomposition (7.20) and $\mathcal{V}_{\mathrm{I}}(\phi)$ the local form,

$$\mathcal{V}_{\mathrm{I}}(\phi) = \int \mathrm{d}x \, V_{\mathrm{I}}\left(\phi(x)\right),$$

the expansion of $\Gamma_1(\varphi)$ in powers of V_{I} takes the form

$$\begin{aligned} \Gamma_1(\varphi) = \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} & \int \mathrm{d}x_1 \ldots \mathrm{d}x_n V_{\mathrm{I}}''(\varphi(x_1)) \Delta(x_1, x_2) \\ & \times V_{\mathrm{I}}''(\varphi(x_2)) \Delta(x_2, x_3) \ldots V_{\mathrm{I}}''(\varphi(x_n)) \Delta(x_n, x_1), \end{aligned} \qquad (7.102)$$

which in terms of Feynman diagrams has the representation of a sum of one-loop diagrams (see figure 7.9).

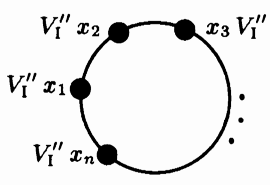


Fig. 7.9

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 (1-particle irreducible in the language of Particle Physics), that is, they cannot be disconnected by cutting only one line. The functional $\Gamma_1(\varphi)$ is the generating functional of one line irreducible one-loop Feynman diagrams.

The generating functional $\Gamma(\varphi)$ at order $\hbar$ can be obtained from the relation

$$\frac{\partial \Gamma}{\partial \hbar} + \frac{\partial \mathcal{W}}{\partial \hbar} = 0\,,$$

for $\hbar \to 0$. Thus,

$$\Gamma(\varphi) = \mathcal{S}(\varphi) - \hbar \mathcal{W}_1(J) = \mathcal{S}(\varphi) + \tfrac{1}{2}\hbar \left[\operatorname{tr}\ln S^{(2)}(x_1, x_2; \phi_c) - \operatorname{tr}\ln S^{(2)}(x_1, x_2; 0)\right]. \quad (7.99)$$

At this order we can replace $\phi_c(J)$ by φ. Setting

$$\Gamma(\varphi) = \mathcal{S}(\varphi) + \hbar \Gamma_1(\varphi) + O\left(\hbar^2\right), \quad (7.100)$$

we obtain for the order $\hbar$ correction $\Gamma_1(\varphi)$:

$$\Gamma_1(\varphi) = \tfrac{1}{2}\operatorname{tr}\left[\ln S^{(2)}(x_1, x_2; \varphi) - \ln S^{(2)}(x_1, x_2; 0)\right]. \quad (7.101)$$

If $\mathcal{S}(\phi)$ has the decomposition (7.20) and $\mathcal{V}_{\mathrm{I}}(\phi)$ the local form,

$$\mathcal{V}_{\mathrm{I}}(\phi) = \int \mathrm{d}x\, V_{\mathrm{I}}\left(\phi(x)\right),$$

the expansion of $\Gamma_1(\varphi)$ in powers of V_{I} takes the form

$$\begin{aligned}\Gamma_1(\varphi) = \frac{1}{2}\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \int & \mathrm{d}x_1 \ldots \mathrm{d}x_n V_{\mathrm{I}}''(\varphi(x_1))\Delta(x_1, x_2) \\ & \times V_{\mathrm{I}}''(\varphi(x_2))\Delta(x_2, x_3) \ldots V_{\mathrm{I}}''(\varphi(x_n))\Delta(x_n, x_1),\end{aligned} \quad (7.102)$$

which in terms of Feynman diagrams has the representation of a sum of one-loop diagrams (see figure 7.9).

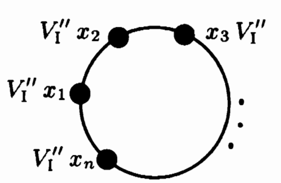


Fig. 7.9

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 (1-particle irreducible in the language of Particle Physics), that is, they cannot be disconnected by cutting only one line. The functional $\Gamma_1(\varphi)$ is the generating functional of one line irreducible one-loop Feynman diagrams.

7.8 Legendre Transformation and 1-Irreducibility

We have shown that the two first orders of the loop expansion of $\Gamma(\varphi)$ contain only one-line irreducible Feynman diagrams. We now suspect that the coefficient of $\hbar^L$ is the generating functional of all L-loop diagrams. By realizing that proper vertices 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 one line irreducible 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 (7.90)). However, we give here a more powerful and completely algebraic proof.

To prove that proper vertices are given in perturbation theory by a sum of one line or one particle irreducible (1PI) Feynman diagrams, we directly use the definition and prove that by cutting one line in all possible ways in a diagram contributing to $\Gamma(\varphi)$, the diagram remains connected.

We consider the modified action:

$$\mathcal{S}_\varepsilon(\phi) = \mathcal{S}(\phi) + \tfrac{1}{2}\varepsilon\left(\int \mathrm{d}x\,\phi(x)\right)^2 = \frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\,\phi(x)\phi(y)\left[K(x,y)+\varepsilon\right] + \mathcal{V}_{\mathrm{I}}(\phi), \quad (7.103)$$

where we have introduced a parameter ε in which we will expand at first order. The corresponding propagator $\Delta_\varepsilon(x,y)$:

$$\int \Delta_\varepsilon(x,z)\left[K(z,y)+\varepsilon\right]\mathrm{d}z = \delta(x-y),$$

can be written as

$$\Delta_\varepsilon(x,y) = \Delta(x,y) - \varepsilon\eta(x)\eta(y) + O\left(\varepsilon^2\right) \quad (7.104)$$

with the definition

$$\eta(x) = \int \Delta(x,z)\mathrm{d}z\,.$$

If we now expand a Feynman diagram with the new propagator $\Delta_\varepsilon(x,y)$ in ε, 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 allow to study irreducibility with respect to cutting two, three, ... lines.

We now calculate the partition function $\mathcal{Z}_\varepsilon(J)$ at first order in ε:

$$\mathcal{Z}_\varepsilon(J) = \int [\mathrm{d}\phi]\left(1-\frac{\varepsilon}{2}\int \mathrm{d}x\,\mathrm{d}y\,\phi(x)\phi(y)\right)\exp\left[-\mathcal{S}(\phi)+J\cdot\phi\right] + O\left(\varepsilon^2\right), \quad (7.105)$$

and, therefore,

$$\mathcal{Z}_\varepsilon(J) = \left[1-\frac{\varepsilon}{2}\int \mathrm{d}x\,\mathrm{d}y\,\frac{\delta^2}{\delta J(x)\delta J(y)} + O\left(\varepsilon^2\right)\right]\mathcal{Z}(J). \quad (7.106)$$

The generating functional,

$$\mathcal{W}_\varepsilon(J) = \ln \mathcal{Z}_\varepsilon(J), \tag{7.107}$$

is then given by

$$\mathcal{W}_\varepsilon(J) = \mathcal{W}(J) - \frac{\varepsilon}{2}\left\{\left[\int \mathrm{d}x \frac{\delta \mathcal{W}}{\delta J(x)}\right]^2 + \int \mathrm{d}x\, \mathrm{d}y \frac{\delta^2 \mathcal{W}}{\delta J(y)\delta J(x)}\right\} + O\left(\varepsilon^2\right). \tag{7.108}$$

It contains at order ε a contribution of the form

$$\left[\int \mathrm{d}x \frac{\delta \mathcal{W}}{\delta J(x)}\right]^2$$

which is disconnected, as expected.

In the Legendre transformation we use the relation (7.73) in the form

$$\frac{\partial \Gamma}{\partial \varepsilon} = -\frac{\partial \mathcal{W}}{\partial \varepsilon},$$

for $\varepsilon = 0$. Therefore,

$$\Gamma_\varepsilon(\varphi) = \Gamma(\varphi) + \frac{\varepsilon}{2}\left\{\left[\int \mathrm{d}x\, \varphi(x)\right]^2 + \int \mathrm{d}x\, \mathrm{d}y \left[\frac{\delta^2 \Gamma}{\delta\varphi(x)\delta\varphi(y)}\right]^{-1}\right\} + O\left(\varepsilon^2\right). \tag{7.109}$$

We see that the first order in ε contains two terms, the term $\frac{1}{2}\varepsilon\left[\int \mathrm{d}x\, \varphi(x)\right]^2$, which we have added explicitly to the action, and a second term which is the connected propagator 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 language of particle physics, one particle irreducible. In the chapters that follow, we use indifferently the terms of proper vertices or 1PI correlation functions. Note finally that in Section 7.5 we have shown that

$$W^{(n)}_{\text{amp.}} = -\Gamma^{(n)} + \text{ reducible terms} \quad (n > 2),$$

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.

7.9 Loop Expansion at Higher Orders

Steepest descent and loop expansion. We now expand the functional integral

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\frac{1}{\hbar}(\mathcal{S}(\phi) - J\cdot\phi)\right], \tag{7.110}$$

with $\mathcal{S}(\phi)$ of the form (7.20), to all orders in perturbation theory. The propagator is $\hbar\Delta$. Any vertex generated by $\mathcal{V}_{\mathrm{I}}(\phi)$ 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), The generating functional,

$$\mathcal{W}_\varepsilon(J) = \ln \mathcal{Z}_\varepsilon(J), \tag{7.107}$$

is then given by

$$\mathcal{W}_\varepsilon(J) = \mathcal{W}(J) - \frac{\varepsilon}{2}\left\{\left[\int \mathrm{d}x \, \frac{\delta \mathcal{W}}{\delta J(x)}\right]^2 + \int \mathrm{d}x \, \mathrm{d}y \, \frac{\delta^2 \mathcal{W}}{\delta J(y)\delta J(x)}\right\} + O\left(\varepsilon^2\right). \tag{7.108}$$

It contains at order ε a contribution of the form

$$\left[\int \mathrm{d}x \frac{\delta \mathcal{W}}{\delta J(x)}\right]^2$$

which is disconnected, as expected.

In the Legendre transformation we use the relation (7.73) in the form

$$\frac{\partial \Gamma}{\partial \varepsilon} = -\frac{\partial \mathcal{W}}{\partial \varepsilon},$$

for $\varepsilon = 0$. Therefore,

$$\Gamma_\varepsilon(\varphi) = \Gamma(\varphi) + \frac{\varepsilon}{2}\left\{\left[\int \mathrm{d}x \, \varphi(x)\right]^2 + \int \mathrm{d}x \, \mathrm{d}y \left[\frac{\delta^2 \Gamma}{\delta\varphi(x)\delta\varphi(y)}\right]^{-1}\right\} + O\left(\varepsilon^2\right). \tag{7.109}$$

We see that the first order in ε contains two terms, the term $\frac{1}{2}\varepsilon\left[\int \mathrm{d}x \, \varphi(x)\right]^2$, which we have added explicitly to the action, and a second term which is the connected propagator 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 language of particle physics, one particle irreducible. In the chapters that follow, we use indifferently the terms of proper vertices or 1PI correlation functions. Note finally that in Section 7.5 we have shown that

$$W^{(n)}_{\text{amp.}} = -\Gamma^{(n)} + \text{ reducible terms} \quad (n > 2),$$

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.

7.9 Loop Expansion at Higher Orders

Steepest descent and loop expansion. We now expand the functional integral

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\frac{1}{\hbar}\left(\mathcal{S}(\phi) - J\cdot\phi\right)\right], \tag{7.110}$$

with $\mathcal{S}(\phi)$ of the form (7.20), to all orders in perturbation theory. The propagator is $\hbar\Delta$. Any vertex generated by $\mathcal{V}_\mathrm{I}(\phi)$ 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

$$\hbar^{I+E-(V+E)+1},$$

the last factor $\hbar$ coming from our normalization of $\mathcal{W}(J)$ (equation (7.89)).

Note that the same result is obtained for one line irreducible Feynman diagrams (i.e. as we prove in the next section, 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. 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:

$$L = I - V + 1 . \tag{7.111}$$

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 get 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 (7.111) is again satisfied. This remark will eventually allow us to relate the number of loops to the number of independent momentum integration variables.

Higher orders calculations. We now indicate how successive terms in the loop expansion can be calculated by applying the steepest descent method to the functional integral (7.110).

The saddle point $\phi_c(J)$ is the solution of the equation

$$\frac{\delta \mathcal{S}}{\delta \phi(x)}(\phi_c) = J(x). \tag{7.112}$$

We change variables in the functional integral (7.86), $\phi \mapsto \chi$, setting

$$\phi(x) = \phi_c(x) + \sqrt{\hbar}\chi(x). \tag{7.113}$$

We expand $\mathcal{S}(\phi)$ in powers of χ:

$$\begin{aligned}\mathcal{S}(\phi_c+\chi) = \mathcal{S}(\phi_c) &+ \frac{\hbar}{2!}\int \mathrm{d}x_1\,\mathrm{d}x_2\,\chi(x_1)\chi(x_2)S^{(2)}(x_1,x_2;\phi_c)\\ &+ \frac{\hbar^{3/2}}{3!}\int \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\chi(x_1)\chi(x_2)\chi(x_3)S^{(3)}(x_1,x_2,x_3;\phi_c)\\ &+ \frac{\hbar^2}{4!}\int \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_4\,\chi(x_1)\ldots\chi(x_4)S^{(4)}(x_1,\ldots,x_4;\phi_c)+\cdots\,, \end{aligned}\tag{7.114}$$

where we have introduced the notation

$$\left.\frac{\delta^n\mathcal{S}(\phi)}{\delta\phi(x_1)\ldots\delta\phi(x_n)}\right|_{\phi=\phi_c} = S^{(n)}(x_1,\ldots,x_n;\phi_c). \tag{7.115}$$

The expansion in powers of $\hbar$ and the integration over χ generate vacuum Feynman diagrams with a ϕ_c-dependent propagator,

$$\int S^{(2)}(x,y;\phi_c)\Delta(y,z;\phi_c)\mathrm{d}y = \delta(x-z), \tag{7.116}$$

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 only removes the one-line reducible 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 two-loop order in Appendix A7.1.

7.10 Statistical and Quantum Interpretation of the 1PI 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 $\Gamma(\varphi)$.

7.10.1 Interpretation and variational principle

Statistical interpretation. We consider the free energy $\mathcal{W}(J)$,

$$\mathrm{e}^{\mathcal{W}(J)} = \int[\mathrm{d}\phi]\,\mathrm{e}^{-\mathcal{S}(\phi)+J\cdot\phi}\,.$$

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 $\mathrm{e}^{-\mathcal{S}(\phi)+J\cdot\phi}$.

We now use the convexity property of the exponential function:

$$\ln\left\langle \mathrm{e}^{-J\cdot\phi}\right\rangle_J = \mathcal{W}(0) - \mathcal{W}(J) \ge \langle -J\cdot\phi\rangle_J = -J\cdot\varphi\,,$$

or introducing the Legendre transform $\Gamma(\varphi)$ of $\mathcal{W}(J)$,

$$\mathcal{W}(0) \geq -\Gamma(\varphi). \tag{7.117}$$

This inequality can be the starting point of a variational principle. Moreover, $\mathcal{W}(0)$ is obtained in the limit $J = 0$ by taking the solution of

$$\frac{\delta\Gamma(\varphi)}{\delta\varphi(x)} = 0\,, \tag{7.118}$$

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.

Note, finally, that we have here 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. We now distinguish in the d-dimensional space $\mathbb{R}^d$ a 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 ourselves to time-independent sources $J(t, x)$:

$$J(t,x) = J(x),$$

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 $T = 1/\beta$:

$$\mathcal{Z} = \operatorname{tr} \mathrm{e}^{-\beta\hat{H}}\,.$$

The inequality (7.117) becomes

$$\ln \operatorname{tr} \mathrm{e}^{-\beta\hat{H}} \geq -\Gamma(\varphi), \tag{7.119}$$

where φ is the thermal expectation value corresponding to the hamiltonian

$$\hat{H}(\phi, J) = \hat{H}(\phi) - \int \mathrm{d}^{d-1}x\, J(x)\hat{\phi}(x), \tag{7.120}$$

$\hat{\phi}(x)$ 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,

$$\mathcal{W}(0) \underset{\beta\to\infty}{\sim} -\beta E_0\,. \tag{7.121}$$

In this limit (7.119) becomes

$$E_0 \leq \frac{1}{\beta}\Gamma(\varphi)\,, \tag{7.122}$$

where again E_0 is obtained by looking for the solution of equation (7.118) which minimizes $\Gamma(\varphi)$.

7.10.2 1PI 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

$$\varphi(x) = \frac{1}{\beta} \int_0^\beta \mathrm{d}t \, \phi(t, x) \,. \tag{7.123}$$

Note that this implies trivially $\varphi(x) = \langle \phi(t, x) \rangle$. We denote by $-\beta \mathcal{G}(\varphi)$ the corresponding free energy

$$\mathrm{e}^{-\beta \mathcal{G}(\varphi)} = \int [\mathrm{d}\phi(t, x)] \exp[-\mathcal{S}(\phi)]. \tag{7.124}$$

We have written the free energy in the form $-\beta \mathcal{G}(\varphi)$ because we know 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

$$\mathrm{e}^{\mathcal{W}(J)} = \int [\mathrm{d}\varphi(x)] \exp \left[-\beta \mathcal{G}(\varphi) + \beta \int \mathrm{d}x \, J(x) \varphi(x) \right] . \tag{7.125}$$

For β large, the functional integral can be calculated by the steepest descent method. The saddle point equation is

$$J(x) = \frac{\delta \mathcal{G}}{\delta \varphi(x)} \,. \tag{7.126}$$

When the equation has several solutions one has to take the stable solution which yields the largest contribution to the free energy. Then,

$$\mathcal{W}(J) \sim -\beta \mathcal{G}(\varphi) + \beta \int \mathrm{d}x \, J(x) \varphi(x). \tag{7.127}$$

After Legendre transformation one finds

$$\beta \mathcal{G}(\varphi) = \Gamma(\varphi),$$

where again $\Gamma(\varphi)$ is the 1PI functional restricted to time-independent fields. Note, however, 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 discussing the minima of $\Gamma(\varphi)$, that is, the minima of a quantity which has convexity properties and can have only one minimum. Actually one discusses the properties of $\mathcal{G}$.

Bibliographical Notes

Covariant perturbation was developed in

S. Tomonoga, *Prog. Theor. Phys.* 1 (1946) 27; J. Schwinger, *Phys. Rev.* 74 (1948) 1439; 75 (1949) 651; 76 (1949) 790; R.P. Feynman, *Phys. Rev.* 76 (1949) 769; F.J. Dyson, *Phys. Rev.* 75 (1949) 486, 1736.

For Wick's theorem see

G.C. Wick, *Phys. Rev.* 80 (1950) 268.

The discussion of the perturbative definition of the functional integral is taken from

A.A. Slavnov, *Theor. Math. Phys.* 22 (1975) 123; J. Zinn-Justin in *Trends in Elementary Particle Physics (Lectures Notes in Physics 37)* Bonn 1974, H. Rollnik and K. Dietz eds. (Springer-Verlag, Berlin 1975).

The quantum field equations appear in

F.J. Dyson, *Phys. Rev.* 75 (1949) 1736; J. Schwinger, *Proc. Natl. Acad. Sci. USA* 37 (1951) 452, 455.

The idea of using generating functionals with sources was originated in

J. Schwinger, *Proc. Natl. Acad. Sci. USA* 37 (1951) 452, 455.

Duality between action and generating functional of connected functions is discussed in

F. Englert and C. De Dominicis, *Nuovo Cim.* 53A (1968) 1007, 1021.

The Legendre transform of connected correlation functions has been discussed in the statistical context in

C. De Dominicis, *J. Math. Phys.* 4 (1963) 255; C. De Dominicis and P.C. Martin, *J. Math. Phys.* 5 (1964) 14, 31;

and introduced in field theory in

G. Jona-Lasinio, *Nuovo Cimento* 34 (1964) 1790.

See, also

J. Goldstone, A. Salam and S. Weinberg, *Phys. Rev. 127* (1962) 965.

A proof of the one-line irreducibility can be found in

S. Coleman, *Laws of Hadronic Matter* (discussion of lecture 3), Erice 1973, A. Zichichi ed. (Academic Press, New York 1975).

The physical interpretation of the 1PI functional is taken from

K. Symanzik, *Commun. Math. Phys.* 16 (1970) 48; see also S. Coleman, reference above.

The loop expansion is discussed in the context of quantum gravity in

B.S. de Witt, *Phys. Rev.* 162 (1967) 1195; 162 (1967) 1239.

It has been shown that the tree approximation is the leading order in $\hbar \to 0$ limit in

Y. Nambu, *Phys. Lett.* 26B (1968) 626.

The one-loop functional is calculated by the steepest descent method applied to the functional integral in

B.W. Lee and J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3121 Appendix B; R. Jackiw, *Phys. Rev.* D9 (1974) 1686.

The proof to all orders that the expansion in powers of $\hbar$ generates a loop expansion is given in

S. Coleman and E. Weinberg, *Phys. Rev.* D7 (1973) 1888.

The use of the solution of the classical field equation as argument of the generating functional or background field method (see Appendix A7.2) was suggested in

B.S. DeWitt in *Relativity, Groups and Topology*, Les Houches 1963, C. DeWitt and B. DeWitt eds. (Gordon and Breach, New York 1964); *Phys. Rev.* 162 (1967) 1195; G. 't Hooft, *Nucl. Phys.* B62 (1973) 444.

APPENDIX A7

A7.1 Two-Loop Calculation

We illustrate the remarks of Section 7.9 by an explicit two-loop calculation. Expanding the interaction terms $S^{(3)}$ and $S^{(4)}$ (equation (7.115)) and integrating term by term over χ (equation (7.113)) we obtain for $\mathcal{W}(J)=\hbar\ln\mathcal{Z}(J)$

$$\mathcal{W}(J)=-\mathcal{S}(\phi_c)+J\cdot\phi_c-\tfrac{1}{2}\hbar\operatorname{tr}\ln S^{(2)}(x_1,x_2;\phi_c)+\hbar^2\mathcal{W}_2(J)+O\left(\hbar^3\right) \qquad (A7.1)$$

with

$$\begin{aligned}\mathcal{W}_2(J)=&-\frac{1}{8}\int \mathrm{d}x_1\ldots\mathrm{d}x_4\, S^{(4)}(x_1,x_2,x_3,x_4)\Delta(x_1,x_2;\phi_c)\Delta(x_3,x_4;\phi_c)\\ &+\int \mathrm{d}x_1\ldots\mathrm{d}y_3\, S^{(3)}(x_1,x_2,x_3;\phi_c)S^{(3)}(y_1,y_2,y_3;\phi_c)\left[\frac{1}{8}\Delta(x_1,x_2;\phi_c)\right.\\ &\left.\times\Delta(y_1,y_2;\phi_c)\Delta(x_3,y_3;\phi_c)+\frac{1}{12}\Delta(x_1,y_1;\phi_c)\Delta(x_2,y_2;\phi_c)\Delta(x_3,y_3;\phi_c)\right], \qquad (A7.2)\end{aligned}$$

where a simple normalization of the functional integral has been chosen.

We now perform the Legendre transformation:

$$\Gamma(\varphi)+\mathcal{W}(J)=\int J(x)\varphi(x)\mathrm{d}x\,,\qquad \varphi(x)=\frac{\delta\mathcal{W}}{\delta J(x)}\,\cdot \qquad (A7.3)$$

We need $\varphi(x)$ only up to order $\hbar$ because expression $\Gamma(\varphi)-J\cdot\varphi$ is stationary in φ:

$$\varphi(x)=\phi_c(J;x)-\frac{\hbar}{2}\frac{\delta}{\delta J(x)}\operatorname{tr}\ln S^{(2)}(x_1,x_2;\phi_c)+O\left(\hbar^2\right). \qquad (A7.4)$$

Using $\delta\operatorname{tr}\ln X=\operatorname{tr}\delta X X^{-1}$, valid for any matrix or operator X, and applying the chain rule, we rewrite the order $\hbar$ correction:

$$\frac{\delta}{\delta J(x)}\operatorname{tr}\ln S^{(2)}(x_1,x_2;\phi_c)=\int \mathrm{d}y\,\mathrm{d}z_1\,\mathrm{d}z_2\frac{\delta\phi_c(y)}{\delta J(x)}S^{(3)}(y,z_1,z_2;\phi_c)\Delta(z_2,z_1,\phi_c).$$

Then, using equations (7.112,A7.4) and definition (7.116), we express ϕ_c in terms of φ:

$$\varphi(x)=\phi_c(J,x)-\frac{\hbar}{2}\int \mathrm{d}y\,\mathrm{d}y_1\,\mathrm{d}y_2\, S^{(3)}(y,y_1,y_2;\varphi)\Delta(x,y;\varphi)\Delta(y_1,y_2;\varphi)+O\left(\hbar^2\right). \qquad (A7.5)$$

We still need $J(x)$ at order $\hbar$. Using equations (7.112) and (A7.5) we find that

$$J(x)=\frac{\delta\mathcal{S}(\varphi)}{\delta\phi(x)}+\frac{\hbar}{2}\int \mathrm{d}y_1\,\mathrm{d}y_2\, S^{(3)}(x,y_1,y_2;\varphi)\Delta(y_1,y_2;\varphi)+O\left(\hbar^2\right). \qquad (A7.6)$$

Equation (A7.3) then yields $\Gamma(\varphi)$ at two-loop order. As expected the reducible part in expression (A7.2) (figure 7.10) cancels and we obtain

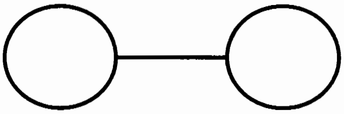


Fig. 7.10 The reducible part at two-loop order.

$$\Gamma(\varphi) = \mathcal{S}(\varphi) + \tfrac{1}{2}\hbar \operatorname{tr} \ln S^{(2)}(x_1, x_2; \varphi) + \hbar^2 \Gamma_2(\varphi) + O\left(\hbar^3\right) \qquad (A7.7)$$

with

$$\begin{aligned} \Gamma_2(\varphi) = {} & \frac{1}{8} \int \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3\, \mathrm{d}x_4\, \Delta(x_1, x_2; \varphi) S^{(4)}(x_1, x_2, x_3, x_4; \varphi) \Delta(x_3, x_4; \varphi) \\ & - \frac{1}{12} \int \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3\, \mathrm{d}y_1\, \mathrm{d}y_2\, \mathrm{d}y_3\, S^{(3)}(x_1, x_2, x_3; \varphi) \Delta(x_1, y_1; \varphi) \Delta(x_2, y_2; \varphi) \\ & \times \Delta(x_3, y_3; \varphi) S^{(3)}(y_1, y_2, y_3; \varphi). \end{aligned} \qquad (A7.8)$$

Figure 7.11 gives a diagrammatic representation of the two-loop terms in the equation.

Fig. 7.11 The two-loop contributions to $\Gamma(\varphi)$.

A7.2 The Background Field Method

In most of our study we have regarded 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 can be given:

(i) Some models are defined on Riemannian manifolds and the fields $\phi_i(x)$ correspond to a particular choice of coordinates on the manifold. For some problems, 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) We have shown that in Particle Physics, normalized S-matrix elements are invariant under a change of field variables (see Section 6.10.2).

Here, we want to introduce a method, the background field method, which has among its main merits, that it allows a more direct calculation of quantities that are rather insensitive to a change of field variables.

We consider a field theory with an action $\mathcal{S}(\phi)$:

$$\mathrm{e}^{\mathcal{W}(J)} = \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{S}(\phi) + J\phi}\,. \qquad (A7.9)$$

We introduce the Legendre transform of $\mathcal{W}(J)$:

$$\mathrm{e}^{-\Gamma(\varphi) + J\varphi} = \int [\mathrm{d}\phi]\, \mathrm{e}^{-\mathcal{S}(\phi) + J\phi}\,. \qquad (A7.10)$$

Using

$$J(x) = \frac{\delta \Gamma}{\delta \varphi(x)},$$

we write equation $(A7.10)$:

$$\mathrm{e}^{-\Gamma(\varphi)} = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, (\phi(x) - \varphi(x)) \frac{\delta \Gamma}{\delta \varphi(x)}\right], \qquad (A7.11)$$

or equivalently translating $\phi(x)$, $\phi \mapsto \varphi + \chi$:

$$\mathrm{e}^{-\Gamma(\varphi)} = \int [\mathrm{d}\chi] \exp\left[-\mathcal{S}(\chi+\varphi) + \int \mathrm{d}x\, \chi(x) \frac{\delta\Gamma}{\delta\varphi(x)}\right]. \tag{A7.12}$$

We now assume that the equation

$$\frac{\delta\Gamma}{\delta\varphi(x)} = 0\,, \tag{A7.13}$$

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:

$$\frac{\delta\mathcal{S}}{\delta\varphi(x)}\left[\varphi_c^{(0)}\right] = 0\,. \tag{A7.14}$$

Then equation ($A7.12$) becomes (see also the discussion at the end of Section 7.9)

$$\mathrm{e}^{-\Gamma(\varphi_c)} = \int [\mathrm{d}\chi]\, \mathrm{e}^{-\mathcal{S}(\chi+\varphi_c)}\,. \tag{A7.15}$$

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 this section. We introduce renormalized quantities (Chapter 10):

$$\Gamma_{\mathrm{r}}(\varphi_c) = -\ln \int [\mathrm{d}\chi] \exp\left[-\mathcal{S}_0(\chi+\varphi_c) + \text{counter-terms}\right],$$

in which $\mathcal{S}_0(\phi)$ is the tree order action. The solution φ_c of

$$\frac{\delta\Gamma_{\mathrm{r}}}{\delta\varphi_c} = 0$$

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 6.10.1 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 Cluster Properties of Connected Feynman Diagrams

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 correlation functions

The propagator can be written (equation (7.19)) as

$$\Delta(x,y) = \frac{1}{(2\pi)^d} \int \mathrm{d}^d p \, \frac{\mathrm{e}^{ip\cdot(x-y)}}{p^2+m^2} \, . \tag{A7.16}$$

To determine the behaviour of Δ for $|x-y|$ large, we rewrite the integral

$$\Delta(x,y) = \frac{1}{(2\pi)^d} \int \mathrm{d}^d p \int_0^{+\infty} \mathrm{d}t \, \mathrm{e}^{ip\cdot(x-y)} \, \mathrm{e}^{-t(p^2+m^2)} \, . \tag{A7.17}$$

We then perform the gaussian integration over the momentum p:

$$\Delta(x,y) = \frac{\pi^{d/2}}{(2\pi)^d} \int \frac{\mathrm{d}t}{t^{d/2}} \exp\left[-tm^2 - \frac{1}{4t}(x-y)^2\right] . \tag{A7.18}$$

The behaviour of Δ for large separation is given by the method of steepest descent. The saddle point is

$$t = \frac{|x-y|}{2m} \, . \tag{A7.19}$$

The gaussian integration over fluctuations around the saddle point finally yields

$$\Delta(x-y) \sim \frac{1}{2m} \left(\frac{m}{2\pi\,|x-y|}\right)^{(d-1)/2} \mathrm{e}^{-m|x-y|} \, . \tag{A7.20}$$

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 figure 7.12).

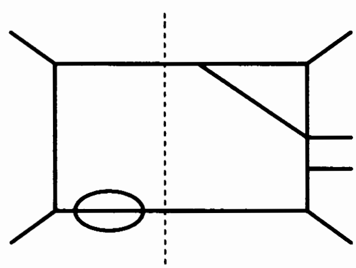


Fig. 7.12 Example with $n = 2$.

In a massless theory instead ($m = 0$) the decay is 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. For example, if, in momentum space, a contribution to the two-point function has an algebraic singularity

$$\tilde{K}^{(2)}(p) \propto (m^2 + p^2)^{-\alpha},$$

a generalization of the previous calculation yields a large distance behaviour

$$K^{(2)}(r) \propto r^{\alpha-d/2-1/2}\,\mathrm{e}^{-mr}\,.$$

If we now consider a 1PI diagram with n internal lines, it yields for r large a contribution

$$K^{(2)}(r) \propto r^{-n(d-1)/2}\,\mathrm{e}^{-nmr}\,.$$

This in turn corresponds to a singularity

$$\tilde{K}^{(2)}(p) \propto (m^2 + p^2)^{-\alpha}, \qquad \alpha = \tfrac{1}{2}[(n-1)d - n - 1].$$

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 RELATIVISTIC FERMIONS

Although we shall discuss mostly boson field theories, occasionally we will consider theories with fermions, in particular, when the fermion nature of fields plays an essential role.

In Chapter 5 we have already introduced the fundamental concepts we need to describe Fermi fields: quantum mechanics with Grassmann variables in Section 5.3 and a representation of the statistical operator $e^{-\beta H}$ for the non-relativistic Fermi gas by functional integrals in the formalism of second quantization. In Section 6.5 we have then obtained an expression for the evolution operator.

To discuss relativistic fermions we have to first recall the transformation properties of fermions under the spin group. Therefore, unlike what we have done with scalar bosons we begin with real time, in the spirit of Chapter 6, a direct euclidean presentation being less intuitive.

We analyse the free action for Dirac fermions, and explain the relation between fields and particles, then derive an expression of the scattering matrix and discuss the non-relativistic limit 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 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 functional 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 8.1 we note the connection between spin and statistics for free fermions. In Section 8.5 we verify the property in a simple perturbative calculation.

For completeness, in Appendix A8, we describe a few properties of the spin group, the algebra of γ matrices, and the corresponding spinors.

8.1 Massive Dirac Fermions

We first consider massive relativistic fermions that generalize the four-dimensional Dirac fermions, in even dimensions d because otherwise the fermion mass breaks space reflection symmetry. We describe in the real time formulation the structure of fermion states in the free field theory, and construct the S-matrix. In real time the spin group associated with the relativistic group $SO(1, d-1)$ (generalized Lorentz group) of space–time pseudo-orthogonal transformations is Spin$(1, d-1)$, groups that we will not discuss extensively, because it is beyond the scope of this work. Instead in Appendix A8 we explain our notation and conventions about γ matrices and present the euclidean analytic continuation Spin(d) of the spin group.

8.1.1 The free massive Dirac fermion

In the formalism of functional integrals, fermion fields $\bar{\psi}_\alpha(x), \psi_\alpha(x)$ are generators of a Grassmann algebra. In a relativistic invariant theory they are in addition spinors, vectors with $2^{d/2}$ components, transforming under the action of the spin group. The free action for a massive Dirac field can be written as

$$\mathcal{A}_0(\bar{\psi}, \psi) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x\, \bar{\psi}_\alpha(t, x) \left[\frac{1}{i}(\gamma_0)_{\alpha\beta}\partial_t + \boldsymbol{\gamma}_{\alpha\beta} \cdot \nabla_x + m\delta_{\alpha\beta}\right] \psi_\beta(t, x)\,, \tag{8.1}$$

where the hermitian γ matrices are defined in Section A8.1.4, the total dimension d is even, and, here, we denote by γ_0 the matrix γ_d associated with the time variable $t \equiv x_0 \equiv -ix_d$, $\boldsymbol{\gamma}$ standing for $(\gamma_1, \ldots, \gamma_{d-1})$.

The action is relativistic invariant in the sense that it is invariant under the transformation of the spin group $\mathrm{Spin}(1, d-1)$ of matrices Λ:

$$\begin{aligned}
(\psi_\Lambda)_\alpha (t, x) &= \Lambda^{-1}_{\alpha\beta} \psi_\beta(\mathbf{R}(t, x)), \\
(\bar{\psi}_\Lambda)_\alpha (t, x) &= \bar{\psi}_\beta(\mathbf{R}(t, x))\Lambda_{\beta\alpha},
\end{aligned} \tag{8.2}$$

where Λ belongs to the spin group, $\mathrm{Spin}(1, d-1)$, and $\mathbf{R}(\Lambda)$ is the corresponding $d \times d$ matrix, element of the group $SO(1, d-1)$. This group preserves the metric $\mathbf{g}$, where $\mathbf{g}$ is the diagonal matrix with coefficients $(+1, -1, \ldots, -1)$ (${}^T\!R$ denotes the transpose of the matrix $\mathbf{R}$):

$$R\,\mathbf{g}\,{}^T\!R = \mathbf{g}\,.$$

After the linear change of variables $\mathbf{R}(t, x) \mapsto (t', x')$ the invariance of the action follows from the identity

$$R_{\mu\nu}\Lambda\tilde{\gamma}_\nu\Lambda^{-1} = \tilde{\gamma}_\mu$$

with $\tilde{\gamma}_0 = \gamma_0$, $\tilde{\boldsymbol{\gamma}} = i\boldsymbol{\gamma}$.

In the relativistic conventions one introduces a field $\bar{\psi}$ which is not the hermitian conjugate of ψ because it has simpler transformation properties under the spin group. Indeed, matrices Λ belonging to the fundamental representation of the spin group, $\mathrm{Spin}(1, d-1)$, satisfy

$$\gamma_0\Lambda^\dagger\gamma_0 = \Lambda^{-1},$$

as one can verify by adapting the expressions of Appendix A8 to real time.

To identify the action (8.1) with an action of the form (6.47*b*) resulting from a hamiltonian formalism, it is convenient, however, to express $\bar{\psi}$ in terms of $\psi^\dagger$. With our conventions

$$\psi^\dagger = -\bar{\psi}\gamma_0\,. \tag{8.3}$$

Then,

$$\mathcal{A}_0(\psi, \psi^\dagger) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x\, \psi^\dagger(t, x)\left[i\partial_t - \gamma_0(\boldsymbol{\gamma} \cdot \nabla_x + m)\right]\psi(t, x)\,.$$

Since the γ matrices are hermitian we find $\mathcal{A}_0 = \mathcal{A}_0^\dagger$ and thus the corresponding hamiltonian is hermitian.

To diagonalize the quadratic form, we proceed by Fourier transformation, setting

$$\psi(t, x) = \int \mathrm{d}^{d-1}\hat{p}\; \mathrm{e}^{i\hat{p}x}\, \tilde{\psi}(t, \hat{p}), \quad \psi^\dagger(t, x) = \int \mathrm{d}^{d-1}\hat{p}\; \mathrm{e}^{-i\hat{p}x}\, \tilde{\psi}^\dagger(t, \hat{p}).$$

The free action becomes

$$\mathcal{A}_0(\psi,\psi^\dagger)=(2\pi)^{d-1}\int \mathrm{d}t\,\mathrm{d}^{d-1}\hat{p}\,\tilde{\psi}^\dagger(t,\hat{p})\left[i\partial_t+h(\hat{p})\right]\tilde{\psi}(t,\hat{p}),$$

where we have introduced the matrix $h(\hat{p})$:

$$h(\hat{p})=-\gamma_0\left(i\gamma\cdot\hat{p}+m\right),\qquad h(\hat{p})=h^\dagger(\hat{p}). \tag{8.4}$$

We see that, 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 simple matrix problem. One verifies

$$h^2(\hat{p})=\omega^2(\hat{p}),\quad \omega(\hat{p})=\sqrt{\hat{p}^2+m^2}\,.$$

The matrix h thus has two eigenvalues $\pm\omega(\hat{p})$. Since

$$\gamma_0\gamma_S h(\hat{p})\gamma_S\gamma_0=-h(\hat{p}),$$

if a spinor $u(\hat{p})$ is an eigenvector with eigenvalue $\omega(\hat{p})$, $\gamma_0\gamma_S u(\hat{p})$ ($\gamma_0\gamma_S$ is associated with time reversal) is an eigenvector with eigenvalue $-\omega(\hat{p})$. The two corresponding subspaces have equal dimensions.

The two orthogonal, hermitian projectors $P_\pm$ on the positive and negative energy sector are

$$P_\pm=\tfrac{1}{2}\left[1\pm h(\hat{p})/\omega(\hat{p})\right],\quad\Rightarrow\ P_++P_-=1\,,\quad P_\pm^2=P_\pm\,,\quad P_+P_-=0\,. \tag{8.5}$$

We note that with the simplest assignment of ψ as the Grassmann field associated with creation operators, states can be created with both positive and negative energies. This means that we have misidentified the reference state, which must be the ground state. As we have shown in Section 6.5, 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 $\psi^\dagger$ must be decomposed into a sum of analytic and anti-analytic components, to ensure that one-particle states have positive energy.

We thus set

$$\chi_-^*(t,\hat{p})=[2\omega(\hat{p})]^{1/2}P_-\tilde{\psi}(t,\hat{p}),\qquad \varphi_+(t,\hat{p})=[2\omega(\hat{p})]^{1/2}P_+\tilde{\psi}(t,\hat{p}) \tag{8.6a}$$

$$\varphi_+^\dagger(t,\hat{p})=[2\omega(\hat{p})]^{1/2}\tilde{\psi}^\dagger(t,\hat{p})P_+\,,\qquad {}^T\chi_-(t,\hat{p})=[2\omega(\hat{p})]^{1/2}\tilde{\psi}^\dagger(t,\hat{p})P_-\,. \tag{8.6b}$$

Conversely,

$$\tilde{\psi}(t,\hat{p})=\frac{1}{\sqrt{2\omega(\hat{p})}}\left(\chi_-^*(t,\hat{p})+\varphi_+(t,\hat{p})\right) \tag{8.7a}$$

$$\tilde{\psi}^\dagger(t,\hat{p})=\frac{1}{\sqrt{2\omega(\hat{p})}}\left({}^T\chi_-(t,\hat{p})+\varphi_+^\dagger(t,\hat{p})\right). \tag{8.7b}$$

The action becomes

$$\begin{aligned}\mathcal{A}_0(\chi_-,\varphi_+)=(2\pi)^{d-1}\int \mathrm{d}t\,\frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})}&\Big[\varphi_+^\dagger(t,\hat{p})\left(i\partial_t+\omega(\hat{p})\right)\varphi_+(t,\hat{p})\\ &+\chi_-^\dagger(t,\hat{p})\left(i\partial_t+\omega(\hat{p})\right)\chi_-(t,\hat{p})\Big]\,.\end{aligned}$$

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 Spin$(d-1)$ (see Appendix A8), χ_- under the conjugated representation. Actually, these representations are equivalent. To verify it we set

$$\varphi_-(t,\hat{p}) = C^\dagger \chi_-(t,\hat{p}), \tag{8.8}$$

where C is a unitary matrix, $C^\dagger C = 1$. The field χ_- satisfies ${}^T P_- \chi_- = \chi_-$. If we can find a matrix C such that

$$C^\dagger \, {}^T P_- C = P_+ \,,$$

then $P_+\varphi_- = \varphi_-$, and the two fields $\varphi_\pm$ have the same transformation properties. Since C is unitary, this equation reduces to

$$C^\dagger \, {}^T h(\hat{p}) C = -h(\hat{p}) \quad \Leftrightarrow \quad C^\dagger \left(i\, {}^T\gamma\cdot\hat{p} + m\right) {}^T\gamma_0 C = -\gamma_0 \left(i\gamma\cdot\hat{p} + m\right) .$$

In Appendix A8.1.7, we construct a charge conjugation matrix C which satisfies (equation (A8.39))

$$C^{-1}\, {}^T\gamma_\mu C = -\gamma_\mu \,.$$

and thus has the required property.

The action then reads

$$\mathcal{A}_0(\varphi_+,\varphi_-) = (2\pi)^{d-1} \int \mathrm{d}t \, \frac{\mathrm{d}^{d-1}\hat{p}}{2\omega(\hat{p})} \sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t,\hat{p})\left(i\partial_t + \omega(\hat{p})\right)\varphi_\epsilon(t,\hat{p}). \tag{8.9}$$

The final form of the action shows that the Dirac field carries two particles transforming under the fundamental representation of the spin group 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).

In these variables, the free hamiltonian reads

$$\mathbf{H}_0 = (2\pi)^{d-1} \int \mathrm{d}^{d-1}\mathrm{d}\hat{p}\,\omega(\hat{p}) \left(\varphi_+(\hat{p})\frac{\delta}{\delta\varphi_+(\hat{p})} - \frac{\delta}{\delta\varphi_-(\hat{p})}\varphi_-(\hat{p})\right).$$

If we now write the hamiltonian in normal-order we find

$$\mathbf{H}_0 = (2\pi)^{d-1} \int \mathrm{d}^{d-1}\mathrm{d}\hat{p}\,\omega(\hat{p}) \left(\varphi_+(\hat{p})\frac{\delta}{\delta\varphi_+(\hat{p})} + \varphi_-(\hat{p})\frac{\delta}{\delta\varphi_-(\hat{p})}\right) + E_0(\mathrm{Dirac})\,.$$

The ground state (vacuum) energy E_0(Dirac) is negative, and proportional to the free scalar vacuum energy (6.68):

$$E_0(\mathrm{Dirac}) = -2^{d/2} E_0(\mathrm{scalar}).$$

Remarks.

(i) Note, therefore, that by adding $2^{d/2}$ scalar bosons of the same mass m to one Dirac fermion, one can construct a theory with zero vacuum energy. One can show that this boson–fermion free theory then has a special fermion-type symmetry called supersymmetry.

(ii) We have seen that the possibility to solve 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 quantum field theory.

8.1.2 Interacting theory and S-matrix

We are now in a situation quite analogous to the scalar case, and the non-relativistic Fermi gas of Section 6.5. The generating functional of n-particle wave functions, which is an element of Fock's space, is a general Grassmann holomorphic function of $\varphi_\pm$.

The particle number operators for both particles commute with the hamiltonian:

$$\mathbf{N}_\pm = (2\pi)^{d-1}\int \mathrm{d}^{d-1}\hat{p}\,\varphi_\pm(\hat{p})\frac{\delta}{\delta\varphi_\pm(\hat{p})},\quad [\mathbf{N}_\pm,\mathbf{H}_0]=0\,,$$

a property which no longer holds in general for a local interacting theory.

S-matrix. The expression of the S-matrix in an interacting theory then follows from a simple extension of the method explained in the scalar case. One has only to be careful of the signs. One verifies that the S-matrix is given by

$$S(\varphi,\bar{\varphi}) = \int [\mathrm{d}\bar{\psi}\mathrm{d}\psi]\exp i\mathcal{A}(\bar{\psi}\sqrt{Z}+\bar{\psi}_c,\psi\sqrt{Z}+\psi_c), \tag{8.10}$$

that is, again a functional 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 (8.7),(8.8):

$$\tilde{\psi}_c(\hat{p}) = C^*\varphi_-(\hat{p})+\varphi_+(\hat{p}),\quad \tilde{\psi}_c^\dagger(\hat{p}) = \varphi_-(\hat{p})+\bar{\varphi}_+(\hat{p})\,{}^T C\,,\quad \bar{\psi}_c = -\psi_c^\dagger\gamma_0\,,$$

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

$$\begin{aligned}\int [\mathrm{d}\bar{\varphi}'(\hat{p})\mathrm{d}\varphi'(\hat{p})]S^*(\varphi',\bar{\varphi})S(\varphi',\bar{\varphi})\exp\left[(2\pi)^{d-1}\int\frac{\mathrm{d}\hat{p}}{2\omega(\hat{p})}\sum_{\varepsilon=\pm}\bar{\varphi}'_\varepsilon(\hat{p})\varphi'_\varepsilon(\hat{p})\right]\\ =\exp\left[(2\pi)^{d-1}\int\frac{\mathrm{d}\hat{p}}{2\omega(\hat{p})}\sum_{\varepsilon=\pm}\varphi_\varepsilon(\hat{p})\bar{\varphi}_\varepsilon(\hat{p})\right].\end{aligned} \tag{8.11}$$

8.1.3 Non-relativistic limit

To establish a connection with the non-relativistic quantum theory we now investigate the low energy low momentum of a theory with massive fermions, in analogy with the discussion for scalar bosons of Section 6.11. Since a fermion mass breaks space reflection symmetry in odd dimensions, we restrict ourselves to even dimensions in what follows.

As an example, we derive the non-relativistic limit of the action of a self-interacting fermion field

$$\mathcal{A}(\bar{\psi},\psi) = \int \mathrm{d}t\,\mathrm{d}x\left[\bar{\psi}\left(\frac{1}{i}\gamma_0\partial_t+\boldsymbol{\gamma}\cdot\nabla_x+m\right)\psi+\frac{1}{2}G\left(\bar{\psi}\psi\right)^2\right], \tag{8.12}$$

which we express in terms of $\psi,\psi^\dagger$ (equation(8.3)):

$$\mathcal{A}(\psi,\psi^\dagger) = \int \mathrm{d}t\,\mathrm{d}x\left[\psi^\dagger(t,x)\left[i\partial_t-\gamma_0(\boldsymbol{\gamma}\cdot\nabla_x+m)\right]\psi(t,x)+\tfrac{1}{2}G\left(\psi^\dagger\gamma_0\psi\right)^2\right].$$

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, we can use the transformations which lead to the action (8.9). In the kinematic part we then expand the one-particle energy

$$\omega(\hat{p}) = \sqrt{\hat{p}^2 + m^2} = m + \hat{p}^2/2m + O(m^{-3}).$$

In the interaction terms we take the non-relativistic limit, neglecting all momentum dependences relative to the mass. The projectors $P_\pm$ defined by equation (8.5) then reduce to

$$P_\pm = \tfrac{1}{2}(1 \pm \gamma_0).$$

The transformation between fields ψ and and $\varphi_\pm$ becomes local. Choosing a different normalization we find

$$\begin{aligned} \varphi_-(t,x) &= P_+ C^\dagger \psi^*(t,x), & \varphi_+(t,x) &= P_+\psi(t,x), \\ \varphi_+^\dagger(t,x) &= \psi^\dagger(t,x)P_+\,, & \varphi_-^\dagger(t,x) &= {}^T\psi(t,x)CP_+\,, \end{aligned}$$

or, conversely,

$$\psi(t,x) = \varphi_+(t,x) + C^*\varphi_-^*(t,x).$$

For γ matrices we choose a basis in which γ_0 is diagonal and restrict, below, the spinor indices to the non-vanishing components of $\varphi_\pm$. Using the relation

$$-\bar{\psi}(t,x)\psi(t,x) = \psi^\dagger(t,x)\gamma_0\psi(t,x) = \varphi_+^\dagger(t,x)\varphi_+(t,x) + \varphi_-^\dagger(t,x)\varphi_-(t,x),$$

we find the action

$$\mathcal{A}(\varphi_\pm, \varphi_\pm^\dagger) = \int \mathrm{d}t\,\mathrm{d}x \left[\sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t,x)\left[i\partial_t + m - \nabla_x^2/2m)\right]\varphi_\epsilon(t,x) + \frac{1}{2}G\left(\sum_{\epsilon=\pm}\varphi_\epsilon^\dagger\varphi_\epsilon\right)^2\right].$$

One then proceeds in analogy with the bosonic case. One translates the one-particle energy by the mass m, setting

$$\varphi_\pm(t,x) \mapsto \mathrm{e}^{imt}\,\varphi_\pm(t,x), \quad \varphi_\pm^\dagger(t,x) \mapsto \mathrm{e}^{-imt}\,\varphi_\pm^\dagger(t,x).$$

One then neglects in the interaction all terms that depend explicitly on time. Here, one finds

$$\mathcal{A}(\varphi_\pm, \varphi_\pm^\dagger) = \int \mathrm{d}t\,\mathrm{d}x \left[\sum_{\epsilon=\pm} \varphi_\epsilon^\dagger(t,x)\left[i\partial_t - \nabla_x^2/2m)\right]\varphi_\epsilon(t,x) + \frac{1}{2}G\left(\sum_{\epsilon=\pm}\varphi_\epsilon^\dagger\varphi_\epsilon\right)^2\right].$$

This action describes a many-body theory of two fermions of the same mass and with spin, the spin playing the role of an external quantum number decoupled from space-time.

Borrowing the result of Section 6.5, and comparing with the action (5.114), we find the non-relativistic hamiltonian $\mathbf{H}$, up to an infinite energy shift. One again verifies that the non-relativistic theory conserves the number of particles and, therefore, sectors with different particle number decouple:

$$\mathbf{N}_{\pm\alpha} = \int \mathrm{d}x\,\varphi_{\pm\alpha}(x)\frac{\delta}{\delta\varphi_{\pm\alpha}(x)} \quad \Rightarrow \quad [\mathbf{N}_{\pm\alpha}, \mathbf{H}] = 0\,,$$

(no summation over the spinor index α and the particle index $\pm$). A general n-particle contribution to the generating functional of wave functions can be written as

$$\Phi(\varphi) = \frac{1}{n!}\int \mathrm{d}x_1 \ldots \mathrm{d}x_n \, \phi_{\varepsilon_1\alpha_1,\varepsilon_2\alpha_2,\ldots,\varepsilon_n\alpha_n}(x_1,\ldots,x_n)\varphi_{\varepsilon_1\alpha_1}(x_1)\varphi_{\varepsilon_2\alpha_2}(x_2)\ldots\varphi_{\varepsilon_n\alpha_n}(x_n),$$

where $\phi_{\varepsilon_1\alpha_1,\varepsilon_2\alpha_2,\ldots,\varepsilon_n\alpha_n}(x_1,\ldots,x_n)$ is a totally antisymmetric wave function.

In the n-particle sector, the hamiltonian reads

$$H_n = -\frac{1}{2m}\sum_{i=1}^{n}\nabla^2_{x_i} - G\sum_{i<j}\delta(x_i - x_j). \tag{8.13}$$

The fermions interact through a two-body $\delta(x)$ 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: in two dimensions the fermion theory we have considered here, is equivalent to the well-known massive Thirring model (see Chapter 32). 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.

8.2 Free Euclidean Relativistic Fermions

We now perform the analytic continuation to euclidean time. This will allow to discuss quantum statistics of relativistic fermions (see Chapter 38), and generally simplify perturbative calculations. We, therefore, first explore in more detail the symmetries of relativistic fermion actions, 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 determine the free action as well as the coupling 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 Appendix A8.

Note that some of these symmetries have a form somewhat different from what one is familiar with in real time quantum field theory. 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.

Generalized Dirac fermions. The free fermion action $\mathcal{S}_0(\bar\psi,\psi)$ for generalized massive Dirac fermions, continuation to imaginary time of the standard action for spinor fields (8.1), can be written as

$$\mathcal{S}_0(\bar\psi,\psi) = -\int \mathrm{d}^d x \, \bar\psi_\alpha(x)\left[(\not\partial)_{\alpha\beta} + m\delta_{\alpha\beta}\right]\psi_\beta(x), \tag{8.14}$$

where the fields $\bar\psi_\alpha(x), \psi_\alpha(x)$ are also generators of a Grassmann algebra. In expression (8.14), we have introduced the traditional notation $\not\partial$ to represent the matrix $\partial_\mu\gamma_\mu$.

Chiral components. We show in Section A8.1.6 that in even dimensions d the spinor representation can be reduced. We can thus define chiral components $\psi_\pm$ of the fermion field

$$\psi_\pm(x) = \tfrac{1}{2}(1 \pm \gamma_S)\psi(x)\,. \tag{8.15}$$

and correspondingly $\bar\psi_\pm(x)$:

$$\bar\psi_\pm(x) = \bar\psi(x)\tfrac{1}{2}\,(1 \pm \gamma_S)\,, \tag{8.16}$$

often denoted by $\psi_{\mathrm{R}}(x)$, $\psi_{\mathrm{L}}(x)$, $\bar\psi_{\mathrm{R}}(x)$, $\bar\psi_{\mathrm{L}}(x)$ 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:

$$\mathcal{S}_0(\bar\psi_-, \psi_+) = -\int \mathrm{d}^d x\, \bar\psi_-(x) \not\partial \psi_+(x), \tag{8.17}$$

because $\bar\psi_- \psi_+ = \bar\psi_+ \psi_- = 0$. To construct an action for a massive propagating fermion the four spinors are required.

8.2.1 Hermitian conjugation

According to the discussion of Section 5.3, hermiticity of the hamiltonian is equivalent to hermiticity of the euclidean action 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 (8.3)) $\psi^\dagger = -\bar\psi\gamma_0$, where γ_0 is the γ matrix associated with the time component.

When one combines these two transformations, one verifies that they can be realized differently, in a way which no longer singles out the time variable. One defines $\bar\psi$ now as the hermitian conjugate of ψ, instead of $\psi^\dagger$, and after hermitian conjugation perform the transformation

$$\psi(\mathbf{x}) \mapsto \gamma_\mu \psi(\tilde{\mathbf{x}}), \quad \bar\psi(\mathbf{x}) \mapsto \bar\psi(\tilde{\mathbf{x}})\gamma_\mu\,, \quad \text{with } \tilde{\mathbf{x}} = P_\mu \mathbf{x}\,, \tag{8.18}$$

where P_μ is the space reflection along the μ axis: acting on a space vector $\mathbf{x}$, P_μ changes the sign of its component μ:

$$P_\mu \mathbf{x} = \tilde{\mathbf{x}} \qquad \text{with} \qquad \tilde{\mathbf{x}} : \begin{cases} \tilde{x}_\mu = -x_\mu, \\ \tilde{x}_\lambda = x \text{ for } \lambda \neq \mu\,. \end{cases} \tag{8.19}$$

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 (8.14). After hermitian conjugation we find

$$\mathcal{S}_0^\dagger(\bar\psi, \psi) = -\int \mathrm{d}^d x\, \bar\psi(x)\,(-\not\partial + m)\,\psi(x),$$

because ∂_μ is anti-hermitian. In the transformation (8.18) 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 integral over ψ and $\bar\psi$ in the functional integral is thus real. Eigenvalues of the operator $\not\partial + m$ are real or appear as complex conjugate pairs.

The action (8.17) has also reflection hermiticity as a symmetry. Indeed,

$$\mathcal{S}_0^\dagger(\bar\psi_-,\psi_+) = \int \mathrm{d}^d x\, \bar\psi_+(x)\not\partial\psi_-(x),$$

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.

8.2.2 Spin group and reflections

We first show that both actions (8.14) and (8.17) are invariant under the transformations of the spin group. We transform the spinors ψ and $\bar\psi$ as

$$\begin{aligned} (\psi_\Lambda)_\alpha(x) &= \Lambda^\dagger_{\alpha\beta}\psi_\beta(\mathbf{R}x)\,, \\ (\bar\psi_\Lambda)_\alpha(x) &= \bar\psi_\beta(\mathbf{R}x)\,\Lambda_{\beta\alpha}, \end{aligned} \tag{8.20}$$

where Λ belong to the spin group Spin(d) (equation ($A8.32$)) and the matrix $\mathbf{R}(\Lambda)$ is the corresponding element of $SO(d)$ (equation ($A8.33$)). After the change of variables $\mathbf{R}x \mapsto x'$ the invariance of the action follows from the identity

$$R_{\mu\nu}\Lambda\gamma_\nu\Lambda^\dagger = \gamma_\mu\,,$$

which is implied by equation ($A8.33$).

Reflections. In even dimensions a reflection along the $\mu = 1$ axis corresponds to the transformation Π_1 (Section A8.1.6):

$$\Pi_1: \quad \psi_{\Pi_1}(\mathbf{x}) = \hat\gamma\gamma_1\psi(\tilde{\mathbf{x}}), \qquad \bar\psi_{\Pi_1}(\mathbf{x}) = \bar\psi(\tilde{\mathbf{x}})\gamma_1\hat\gamma^\dagger \quad \text{with} \quad \tilde{\mathbf{x}} = (-x_1, x_2, \ldots, x_d). \tag{8.21}$$

The mass term in action (8.14) is clearly invariant. In the term $\not\partial$ the space reflection changes ∂_1 in $-\partial_1$, but then $\hat\gamma\gamma_1$ anticommutes with γ_1 and commutes with all other γ_μ matrices. The total action (8.14) is thus invariant.

The action (8.17), in contrast with the action (8.14), is not invariant under reflection, since reflection exchanges chiral components.

In odd dimensions total reflection $\tilde{\mathbf{x}} = -\mathbf{x}$ can be implemented by

$$\hat\Pi: \quad \psi_{\hat\Pi}(\mathbf{x}) = \psi(\tilde{\mathbf{x}}), \quad \bar\psi_{\hat\Pi}(\mathbf{x}) = -\bar\psi(\tilde{\mathbf{x}}), \tag{8.22}$$

a transformation that does not belong to the spin group but commutes with it. We then note that the fermion mass term is not invariant under reflection: in odd dimensions a *mass term* for a spinor fermion *violates parity conservation.*

8.2.3 Charge conjugation, charge conservation and chiral symmetry

Charge conjugation. We introduce a unitary matrix C and transform spinors as:

$$\psi_\alpha(x) = \bar{\psi}'_\beta(x) C^\dagger_{\beta\alpha}, \qquad \bar{\psi}_\alpha(x) = -C_{\alpha\beta}\psi'_\beta(x). \tag{8.23}$$

As a function of the new fields ψ' and $\bar{\psi}'$, the action (8.14) now reads

$$\mathcal{S}_0(\bar{\psi}', \psi') = -\int \mathrm{d}^d x\, \bar{\psi}'(x)\left(-C^{\dagger\,T}\!\not{\partial} C + m\right)\psi'(x). \tag{8.24}$$

The action (8.14) is thus invariant if the matrix C satisfies

$$C^{\dagger\,T}\gamma_\mu C = -\gamma_\mu \,.$$

We recognize the definition of the charge conjugation matrix (A8.39).

We also consider the action (8.17). The transformation (8.23) leads to

$$\begin{aligned}\not{\partial}(1+\gamma_S) \mapsto -C^{\dagger\,T}(1+\gamma_S)\,^T\!\not{\partial}\, C &= -C^\dagger(1+\gamma_S)\,^T\!\not{\partial}\, C = -C^{\dagger\,T}\!\not{\partial}(1-\gamma_S)C \\ &= \not{\partial} C^\dagger(1-\gamma_S)C \,.\end{aligned}$$

(The matrix γ_S is symmetric.) We have shown in Section A8.1.7 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 (8.17) is invariant. If the dimension is a multiple of four then,

$$\not{\partial}(1+\gamma_S) \mapsto \not{\partial}\,(1-\gamma_S)\,,$$

instead, and charge conjugation is not a symmetry. However, charge conjugation multiplied by space reflection, which exchanges chiral components, is a symmetry.

Finally, to justify 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

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^d x\, \bar{\psi}(x)\left(\not{\partial} + m + ie\not{A}\right)\psi(x).$$

After charge conjugation, as a consequence of equation (A8.39) the sign of the charge e has changed.

Odd dimensions. In odd dimensions if the action is reflection symmetric the mass term is absent. Charge conjugation can then be implemented by the matrix $\tilde{C}$ (definition (A8.41)) that satisfies

$$\tilde{C}^{\dagger\,T}\gamma_\mu \tilde{C} = \gamma_\mu \,,$$

and the transformations

$$\psi_\alpha(x) = \bar{\psi}'_\beta(x)\tilde{C}^\dagger_{\beta\alpha}, \qquad \bar{\psi}_\alpha(x) = \tilde{C}_{\alpha\beta}\psi'_\beta(x). \tag{8.25}$$

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 $\bar{\psi} = C\psi$. They correspond to neutral fermion fields, and are called Majorana spinors.

Fermion number conservation. If we assign a fermion number $+1$ to ψ and -1 to $\bar{\psi}$ we see that the action (8.14) conserves fermion number. To fermion number conservation corresponds a $U(1)$ invariance of the action:

$$\psi_\theta(x) = \mathrm{e}^{i\theta}\,\psi(x), \qquad \bar{\psi}_\theta(x) = \mathrm{e}^{-i\theta}\,\bar{\psi}(x). \tag{8.26}$$

For charged fermions the fermion number is proportional to the electric charge.
Majorana spinors violate fermion number conservation.

Chiral symmetry. In even dimensions, the free fermion action (8.14) possesses in the massless limit $m = 0$ an important additional $U(1)$ symmetry called *chiral symmetry*:

$$\psi_\theta(x) = \mathrm{e}^{i\theta\gamma_S}\,\psi(x), \qquad \bar{\psi}_\theta(x) = \bar{\psi}(x)\,\mathrm{e}^{i\theta\gamma_S}\ . \tag{8.27}$$

Such a symmetry has deep consequences and shows 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.

8.3 Partition Function. Correlations

In Section 5.6 we have derived an expression for the statistical operator of a system of non-relativistic fermions in the form of a functional integral. Here we generalize the expression to a relativistic quantum field theory, using the formalism of euclidean fermions introduced in Section 8.2. 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 5.3 and 8.2 we infer that the partition function for self-interacting massive Dirac fermions is given by an integral over Grassmann fields of the form

$$\mathcal{Z}(\beta) = \int [\mathrm{d}\psi(t,x)\mathrm{d}\bar{\psi}(t,x)] \exp\left[-\mathcal{S}(\bar{\psi},\psi) + \mu\int_0^\beta \mathrm{d}t \int \mathrm{d}^{d-1}x\,\bar{\psi}(t,x)\gamma_d\psi(t,x)\right] \tag{8.28}$$

with

$$\mathcal{S}(\bar{\psi},\psi) = -\int_0^\beta \mathrm{d}t \int \mathrm{d}^{d-1}x\left[\bar{\psi}(t,x)(\not{\partial} + m)\psi(t,x) + V\big(\bar{\psi}(t,x),\psi(t,x)\big)\right], \tag{8.29}$$

where t is the euclidean time, β the inverse temperature and μ the chemical potential. Fermion fields satisfy anti-periodic boundary conditions in the time direction:

$$\psi(t \equiv x_d = 0, x) = -\psi(\beta, x), \quad \bar{\psi}(0,x) = -\bar{\psi}(\beta,x).$$

Note that the term coupled to the chemical potential μ is hermitian, has the correct non-relativistic limit, is proportional to the conserved fermion charge (Section A13.1) and thus corresponds to a quantum operator which commutes with the hamiltonian.

In what follows we work for simplicity at zero temperature and in zero chemical potential, the finite temperature field theory being discussed in Chapter 38. In this limit, the boundary conditions play no role, and we no longer distinguish between space and time; x denotes all d coordinates.

We now introduce Grassmann sources $\bar{\eta}, \eta$ and consider a more general functional integral,

$$\mathcal{Z}(\bar{\eta},\eta) = \int [\mathrm{d}\psi(x)\mathrm{d}\bar{\psi}(x)] \exp\left[-\mathcal{S}(\bar{\psi},\psi) + \int \mathrm{d}^d x\, \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right], \tag{8.30}$$

where $\mathcal{Z}(\bar{\eta},\eta)$ is the generating functional of ψ, $\bar{\psi}$ field correlation functions since

$$\prod_{i=1}^{n} \frac{\delta}{\delta\eta(x_i)} \prod_{j=1}^{n} \frac{\delta}{\delta\bar{\eta}(y_j)} \mathcal{Z}(\eta,\bar{\eta})\bigg|_{\eta=\bar{\eta}=0} = \mathcal{Z}(0)\left[(-1)^n \left\langle \bar{\psi}(x_1)\ldots\bar{\psi}(x_n)\psi(y_1)\ldots\psi(y_n)\right\rangle\right].$$

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 a gaussian integral with external sources:

$$\mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta) = \int [\mathrm{d}\psi(x)\mathrm{d}\bar{\psi}(x)] \exp\left[-\mathcal{S}_0(\bar{\psi},\psi) + \int \mathrm{d}^d x \left(\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right)\right], \tag{8.31}$$

where $\mathcal{S}_0(\bar{\psi},\psi)$ is the free action (8.14):

$$\mathcal{S}_0(\bar{\psi},\psi) = -\int \mathrm{d}^d x\, \bar{\psi}(x)\left(\not{\partial} + m\right)\psi(x). \tag{8.32}$$

As usual we shift variables $\psi \mapsto \psi'$ to eliminate linear terms:

$$\psi(x) + \left(\not{\partial} + m\right)^{-1}\eta(x) = \psi'(x), \quad \bar{\psi}(x) + \bar{\eta}(x)\left(\not{\partial} + m\right)^{-1} = \bar{\psi}'(x). \tag{8.33}$$

Normalizing the functional integral (8.31) by $\mathcal{Z}(0,0) = 1$, we obtain

$$\mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta) = \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}^d y\, \bar{\eta}(y)\Delta_{\mathrm{F}}(y,x)\eta(x)\right], \tag{8.34}$$

in which the fermion propagator Δ_{F} is given by

$$\Delta_{\mathrm{F}}(y,x) = \frac{1}{(2\pi)^d} \int \mathrm{d}^d p\, \mathrm{e}^{-ip(x-y)} \frac{(m - i\not{p})}{p^2 + m^2}. \tag{8.35}$$

One verifies that on mass-shell ($p^2 = -m^2$), $m - i\not{p}$ 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 Spin$(d-1)$, the subgroup of Spin(d) which leaves the momentum p invariant.

The fermion two-point correlation function in a free or gaussian theory, then, is

$$\left\langle \bar{\psi}_\alpha(x)\psi_\beta(y)\right\rangle_0 = -\frac{\delta}{\eta_\alpha(x)} \frac{\delta}{\bar{\eta}_\beta(y)} \mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta) = \left(\Delta_{\mathrm{F}}\right)_{\beta\alpha}(y,x). \tag{8.36}$$

Generalization of equation (1.79) to the action (8.29) then yields the identity

$$\mathcal{Z}(\bar{\eta},\eta) = \exp\left[\int \mathrm{d}^d x\, V\left(-\frac{\delta}{\delta\eta(x)}, \frac{\delta}{\delta\bar{\eta}(x)}\right)\right] \mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta). \tag{8.37}$$

This identity leads to the perturbative expansion of a field theory with self-interacting fermions. Alternatively, the functional integral (8.30) can be expanded in powers of $V, \eta, \bar{\eta}$ and all terms evaluated with the corresponding Wick's theorem for fermion fields, a simple generalization of the form (1.77), which can be written in terms of free field expectation values as

$$\left\langle \prod_{i=1,n} \bar{\psi}_{\alpha_i}(x_i)\psi_{\beta_i}(y_i) \right\rangle_0 = \sum_{\substack{\text{permutations} \\ P \text{ of } \{1,2,\ldots,n\}}} \epsilon(P) \prod_{i=1,n} \left\langle \bar{\psi}_{\alpha_{P(i)}}(x_{P(i)})\psi_{\beta_i}(y_i)\right\rangle_0 . \tag{8.38}$$

Theories with bosons and fermions. Many field theories involve both fermions and bosons. A simple example which will be studied later is

$$\mathcal{Z}(\bar{\eta},\eta,J) = \int [\mathrm{d}\psi\, \mathrm{d}\bar{\psi}\, \mathrm{d}\phi] \exp\Big\{-\mathcal{S}(\bar{\psi},\psi,\phi) + \int \mathrm{d}^d x \left[\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x) + J(x)\phi(x)\right]\Big\}, \tag{8.39}$$

in which ψ, $\bar{\psi}$, η, $\bar{\eta}$ are Grassmann fields and ϕ and J usual real fields, and in which the action $\mathcal{S}(\bar{\psi},\psi,\phi)$ has the form

$$\mathcal{S} = \int \mathrm{d}^d x \left\{-\bar{\psi}(x)\left[\not{\partial} + M + g\phi(x)\right]\psi(x) + \tfrac{1}{2}\left(\partial_\mu\phi(x)\right)^2 + \tfrac{1}{2}m^2\phi^2(x) + \tfrac{1}{24}\lambda\phi^4(x)\right\}. \tag{8.40}$$

The functional $\mathcal{Z}(\bar{\eta},\eta,J)$ in equation (8.39) generates both ϕ field and ψ, $\bar{\psi}$ field correlation functions.

Perturbation theory is generated by

$$\mathcal{Z}(\bar{\eta},\eta,J) = \exp\left[-\int \mathrm{d}^d x \left(\frac{\lambda}{24}\left(\frac{\delta}{\delta J(x)}\right)^4 + g\frac{\delta}{\delta J(x)}\frac{\delta}{\delta\eta(x)}\frac{\delta}{\delta\bar{\eta}(x)}\right)\right] \mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta,J), \tag{8.41}$$

in which $\mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta,J)$ is the product of free fermion and free boson functionals:

$$\mathcal{Z}_{\mathrm{G}}(\bar{\eta},\eta,J) = \exp\left[\int \mathrm{d}^d x\, \mathrm{d}^d y \left(\tfrac{1}{2}J(x)\Delta(x,y)J(y) - \bar{\eta}(x)\Delta_{\mathrm{F}}(x,y)\eta(y)\right)\right]. \tag{8.42}$$

All algebraic transformations we have performed on expression (7.22) can easily be generalized to the representation (8.41), 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 fermions.

In the example (8.40), the integral over fermions is gaussian and can also be performed explicitly. This leads to a scalar field theory with additional non-local interactions:

$$\int [\mathrm{d}\psi\, \mathrm{d}\bar{\psi}] \exp \int \mathrm{d}^d x \left[\bar{\psi}(x)\left(\not{\partial} + M + g\phi(x)\right)\psi(x) + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right] \propto \exp\left[-\mathcal{S}_{\mathrm{F}}(\phi,\eta,\bar{\eta})\right] \tag{8.43}$$

with

$$\mathcal{S}_{\mathrm{F}} = -\operatorname{tr}\ln\left[\not{\partial} + M + g\phi(x)\right] + \int \mathrm{d}^d x \, \mathrm{d}^d y \, \bar{\eta}(y) \left[\not{\partial} + M + g\phi(\cdot)\right]^{-1}(y,x)\eta(x) \,. \tag{8.44}$$

The expansion of $\mathcal{S}_{\mathrm{F}}(\phi, 0, 0)$ in powers of ϕ, generates a set of one fermion loop Feynman diagrams (see Section 7.7). A similar integral over boson fields would have generated a contribution of the form $+\operatorname{tr}\ln$. Hence, compared to boson loops, *fermion loops* are multiplied by an additional *minus sign.*

8.4 Generating Functionals

We have discussed connected functions and proper vertices only 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 8.3, we write the source terms in the functional integral $\bar{\eta}\psi + \bar{\psi}\eta$, then we define the Legendre transform of $\mathcal{W}$ by

$$\Gamma(\bar{\psi}, \psi) + \mathcal{W}(\eta, \bar{\eta}) = \int \mathrm{d}x \, \left[\bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right], \tag{8.45a}$$

$$\psi(x) = \frac{\delta \mathcal{W}}{\delta \bar{\eta}(x)}, \quad \bar{\psi}(x) = -\frac{\delta \mathcal{W}}{\delta \eta(x)} \,. \tag{8.45b}$$

The equations (8.45b) 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 easily verifies 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. We recall here, however, that gaussian integration over fermion fields yields a determinant instead of the inverse of a determinant for a complex scalar. This implies that fermion loops are affected by an additional minus sign compared to boson loops. Let us, as an example, give the one-loop results corresponding to the boson–fermion action:

$$\mathcal{S}(\phi, \bar{\psi}, \psi) = \int \mathrm{d}^d x \left[-\bar{\psi}(x)\left(\not{\partial} + A(\phi)\right)\psi(x) + \mathcal{S}_{\mathrm{B}}(\phi)\right]. \tag{8.46}$$

We have to evaluate the functional integral:

$$\mathcal{Z}(J, \bar{\eta}, \eta) = \int \left[\mathrm{d}\phi \, \mathrm{d}\bar{\psi} \mathrm{d}\psi\right] \exp\left[-\mathcal{S}(\phi, \bar{\psi}, \psi) + \int \mathrm{d}^d x \left[J(x)\phi(x) + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)\right]\right]. \tag{8.47}$$

We first again look 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 functional integral over boson and fermion fields. The result can be deduced 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, we find for connected correlation functions

$$\begin{aligned}\mathcal{W}(J,\bar\eta,\eta) &= -\mathcal{S}(\phi_c,\bar\psi_c,\psi_c) + \int \mathrm{d}^d x\,\left[J(x)\phi_c(x) + \bar\eta(x)\psi_c(x) + \bar\psi_c(x)\eta(x)\right] \\ &\quad - \frac{1}{2}\operatorname{tr}\ln\left[\frac{\delta^2\mathcal{S}_\mathrm{B}}{\delta\phi_c(x)\delta\phi_c(y)} - \bar\psi_c\frac{\delta^2 A}{\delta\phi_c(x)\delta\phi_c(y)}\psi_c + 2\bar\psi_c\frac{\delta A}{\delta\phi_c(x)}\left[\not{\partial} + A(\phi_c)\right]^{-1}\frac{\delta A}{\delta\phi_c(y)}\psi_c\right] \\ &\quad + \operatorname{tr}\ln\left[\not{\partial} + A(\phi_c)\right],\end{aligned}$$

where the fields $\phi_c, \bar\psi_c, \psi_c$ are solutions of the field equations

$$\frac{\delta\mathcal{S}_\mathrm{B}}{\delta\phi_c(x)} - \bar\psi_c\frac{\delta A}{\delta\phi_c(x)}\psi_c - J(x) = 0\,, \tag{8.48a}$$

$$\left[\not{\partial} + A(\phi_c)\right]\psi_c(x) + \eta(x) = 0\,, \tag{8.48b}$$

$$\bar\psi_c(x)\left[\not{\partial} + A(\phi_c)\right] + \bar\eta(x) = 0 \tag{8.48c}$$

(by convention the operator ∂_μ in (8.48*c*) acts on the left with a minus sign).

A Legendre transformation then yields the 1PI functional at one-loop order:

$$\begin{aligned}\Gamma(\varphi,\bar\psi,\psi) &= \mathcal{S}(\varphi,\bar\psi,\psi) - \operatorname{tr}\ln\left(\not{\partial} + A(\varphi)\right) \\ &\quad + \frac{1}{2}\operatorname{tr}\ln\left\{\frac{\delta^2\mathcal{S}_\mathrm{B}}{\delta\varphi(x)\delta\varphi(y)} + \bar\psi\left[2\frac{\delta A}{\delta\varphi(x)}\left[\not{\partial} + A(\varphi)\right]^{-1}\frac{\delta A}{\delta\varphi(y)} - \frac{\delta^2 A}{\delta\varphi(x)\delta\varphi(y)}\right]\psi\right\}.\end{aligned} \tag{8.49}$$

8.5 Connection between Spin and Statistics

We have noticed in Section 8.1 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 degree 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, by an explicit calculation.

We have shown in Section A2.1 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 6.9. All possible intermediate states contribute with the same sign. This result can easily be generalized to the discontinuity in the physical domain of diagonal scattering amplitudes. Let us then 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

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x\left[\tfrac{1}{2}\left(\partial_\mu\phi(x)\right)^2 + \tfrac{1}{2}m^2\phi^2(x) + \tfrac{1}{2}g\phi^2(x)\chi(x)\right].$$

The integration over ϕ yields a (non-local) contribution to the χ action ($\ln\det = \operatorname{tr}\ln$):

$$\delta\mathcal{S}(\chi) = \tfrac{1}{2}\operatorname{tr}\ln(-\nabla_x^2 + m^2 + g\chi).$$

If we expand this expression to order g^2 we find a term linear in χ which just shifts the χ field expectation value and a quadratic term which modifies the two-point function in the gaussian approximation. One verifies that the corresponding contribution to $\delta W^{(2)}_{\chi^2}$, the χ two-point function in Fourier variables, is

$$\delta W^{(2)}_{\chi^2} = \tfrac{1}{2} g^2 \Delta^2_\chi(p) B(p) + \text{const.}$$

with

$$B(p) = \frac{1}{(2\pi)^d} \int \frac{\mathrm{d}^d q}{(q^2+m^2)\left[(p+q)^2+m^2\right]} . \tag{8.50}$$

A convenient integral representation of the diagram is given at the end of Section 11.5 (equation (11.66)), which indeed shows that the function has a cut for $p^2 < -(2m)^2$, the region of physical ϕ particle emission, with a negative imaginary part,

$$B(p)|_{p^2+i\epsilon} - B(p)|_{p^2-i\epsilon} = -2ig^2 \frac{\pi^{(3-d)/2} 2^{4-2d}}{\Gamma[\frac{1}{2}(d-1)]} s^{-1/2}(s-4m^2)^{(d-3)/2} ,$$

a result which is consistent with the representation (6.90) for $\delta W^{(2)}_{\chi^2}$.

If we instead consider the contribution coming from scalar fermions, the tr ln is replaced by $-$ tr ln, and the contribution has the opposite sign, which violates hermiticity.

Spinor fermions. Let us instead calculate the contribution of spinor fermions. After gaussian integration the contribution to the χ action is now

$$\delta S(\chi) = -\operatorname{tr}\ln(\not\partial + m + g\chi),$$

which expanded to order g^2 yields a modification of the χ two-point function,

$$\begin{aligned} \delta W^{(2)}_{\chi^2}(p) &= -\frac{g^2}{(2\pi)^d} \operatorname{tr} \int \frac{\mathrm{d}^d q\,(-i\not q + m)(-i\not p - i\not q + m)}{(q^2+m^2)\left[(p+q)^2+m^2\right]} \\ &= -\frac{g^2}{(2\pi)^d} \operatorname{tr}\mathbf{1} \int \frac{\mathrm{d}^d q\,(m^2 - pq - p^2)}{(q^2+m^2)\left[(p+q)^2+m^2\right]} . \end{aligned}$$

We now use the identity

$$m^2 - pq - q^2 = 2m^2 + \tfrac{1}{2}p^2 - \tfrac{1}{2}\left[(p+q)^2 + m^2 + q^2 + m^2\right].$$

The two terms inside the brackets cancel a denominator and thus yield a constant (in general divergent, see Chapter 9) result, which has no discontinuity:

$$\delta W^{(2)}_{\chi^2}(p) = -g^2 \operatorname{tr}\mathbf{1}(p^2+4m^2)\Delta^2_\chi(p)B(p) + \text{ const.} . \tag{8.51}$$

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.

Bibliographical Notes

Most references of Chapter 7 are still relevant here. Two important references for this chapter again are

F.A. Berezin, *The Method of Second Quantization* (Academic Press, New York 1966);

L. D. Faddeev in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

Besides the work of C. Itzykson and J.B. Zuber, quoted in the introduction, another general reference here is

J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics*, (McGraw-Hill, New-York 1964).

The non-relativistic limit is considered in

L.L. Foldy and S.A. Wouthuysen, *Phys. Rev.* 78 (1950) 29.

For a discussion of the *PCT* theorem and the spin-statistics connection, starting from first principles see

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

The appendix A8 contains an elementary introduction to spinors and spin group. For details see

Y. Choquet-Bruhat and 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 and C. Itzykson eds. (Gordon and Breach, New York 1973).

APPENDIX A8
EUCLIDEAN FERMIONS, SPIN GROUP AND γ MATRICES

This appendix assumes a minimal familiarity with Dirac's equation and the action of the free fermion field. These questions are well discussed in standard textbooks of Particle Physics (see also Section 8.1). We want to, briefly, describe here the formalism of euclidean fermions, analytic continuation to imaginary time of relativistic fermions with spin. As we have noticed in Chapter 7, in this continuation 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 Spin(d), a group which 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. We finally discuss the calculation of traces of products of γ-matrices, quantities that appear in perturbation theory with fermions, and define the Fierz transformation.

A8.1 Spin Group. Dirac γ Matrices

In this section, we begin with some 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.

A8.1.1 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

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\,\delta_{\mu\nu}\mathbf{1}\,, \tag{A8.1}$$

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 1.5, acting on Grassmann algebras (see equation (1.49)).

It follows from the relations ($A8.1$) 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 $\mu_1 < \mu_2 < \cdots < \mu_p$.

Automorphism. As in the example of Grassmann algebras we define an automorphism P in $\mathcal{C}(d)$ by

$$\mathrm{P}(\gamma_\mu) = -\gamma_\mu\,. \tag{A8.2}$$

It splits $\mathcal{C}(d)$ into two vector spaces, $\mathcal{C}_-(d)$ and $\mathcal{C}_+(d)$, containing odd and even elements, respectively,

$$\mathrm{P}(\mathcal{C}_\pm) = \pm\mathcal{C}_\pm\,,$$

where only $\mathcal{C}_+(d)$ is a subalgebra.

Clifford algebra and orthogonal group. Let us perform a linear transformation acting on the generators $\gamma \mapsto \gamma'$,

$$\gamma'_\mu = R_{\mu\nu}\gamma_\nu\,, \quad \det \mathbf{R} \neq 0\,, \tag{A8.3}$$

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 (A8.1),

$$\gamma'_\mu\gamma'_\nu + \gamma'_\nu\gamma'_\mu = R_{\mu\rho}R_{\nu\sigma}\left(\gamma_\rho\gamma_\sigma + \gamma_\sigma\gamma_\rho\right) = 2R_{\mu\rho}R_{\nu\rho}\mathbf{1} = 2\delta_{\mu\nu}\mathbf{1}\,.$$

Therefore, the matrix $\mathbf{R}$ must be orthogonal

$$R_{\mu\rho}R_{\nu\rho} = \delta_{\mu\nu}\,. \tag{A8.4}$$

The relations (A8.1) are invariant under the orthogonal group $O(d)$ (rotations–reflections in the euclidean d-dimensional space).

Remark. If the tensor $\delta_{\mu\nu}$ is replaced in the r.h.s. of equation (A8.1) 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 $+1, -1, \ldots, -1$ one obtains the relativistic group $O(1, d-1)$, generalization of the Lorentz group $O(1,3)$.

Product of all generators. In $C(d)$ one element plays a special role, the product of all generators. We thus define

$$\hat\gamma = \gamma_1\gamma_2\ldots\gamma_d\,. \tag{A8.5}$$

Then,

$$\hat\gamma^2 = (-1)^{[d/2]}\mathbf{1}\,, \tag{A8.6}$$

where $[d/2]$ is the integer part of $d/2$.

An orthogonal transformation of matrix $\mathbf{R}$ acting on the generators transforms $\hat\gamma$ into $\hat\gamma \det \mathbf{R}$.

If d is even, $\hat\gamma$ commutes with the elements of C_+ and anticommutes with those of C_-, that is, for all elements c of $C(d)$,

$$\hat\gamma c = \mathrm{P}(c)\hat\gamma\,. \tag{A8.7}$$

If d is odd, $\hat\gamma$ instead commutes with all elements c of $C(d)$,

$$\hat\gamma c = c\hat\gamma\,. \tag{A8.8}$$

Centre of the algebra. One actually verifies that when d is even the centre of the algebra, that is, he set of elements which commute with $C(d)$, contains only $r\mathbf{1}$, $r \in \mathbb{R}$. When d is odd, the elements which commute with $C(d)$ are linear combinations of $\mathbf{1}$ and $\hat\gamma$.

Dimension d odd. Since $\hat\gamma$ commutes with all elements of $C(d)$ we can construct an algebra homomorphic to $C(d)$ by imposing to the generators, in addition to the relations (A8.1), the relation

$$\hat\gamma = i^{(d-1)/2}\mathbf{1}, \tag{A8.9}$$

consistent with the identity (A8.6). This has several consequences:

(i) The group of symmetry of the relations (A8.1) and (A8.9) is reduced to orthogonal matrices of determinant one, thus belonging to the group $SO(d)$ (rotations) and the transformation P (A8.2) is no longer an automorphism.

(ii) If $d = 1 \pmod 4$ then $\hat\gamma = \pm\mathbf{1}$ and the algebra is isomorphic to $C(d-1)$ (a real vector space of dimension 2^{d-1}). If $d = 3 \pmod 4$ then $\hat\gamma = \pm i\mathbf{1}$ and the algebra is isomorphic to the complexified form of $C(d-1)$ (still a vector space of dimension 2^d on $\mathbb{R}$).

A8.1.2 Clifford algebra and group structure

We now consider the group $\mathfrak{G}(d)$ of invertible elements Λ of $C(d)$ that satisfy

$$\Lambda^{-1}\gamma_\mu \Lambda = R_{\mu\nu}\gamma_\nu \,, \tag{A8.10}$$

where $\mathbf{R}$ of coefficients $R_{\mu\nu}$ is a real matrix. This defining relation induces a homomorphism of groups. Indeed, if

$$\Lambda_1 \mapsto \mathbf{R}_1 \,, \qquad \Lambda_2 \mapsto \mathbf{R}_2 \,,$$

to the product $\Lambda_1\Lambda_2$ corresponds the product of real matrices $\mathbf{R}_1\mathbf{R}_2$.

Moreover, the relation ($A8.10$) implies

$$\Lambda^{-1}\gamma_\mu\gamma_\nu\Lambda = R_{\mu\rho}R_{\nu\sigma}\gamma_\rho\gamma_\sigma \,.$$

Adding the symmetric relation $\mu \leftrightarrow \nu$, and using the relations ($A8.1$), we obtain

$$R_{\mu\rho}R_{\nu\rho} = \delta_{\mu\nu} \,.$$

The real matrices $\mathbf{R}$ thus form a group, subgroup of the orthogonal group $O(d)$ (rotation–reflection) of transformations ($A8.3$).

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

$$\Lambda^{-1}\hat\gamma\Lambda = \det\mathbf{R}\,\hat\gamma \;\Leftrightarrow\; \hat\gamma\Lambda = \det\mathbf{R}\,\Lambda\hat\gamma \,. \tag{A8.11}$$

Comparing with the two properties ($A8.7$,$A8.8$) 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 $\hat\gamma$ commutes with all other elements of $C(d)$ and thus the relation ($A8.11$) 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 $\mathbf{R}^\rho$:

$$R^\rho_{\mu\nu} = 2\delta_{\rho\mu}\delta_{\rho\nu} - \delta_{\mu\nu} \,, \tag{A8.12}$$

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.

(vi) The element $\hat\gamma$ also belongs to the group $\mathfrak{G}(d)$ and for d even corresponds to the rotation matrix $-\mathbf{1}$ since from ($A8.7$)

$$(\hat\gamma)^{-1}\gamma_\mu\hat\gamma = -\gamma_\mu \,. \tag{A8.13}$$

Finally, since γ_μ is associated with the orthogonal matrix ($A8.12$), the product Π_μ,

$$\Pi_\mu = \hat\gamma\gamma_\mu \,, \tag{A8.14}$$

corresponds to a reflection P_μ along the μ axis, $x_\mu \mapsto -x_\mu$ (as defined by (8.19)). Note that

$$P_\mu^2 = \mathbf{1} \quad \text{but} \quad (\Pi_\mu)^2 = (-1)^{d/2+1}\mathbf{1} \,.$$

A8.1.3 Spin group and Lie algebra

We now show by explicit construction that the whole group $SO(d)$ can be generated by a subgroup of $\mathfrak{G}(d)$, the spin group 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 Spin(d) are Lie groups and we can discuss their Lie algebras.

We consider the elements

$$\tilde{\boldsymbol{\sigma}}_{\mu\nu} = \tfrac{1}{2}[\gamma_\mu, \gamma_\nu]. \tag{A8.15}$$

Only $d(d-1)/2$ elements $\tilde{\boldsymbol{\sigma}}_{\mu\nu}$ are linearly independent; a basis is, for example, $\tilde{\boldsymbol{\sigma}}_{\lambda\mu}$ with $\mu < \nu$. Using the relations ($A8.1$) one verifies that for $\mu \neq \nu$ (no summation on μ, ν being implied)

$$\tilde{\boldsymbol{\sigma}}^2_{\mu\nu} = (\gamma_\mu \gamma_\nu)^2 = -\mathbf{1}\,.$$

To $\tilde{\boldsymbol{\sigma}}_{\mu\nu}$ $(\mu \neq \nu)$ we associate the elements Λ:

$$\begin{aligned}\Lambda(\theta) &= \exp[-\tfrac{1}{2}\theta\tilde{\boldsymbol{\sigma}}_{\mu\nu}]\,, \quad \theta \in \mathbb{R} \\ &= \cos(\theta/2)\,\mathbf{1} - \sin(\theta/2)\,\tilde{\boldsymbol{\sigma}}_{\mu\nu}\,.\end{aligned} \tag{A8.16}$$

The elements $\Lambda(\theta)$ for a given matrix $\tilde{\boldsymbol{\sigma}}_{\mu\nu}$ 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 in the (μ, ν) plane.

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 $\mathbf{T}^{\rho\sigma}$:

$$\left(\mathbf{T}^{\rho\sigma}\right)_{\alpha\beta} = \delta_{\rho\alpha}\delta_{\sigma\beta} - \delta_{\rho\beta}\delta_{\sigma\alpha}\,, \tag{A8.17}$$

where only $d(d-1)/2$ are independent. If $\theta_{\rho\sigma}$ is an arbitrary antisymmetric matrix, with this normalization

$$\tfrac{1}{2}\left(\mathbf{T}^{\rho\sigma}\right)_{\alpha\beta}\theta_{\rho\sigma} = \theta_{\alpha\beta}\,. \tag{A8.18}$$

A general matrix $\mathbf{R}$ of $SO(d)$ can thus be written as

$$\mathbf{R} = \mathrm{e}^{\boldsymbol{\theta}} = \exp\left[\tfrac{1}{2}\mathbf{T}^{\rho\sigma}\theta_{\rho\sigma}\right]. \tag{A8.19}$$

Spin group. It follows from the homomorphism between groups that the matrices $\tilde{\boldsymbol{\sigma}}_{\mu\nu}/2$ 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* Spin(d), subgroup of the group $\mathfrak{G}(d)$,

$$\Lambda = \exp\left[\tfrac{1}{4}\theta_{\mu\nu}\tilde{\boldsymbol{\sigma}}_{\mu\nu}\right]. \tag{A8.20}$$

The only non-trivial element of the centre that is contained in the spin group is $-\mathbf{1}$, which corresponds in ($A8.16$) 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 above, for d even the addition of one reflection $\Pi_\mu = \hat{\gamma}\gamma_\mu$ allows to generate the whole $O(d)$ group.

Therefore, the transformations ($A8.10$) implement the transformations ($A8.3$) provided in the odd case one restricts the Clifford algebra by the relation ($A8.9$).

A8.1.4 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 for the matrices representing the generators *the same notation* γ_μ, $\mu = 1, \dots, d$. We first deal with even-dimensional spaces and then generalize to odd-dimensional spaces, in the framework of the constraint ($A8.9$).

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 ($A8.1$).

For $d = 2$, the standard Pauli matrices realize the algebra $\mathcal{C}(2)$:

$$\gamma_1^{(d=2)} \equiv \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_2^{(d=2)} \equiv \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \qquad (A8.21)$$

The third Pauli matrix σ_3 is proportional to the matrix $\hat{\gamma}$ (equation ($A8.5$)). We define

$$\gamma_S^{(d=2)} \equiv \gamma_3^{(d=2)} = -i\hat{\gamma}^{(d=2)} = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (A8.22)$$

The three matrices are hermitian:

$$\gamma_i = \gamma_i^\dagger .$$

The matrices γ_1 and γ_3 are symmetric, and γ_2 is antisymmetric:

$$\gamma_1 = {}^T\gamma_1\,, \qquad \gamma_3 = {}^T\gamma_3\,, \qquad \gamma_2 = -\,{}^T\gamma_2\,,$$

where we have denoted by ${}^T\gamma$ the transpose of the matrix γ.

Quite generally we define

$$\gamma_S^{(d)} \equiv \gamma_{d+1}^{(d)} = i^{-d/2}\hat{\gamma}^{(d)}\,, \qquad (A8.23)$$

where $\hat{\gamma}$ is the product of all generators (definition ($A8.5$)). The matrix γ_S then satisfies

$$\gamma_S^2 = \mathbf{1}\,, \quad \gamma_S\gamma_\mu + \gamma_\mu\gamma_S = 0 \text{ for } \mu \le d\,. \qquad (A8.24)$$

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

$$\gamma_i^{(d+2)} = \sigma_1 \otimes \gamma_i^{(d)} = \begin{pmatrix} 0 & \gamma_i^{(d)} \\ \gamma_i^{(d)} & 0 \end{pmatrix}, \quad 1 \le i \le d+1\,, \qquad (A8.25)$$

$$\gamma_{d+2}^{(d+2)} = \sigma_2 \otimes \mathbf{1}_d = \begin{pmatrix} 0 & -i\mathbf{1}_d \\ i\mathbf{1}_d & 0 \end{pmatrix}, \qquad (A8.26)$$

in which $\mathbf{1}_d$ is the unit matrix in $2^{d/2}$ dimensions.

From the definition ($A8.23$), it follows that $\gamma_S^{(d+2)}$ is then given by

$$\gamma_S^{(d+2)} \equiv \gamma_{d+3}^{(d+2)} = \sigma_3 \otimes \mathbf{1}_d = \begin{pmatrix} \mathbf{1}_d & 0 \\ 0 & -\mathbf{1}_d \end{pmatrix}. \tag{A8.27}$$

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 ($A8.1$), the $\gamma_i^{(d+2)}$ matrices satisfy the same relations.

By inspection, we see that the γ matrices are all hermitian. In addition

$$^T\gamma_i^{(d+2)} = \begin{pmatrix} 0 & {}^T\gamma_i^{(d)} \\ {}^T\gamma_i^{(d)} & 0 \end{pmatrix}, \quad 1 \le i \le d+1 .$$

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 odd index are symmetric, all matrices with even index are antisymmetric:

$$^T\gamma_i = (-1)^{i+1}\gamma_i , \quad 1 \le i \le d+1 . \tag{A8.28}$$

Finally the relations ($A8.1$) and ($A8.24$) can be summarized by

$$\gamma_i\gamma_j + \gamma_j\gamma_i = 2\,\delta_{ij} , \qquad \text{for } i,j = 1,\ldots,d,d+1 . \tag{A8.29}$$

We will use greek letters $\mu\ \nu \ldots$ to indicate that we exclude the value $d+1$ for the index.

Space of odd dimensions d. The relation ($A8.29$) 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 ($A8.23$) the product of γ matrices $\hat{\gamma}$ satisfies the relation ($A8.9$),

$$\hat{\gamma} = \gamma_1\gamma_2 \ldots \gamma_d = i^{d/2}.$$

A8.1.5 Spin group: a unitary representation

When the generators of the Clifford algebra are represented by hermitian matrices the generators $\tilde{\boldsymbol{\sigma}}_{\mu\nu}$ of the spin group Spin(d) are represented by anti-hermitian, traceless, matrices (for which we use below the same notation). The complex vectors on which the representation acts are called *spinors.*

Instead of the anti-hermitian matrices $\tilde{\boldsymbol{\sigma}}_{\mu\nu}$ one usually defines hermitian matrices $\boldsymbol{\sigma}_{\mu\nu}$:

$$\boldsymbol{\sigma}_{\mu\nu} = \frac{1}{i}\tilde{\boldsymbol{\sigma}}_{\mu\nu} = \frac{1}{2i}\left[\gamma_\mu, \gamma_\nu\right]. \tag{A8.30}$$

Then for $\mu \neq \nu$ the matrices $\boldsymbol{\sigma}_{\mu\nu}$ have the property

$$\boldsymbol{\sigma}^2_{\mu\nu} = \mathbf{1}. \tag{A8.31}$$

The matrices belonging to the representation of the spin group Spin(d) can be written as

$$\Lambda = \exp\left(\tfrac{i}{4}\sigma_{\mu\nu}\theta_{\mu\nu}\right), \tag{A8.32}$$

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 one: the representation of the spin group Spin(d) is a unitary group, subgroup of the unitary group $SU(2^{[d/2]})$.

The relation (A8.10) can be written as

$$\Lambda\gamma_\mu\Lambda^\dagger = \gamma_\nu R_{\nu\mu}\,. \tag{A8.33}$$

Introducing a space vector p_μ, we can write the equation in an equivalent form:

$$\Lambda p_\mu\gamma_\mu\Lambda^\dagger = (p_R)_\mu\gamma_\mu \quad \text{with} \quad (p_R)_\mu = R_{\mu\nu}p_\nu\,. \tag{A8.34}$$

In this form the equation shows explicitly that the group $SO(d)$ is isomorphic to the adjoint representation of the spin group 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 (A8.33) in terms of Λ by taking a trace:

$$R_{\mu\nu} = \operatorname{tr}\left(\Lambda^\dagger\gamma_\mu\Lambda\gamma_\nu\right)/\operatorname{tr}\mathbf{1}\,.$$

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 $SU(2) \times SU(2)$.

A8.1.6 Reflections and chiral components

Reducibility. In even dimensions we infer from equation (A8.7) that γ_S, which is proportional to $\hat{\gamma}$, commutes with all transformations of the spin group:

$$[\Lambda, \gamma_S] = 0\,. \tag{A8.35}$$

Therefore, the unitary representation of Spin(d) is reducible. The projection is 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 $\psi_\pm$, the chiral components of the spinor ψ:

$$\psi_\pm = \tfrac{1}{2}(\mathbf{1} \pm \gamma_S)\psi\,. \tag{A8.36}$$

Space reflections and chiral components. To obtain the full orthogonal group we have still to represent reflections. We have seen that this can be achieved with elements of the Clifford algebra only in spaces of even dimensions. Then the elements $\pm\Pi_\mu$, ($\Pi_\mu = \hat{\gamma}\gamma_\mu$, see equation ($A$8.14)) correspond to reflections P_μ (equation (8.19)) that act on a position vector $\mathbf{x}$ by changing the sign of the μ component. The anticommutation relation,

$$\gamma_S\Pi_\mu = -\gamma_S\Pi_\mu\,, \tag{A8.37}$$

implies

$$\Pi_\mu\,\tfrac{1}{2}\,(1 + \gamma_S) = \tfrac{1}{2}\,(1 - \gamma_S)\,\Pi_\mu. \tag{A8.38}$$

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.

A8.1.7 Charge conjugation

We exhibit unitary matrices C such that

$$C^\dagger\, {}^T\gamma_\mu\, C = -\gamma_\mu \quad \Leftrightarrow \quad C\,\gamma_\mu\, C^\dagger = -\,{}^T\gamma_\mu\,. \qquad (A8.39)$$

Note that since the matrices γ_μ are hermitian ${}^T\gamma_\mu = \gamma_\mu^*$.

Even dimensions. In the representation of Section A8.1.4, for d even we can take

$$C = \pm\,(\hat\gamma)^{d/2} \prod_{\text{all }\mu\text{ odd}} \gamma_\mu\,, \quad \Rightarrow C^\dagger C = \mathbf{1}\,. \qquad (A8.40)$$

Examples

$$d = 2: \quad C = \sigma_2\,, \qquad d = 4\,: \quad C = \gamma_1\gamma_3\,.$$

Note that the unitary matrix

$$\tilde C = C\hat\gamma\,, \qquad (A8.41)$$

then satisfies

$$\begin{aligned}\tilde C^\dagger\, {}^T\gamma_\mu\, \tilde C &= \hat\gamma^\dagger C^\dagger\, {}^T\gamma_\mu\, C\hat\gamma = -\hat\gamma^\dagger\gamma_\mu\hat\gamma \\ &= \gamma_\mu\,. \end{aligned} \qquad (A8.42)$$

The matrix γ_S is symmetric. Under C or $\tilde C$, it transforms like

$$C^\dagger\, {}^T\gamma_S C = (-1)^{d/2}\,\gamma_S\,.$$

Odd dimensions. In odd dimensions γ_S becomes γ_d. We see that the property (A8.39) 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 (A8.42)).

Spin group: conjugate representation. We now apply the transformation (A8.39) to an element Λ of the unitary representation of the spin group, using the form (A8.32),

$$C \exp\left(\tfrac{1}{8}\theta_{\mu\nu}[\gamma_\mu,\gamma_\nu]\right) C^\dagger = \exp\left(\tfrac{1}{8}\theta_{\mu\nu}[\gamma_\mu^*,\gamma_\nu^*]\right),$$

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

$$C\hat\gamma C^\dagger = \hat\gamma^*,$$

and this extends the property to reflections.

For reasons which will become clear later we call these transformations charge conjugation.

Charge conjugation and chiral components. The transformation properties of γ_S imply

$$\begin{cases} C^\dagger(1 + {}^T\gamma_S)C = (1+\gamma_S) & \text{for } d = 0 \pmod 4, \\ C^\dagger(1 + {}^T\gamma_S)C = (1-\gamma_S) & \text{for } d = 2 \pmod 4). \end{cases}$$

If the dimension is a multiple of four C respects chirality, otherwise it exchanges chiral components. Charge conjugation multiplied by a reflection like $C\Pi_\mu$ has the opposite property.

A8.1.7 Charge conjugation

We exhibit unitary matrices C such that

$$C^{\dagger}\,{}^{T}\gamma_{\mu}\,C = -\gamma_{\mu} \quad \Leftrightarrow \quad C\,\gamma_{\mu}\,C^{\dagger} = -\,{}^{T}\gamma_{\mu}\,. \tag{A8.39}$$

Note that since the matrices γ_{μ} are hermitian ${}^{T}\gamma_{\mu} = \gamma_{\mu}^{*}$.

Even dimensions. In the representation of Section A8.1.4, for d even we can take

$$C = \pm\,(\hat{\gamma})^{d/2} \prod_{\text{all } \mu \text{ odd}} \gamma_{\mu}\,, \quad \Rightarrow C^{\dagger}C = \mathbf{1}\,. \tag{A8.40}$$

Examples

$$d = 2: \quad C = \sigma_2\,, \qquad d = 4: \quad C = \gamma_1\gamma_3\,.$$

Note that the unitary matrix

$$\tilde{C} = C\hat{\gamma}\,, \tag{A8.41}$$

then satisfies

$$\begin{aligned}\tilde{C}^{\dagger}\,{}^{T}\gamma_{\mu}\,\tilde{C} &= \hat{\gamma}^{\dagger}C^{\dagger}\,{}^{T}\gamma_{\mu}\,C\hat{\gamma} = -\hat{\gamma}^{\dagger}\gamma_{\mu}\hat{\gamma} \\ &= \gamma_{\mu}\,.\end{aligned} \tag{A8.42}$$

The matrix γ_S is symmetric. Under C or $\tilde{C}$, it transforms like

$$C^{\dagger}\,{}^{T}\gamma_S C = (-1)^{d/2}\,\gamma_S\,.$$

Odd dimensions. In odd dimensions γ_S becomes γ_d. We see that the property ($A8.39$) 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 ($A8.42$)).

Spin group: conjugate representation. We now apply the transformation ($A8.39$) to an element Λ of the unitary representation of the spin group, using the form ($A8.32$),

$$C \exp\left(\tfrac{1}{8}\theta_{\mu\nu}[\gamma_{\mu},\gamma_{\nu}]\right) C^{\dagger} = \exp\left(\tfrac{1}{8}\theta_{\mu\nu}[\gamma_{\mu}^{*},\gamma_{\nu}^{*}]\right),$$

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

$$C\hat{\gamma}C^{\dagger} = \hat{\gamma}^{*},$$

and this extends the property to reflections.

For reasons which will become clear later we call these transformations charge conjugation.

Charge conjugation and chiral components. The transformation properties of γ_S imply

$$\begin{cases} C^{\dagger}(1 + {}^{T}\gamma_S)C = (1 + \gamma_S) & \text{for } d = 0 \pmod 4, \\ C^{\dagger}(1 + {}^{T}\gamma_S)C = (1 - \gamma_S) & \text{for } d = 2 \pmod 4. \end{cases}$$

If the dimension is a multiple of four C respects chirality, otherwise it exchanges chiral components. Charge conjugation multiplied by a reflection like $C\Pi_{\mu}$ has the opposite property.

A8.2 The Example of Dimension 4

Since the dimension 4 plays a special role we specialize some of the previous results to the example of the spin group Spin(4) which is homomorphic to $SO(4)$. With the conventions of Section A8.1.4 the 4×4 γ matrices take the form

$$\gamma_{i=1,2,3} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \qquad \gamma_4 = \begin{pmatrix} 0 & -i\mathbf{1}_2 \\ i\mathbf{1}_2 & 0 \end{pmatrix}$$

(σ_i are the three Pauli matrices) and $\gamma_S \equiv \gamma_5$,

$$\gamma_5 = -\hat{\gamma} = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix}.$$

The matrices $\boldsymbol{\sigma}_{\mu\nu}$ then become

$$\boldsymbol{\sigma}_{ij} = \epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \text{ for } i,j,k \le 3\,, \qquad \boldsymbol{\sigma}_{i4} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \text{ for } i \le 3\,.$$

We recognize in the matrices

$$\boldsymbol{\sigma}_i^{\pm} = \tfrac{1}{4}\epsilon_{ijk}\boldsymbol{\sigma}_{jk} \pm \tfrac{1}{2}\boldsymbol{\sigma}_{i4}\,,$$

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

$$U^* = \sigma_2 U \sigma_2 \;\; \forall\, U \in SU(2)$$

(see also Section A8.1.7) and thus if φ and χ are two $SU(2)$ spinors the combination

$$\varphi_\alpha \left(\sigma_2\right)_{\alpha\beta} \chi_\beta = -i\epsilon_{\alpha\beta}\varphi_\alpha\chi_\beta\,,$$

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 $\gamma_1\gamma_3 = \mathbf{1}_2 \otimes \sigma_2$.

A8.3 Traces of Products of γ Matrices

Perturbative calculations involving relativistic fermions often require the calculation of traces of products of γ matrices, which we, therefore, 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 ($A8.1$) 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 thus define

$$\operatorname{tr} \mathbf{1} = N\,, \quad N \in \mathbb{R}_+\,. \tag{A8.43}$$

A8.2 The Example of Dimension 4

Since the dimension 4 plays a special role we specialize some of the previous results to the example of the spin group Spin(4) which is homomorphic to $SO(4)$. With the conventions of Section A8.1.4 the 4×4 γ matrices take the form

$$\gamma_{i=1,2,3} = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \qquad \gamma_4 = \begin{pmatrix} 0 & -i\mathbf{1}_2 \\ i\mathbf{1}_2 & 0 \end{pmatrix}$$

(σ_i are the three Pauli matrices) and $\gamma_S \equiv \gamma_5$,

$$\gamma_5 = -\hat{\gamma} = \begin{pmatrix} \mathbf{1}_2 & 0 \\ 0 & -\mathbf{1}_2 \end{pmatrix}.$$

The matrices $\boldsymbol{\sigma}_{\mu\nu}$ then become

$$\boldsymbol{\sigma}_{ij} = \epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \text{ for } i,j,k \le 3\,, \qquad \boldsymbol{\sigma}_{i4} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \text{ for } i \le 3\,.$$

We recognize in the matrices

$$\boldsymbol{\sigma}_i^{\pm} = \tfrac{1}{4}\epsilon_{ijk}\boldsymbol{\sigma}_{jk} \pm \tfrac{1}{2}\boldsymbol{\sigma}_{i4}\,,$$

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

$$U^* = \sigma_2 U \sigma_2 \;\; \forall\, U \in SU(2)$$

(see also Section A8.1.7) and thus if φ and χ are two $SU(2)$ spinors the combination

$$\varphi_\alpha \left(\sigma_2\right)_{\alpha\beta} \chi_\beta = -i\epsilon_{\alpha\beta}\varphi_\alpha\chi_\beta\,,$$

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 $\gamma_1\gamma_3 = \mathbf{1}_2 \otimes \sigma_2$.

A8.3 Traces of Products of γ Matrices

Perturbative calculations involving relativistic fermions often require the calculation of traces of products of γ matrices, which we, therefore, 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 ($A8.1$) 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 thus define

$$\operatorname{tr}\mathbf{1} = N\,, \quad N \in \mathbb{R}_+\,. \tag{A8.43}$$

Since we work in the framework of the relation ($A8.9$) we can restrict the analysis to *even dimensions.*

Traces of odd elements. We first consider odd elements, which are such that $\mathrm{P}(c) = -c$ (definition ($A8.2$)). By using the relations ($A8.1$) 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 (no sum over μ)

$$\operatorname{tr}\Gamma^A = \operatorname{tr}\gamma_\mu^2\Gamma^A = -\operatorname{tr}\gamma_\mu\Gamma^A\gamma_\mu = -\operatorname{tr}\Gamma^A \;\Rightarrow \operatorname{tr}\Gamma^A = 0\,.$$

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, $\operatorname{tr}\gamma_{\mu_1}\ldots\gamma_{\mu_{2n}}$, we successively commute $\gamma_{\mu_{2n}}$ through all other factors $\gamma_{\mu_1},\ldots,\gamma_{\mu_{2n-1}}$, using the commutation relations ($A8.1$). 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

$$\begin{aligned}\operatorname{tr}\gamma_{\mu_1}\ldots\gamma_{\mu_{2n}} &= \delta_{\mu_1\mu_{2n}}\operatorname{tr}\left(\gamma_{\mu_2}\ldots\gamma_{\mu_{2n-1}}\right) - \delta_{\mu_2\mu_{2n}} \\ &\quad\times\operatorname{tr}\left(\gamma_{\mu_1}\gamma_{\mu_3}\ldots\gamma_{\mu_{2n-1}}\right) + \cdots + \delta_{\mu_{2n-1}\mu_{2n}}\operatorname{tr}\left(\gamma_{\mu_1}\ldots\gamma_{\mu_{2n-2}}\right).\end{aligned} \tag{A8.44}$$

We, therefore, prove by induction Wick's theorem for the trace of a product of an even number of generators γ_μ:

$$\operatorname{tr}\gamma_{\mu_1}\ldots\gamma_{\mu_{2n}} = N \sum_{\substack{\text{all possible pairings}\\ \text{of } (1,2,\ldots,2n)}} \epsilon(P)\delta_{\mu_{P_1}\mu_{P_2}}\ldots\delta_{\mu_{P_{2n-1}}\mu_{P_{2n}}}, \tag{A8.45}$$

in which $\epsilon(P)$ is the signature of the permutation P when $P_{2m-1} < P_{2m}$ for $1 \le m \le n$.

The element γ_S. Of direct interest, and because $\gamma_S = \gamma_{d+1}$ of interest for odd dimensions, are the calculations of traces of products of the form $\operatorname{tr}\gamma_S\gamma_{\mu_1}\ldots\gamma_{\mu_{2n}}$ (the trace vanishes for an odd number of γ_μ generators). From Wick's theorem it follows immediately that

$$\operatorname{tr}\gamma_S\gamma_{\mu_1}\ldots\gamma_{\mu_{2n}} = 0 \text{ for } 2n < d\,.$$

For $2n = d$ from $\gamma_S = i^{-d/2}\hat\gamma$ one infers

$$\operatorname{tr}\gamma_S\gamma_{\mu_1}\ldots\gamma_{\mu_d} = N\, i^{d/2}\epsilon_{\mu_1\ldots\mu_d}, \tag{A8.46}$$

in which $\epsilon_{\mu_1\ldots\mu_d}$ is the completely antisymmetric tensor normalized by

$$\epsilon_{12\ldots d} = 1\,. \tag{A8.47}$$

We shall see later that relation ($A8.46$), 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.

A8.4 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 call these matrices 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. Let us express that any $2^{d/2} \times 2^{d/2}$ matrix $\mathbf{X}$ can be expanded on the basis of Γ matrices:

$$X_{ab} = N^{-1} \operatorname{tr} \mathbf{X} \Gamma^A \Gamma^A_{ab} \,, \tag{A8.48}$$

in which N is the trace of $\mathbf{1}$. We now choose a matrix $\mathbf{X}$ of the form

$$X_{ab} = \Gamma^B_{cb} \Gamma^C_{ad} \,. \tag{A8.49}$$

Identity ($A8.48$) becomes

$$\Gamma^B_{cb} \Gamma^C_{ad} = N^{-1} \left(\Gamma^B \Gamma^A \Gamma^C\right)_{cd} \Gamma^A_{ab} . \tag{A8.50}$$

By expanding the product $\Gamma^B \Gamma^A \Gamma^C$ on the basis of Γ matrices,

$$\Gamma^B \Gamma^A \Gamma^C = N M^{BC}_{AD} \Gamma^D \,,$$

we obtain the decomposition of any element of one basis onto another:

$$\Gamma^B_{cb} \Gamma^C_{ad} = M^{BC}_{AD} \Gamma^D_{cd} \Gamma^A_{ab} .$$

Examples

(i) For $d = 2$ a basis is $\mathbf{1}$ and σ_i. We leave as an exercise to verify that the subset $\mathbf{1} \otimes \mathbf{1}$ and $\sigma_i \otimes \sigma_i$ transforms into itself with a matrix $\mathbf{M}_2$:

$$\mathbf{M}_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} .$$

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 leave as an exercise 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 a matrix $\mathbf{M}_4$ of square $\mathbf{1}$:

$$\mathbf{M}_4 = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 4 & -2 & -4 & 2 & 0 \\ 1 & -1 & 1 & -1 & 1 \\ 4 & 2 & -4 & -2 & 0 \\ 6 & 0 & 6 & 0 & -2 \end{pmatrix} .$$

9 QUANTUM FIELD THEORY: DIVERGENCES AND REGULARIZATION

We have explained in Chapters 7 and 8 how in quantum field theory physical quantities can be calculated as power series in the various interactions. We now consider specifically *local* relativistic field theories: the action is the integral of the lagrangian density that is a function of fields and their derivatives. We show that, as a consequence, infinities appear in perturbative calculations due to severe short distance singularities, or after Fourier transformation, to integrals diverging at large momenta: one speaks of UV divergences. These divergences are peculiar to local Quantum Field Theory: in contrast with Classical Mechanics or non-relativistic Quantum Mechanics with a finite number of particles, a straightforward construction of a quantum field theory of point-like objects is impossible.

In Chapter 10, we shall explain how to deal with this problem, at least in the framework of formal perturbation theory. The study of critical phenomena by renormalization group methods will eventually suggest a possible interpretation of these divergences.

In this chapter we first display the problem, using the example of the ϕ^3 field theory at one-loop order. We then systematically characterize divergences by power counting: this leads to a classification of local field theories into three families, super-renormalizable, renormalizable and non-renormalizable. We show that divergences of correlation functions involving composite operators require a separate analysis.

Our discussion is based on the momentum representation of Feynman diagrams for several reasons. It is easier to study the behaviour of integrals at large momenta than the singularities of distributions in space variables. Moreover, the explicit expressions of Section 7.6 show that amputation and Legendre transformation are in momentum representation purely algebraic operations in the sense that they involve no momentum integration because propagators are then diagonal. Therefore, the divergences of connected correlation functions can easily be deduced from the divergences of proper vertices. Since a short analysis reveals that the structure of divergences of proper vertices is simpler, we examine in what follows the divergences of Feynman diagrams contributing to the generating functional $\Gamma(\varphi)$, that is, 1PI Feynman diagrams.

To understand the structure of divergences, one modifies the field theory at large momentum, short distance or otherwise in such a way that the new Feynman diagrams become well-defined finite quantities. The modifications must be such that when one control parameter reaches some limit (for example, the momentum cut-off is sent to infinity), the original perturbation theory is formally recovered. This procedure is called *regularization*. It allows isolating well-defined divergent parts of diagrams and deal with them with renormalization as will be explained in Chapter 10. Note that from the point of view of Particle Physics all these modifications alter in some essential way the physical properties of the theory, and thus can only be considered as intermediate steps in the removal of divergences.

Many regularization methods have been introduced in the literature, but we describe only three of them here: momentum or Pauli–Villars, lattice and dimensional regularizations, which have different advantages and shortcomings. In specific applications two (not necessarily compatible) criteria are important: (i) in some theories, symmetries play a crucial role and it is essential that the regularization preserves the symmetry; (ii) some regularization schemes lead to simpler explicit calculations of Feynman diagrams.

9.1 Divergences at One-Loop Order: The ϕ^3 Field Theory

The action for a scalar field ϕ with a ϕ^3 self-interaction can be written as

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi(x)\right)^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{3!} g\, \phi^3(x)\right], \tag{9.1}$$

where g is a constant. This theory is somewhat unphysical because the potential is not bounded from below. However, it has a well-defined perturbative expansion where this non-perturbative pathology does not show up (see Chapter 40). Moreover, when g is purely imaginary, it makes sense beyond perturbation theory and describes in classical statistical physics universal properties of the Yang–Lee edge singularity of the Ising model.

9.1.1 Perturbation theory at one-loop order

The 1PI functional $\Gamma(\varphi)$ has been calculated at one-loop order in Section 7.7.

Tree approximation. In the tree approximation the 1PI functional $\Gamma(\varphi)$ reduces to $\mathcal{S}(\varphi)$. The inverse or 1PI two-point function is thus

$$\Gamma^{(2)}_{\text{tree}}(x,y) = \left(-\nabla_x^2 + m^2\right)\delta(x-y),$$

and after Fourier transformation,

$$\tilde{\Gamma}^{(2)}_{\text{tree}}(p) = p^2 + m^2 . \tag{9.2}$$

More generally, the Fourier components of the 1PI n-point functions are

$$\tilde{\Gamma}^{(3)}_{\text{tree}}(p_1, p_2, -p_1 - p_2) = g\,, \quad \tilde{\Gamma}^{(n)}_{\text{tree}}(p_1, \ldots, p_n) = 0 \quad \text{for } n > 3\,. \tag{9.3}$$

One-loop order. Using equation (7.101), we now calculate the one-loop contribution $\Gamma_1(\varphi)$:

$$\Gamma_1(\varphi) = \tfrac{1}{2}\operatorname{tr}\ln\left[1 + g\left(-\nabla^2 + m^2\right)^{-1}\varphi\right]. \tag{9.4}$$

The expansion of $\Gamma_1(\varphi)$ in powers of φ generates all one-loop contributions to the 1PI functions $\Gamma^{(n)}$. After Fourier transformation we find

$$\begin{aligned}\tilde{\Gamma}^{(n)}_{1\,\text{loop}} = & -\frac{(n-1)!}{2}(-g)^n \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} \frac{1}{(q+p_1)^2 + m^2} \cdots \\ & \times \frac{1}{(q + p_1 + \cdots + p_{n-1})^2 + m^2},\end{aligned} \tag{9.5}$$

which is the quantity represented by the Feynman diagram of figure 9.1.

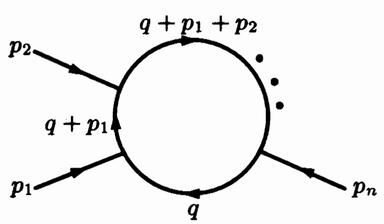


Fig. 9.1 One-loop 1PI diagrams.

The integrand in expression (9.5) behaves, for large momentum q, like $1/q^{2n}$, and the integral thus diverges for $2n \leq d$. Except for $d = 1$ (quantum mechanics) divergences appear. For $d = 2$ the one-point function, which has no momentum dependence, diverges like $\int d^2q/q^2$.

With increasing dimension d more correlation functions diverge. For $d = 6$, the one, two and three-point functions diverge. If the momentum integrals are cut at a large momentum $|\mathbf{q}| = O(\Lambda)$, then the one- and two-point functions diverge like powers of Λ, while $\widetilde{\Gamma}^{(3)}$ diverges logarithmically.

To calculate the contributions which diverge with Λ 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 $d \geq k + 2n$.

Therefore, the divergent part of a one-loop contribution to the n-point function is a polynomial of degree $d - 2n$.

The first important observation is that since the divergences are polynomials in the external momenta, the corresponding divergent contribution $\Gamma_1^{\text{div}}(\varphi)$ to the 1PI functional $\Gamma(\varphi)$ is *local*, that is, takes the form of the space integral of a function of the field and its derivatives, like the action itself (see Section A9.2 for a direct calculation).

To isolate more precisely a divergent part, it is necessary to introduce one of the regularization methods discussed starting in Section 9.5. For $d = 6$, cutting the momentum integral according to Schwinger's regularization (see equation (A9.16)), one finds

$$\begin{aligned}
\widetilde{\Gamma}^{(1)}_{1\,\text{loop}} &= \frac{g}{2^7\pi^3}\left[\tfrac{1}{2}\Lambda^4 - \tfrac{1}{2}m^2\Lambda^2 + m^4\ln(\Lambda/m) + O(1)\right], \\
\widetilde{\Gamma}^{(2)}_{1\,\text{loop}} &= -\frac{g^2}{2^7\pi^3}\left[\Lambda^2 - \left(2m^2 + p^2/3\right)\ln(\Lambda/m) + O(1)\right], \\
\widetilde{\Gamma}^{(3)}_{1\,\text{loop}} &= \frac{g^3}{2^6\pi^3}\ln(\Lambda/m) + O(1).
\end{aligned} \tag{9.6}$$

The three divergent one-loop diagrams are displayed in figure 9.2.

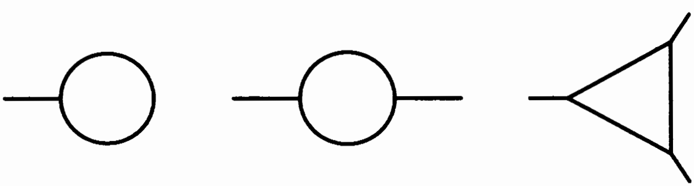


Fig. 9.2 Divergent one-loop diagrams in a $\phi^3_{d=6}$ field theory.

We note that the dimension 6 is special in the following sense: the 1PI correlation functions that diverge are all those which are already non-trivial in the tree approximation (a term linear in ϕ can be added to the action (9.1) by translating ϕ by a constant). Moreover, the divergent terms and the tree approximation have the same momentum dependence.

By contrast, for $d \geq 8$ the four-point function, which vanishes in the tree approximation, is also divergent.

9.1.2 Empirical removal of divergences at one-loop order

For $d \leq 6$ dimensions, the divergent parts of the one-loop correlation functions have the structure of the original action. For example in $d = 6$, $\Gamma_1^{\text{div.}}(\varphi)$, the divergent part at one-loop of $\Gamma(\varphi)$, has the structure

$$\Gamma_1^{\text{div}}(\varphi) = \int \mathrm{d}^6 x \left[\frac{1}{2} g^2 a_0(\Lambda)\left(\partial_\mu \varphi(x)\right)^2 + g a_1(\Lambda)\varphi(x) + \frac{1}{2} g^2 a_2(\Lambda)\varphi^2(x) \right.$$
$$\left. + \frac{1}{3!} g^3 a_3(\Lambda)\varphi^3(x)\right]. \tag{9.7}$$

The functions $a_i(\Lambda)$ follow from equations (9.6) and are, therefore, defined only up to additive finite parts. For later purposes (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, we find

$$\begin{aligned} 2^7\pi^3 a_0(\Lambda) &= \tfrac{1}{3}\ln(\Lambda/m), \\ 2^7\pi^3 a_1(\Lambda) &= \tfrac{1}{2}\Lambda^4 - \tfrac{1}{2}m^2\Lambda^2 + m^4\ln(\Lambda/m), \\ 2^7\pi^3 a_2(\Lambda) &= -\Lambda^2 + 2m^2\ln(\Lambda/m), \\ 2^7\pi^3 a_3(\Lambda) &= 2\ln(\Lambda/m). \end{aligned} \tag{9.8}$$

If we now add the local contribution $-\Gamma_1^{\text{div}}(\varphi)$ to the initial action $\mathcal{S}(\phi)$, the new action,

$$\mathcal{S}_1(\phi) = \mathcal{S}(\phi) - \Gamma_1^{\text{div}}(\varphi),$$

differs from the original action by its parametrization, but involves the same monomials of the field and, therefore, depends on the same number of parameters:

$$\mathcal{S}_1(\phi) = \int \mathrm{d}^6 x \left[\frac{1}{2}\left(1 - g^2 a_0(\Lambda)\right)\left(\partial_\mu\phi\right)^2 - g a_1(\Lambda)\phi + \frac{1}{2}\left(m^2 - g^2 a_2(\Lambda)\right)\phi^2 \right.$$
$$\left. + \frac{1}{3!}\left(g - g^3 a_3(\Lambda)\right)\phi^3\right].$$

The new 1PI functional $\Gamma(\varphi)$ at one-loop order is then

$$\begin{aligned} \Gamma(\varphi) &= \mathcal{S}_1(\varphi) + \Gamma_1(\varphi) + O\,(\text{two loops}) \\ &= \mathcal{S}(\varphi) - \Gamma_1^{\text{div}}(\varphi) + \Gamma_1(\varphi) + O\,(\text{two loops}), \end{aligned}$$

where $\Gamma_1(\varphi)$ is the sum of one-loop diagrams calculated only with $\mathcal{S}(\phi)$. Therefore, $\Gamma(\varphi)$ now has a limit at one-loop order when the cut-off becomes infinite.

Finally, note that a change in the definition of the divergent part changes $\Gamma_1^{\text{div}}(\varphi)$ by a finite local polynomial, and the conclusions are the same.

The renormalization idea. To solve, at least formally, the problem of the divergences of perturbation theory, one tries to generalize the previous method to all orders in the loop expansion and to more general field theories. One thus introduces a large momentum cut-off in the theory, or equivalently modify the field theory at short distance, as we have done here, to characterize the divergences of Feynman diagrams (the physical reality of such a cut-off is from this point of view irrelevant). One then investigates the possibility of choosing the original parameters of the theory as functions of the cut-off in such a way that correlation functions have a finite large cut-off limit. When such a limit exists, one can show that it is independent of the cut-off procedure (under some general conditions). The local field theories, for which this procedure works, are called renormalizable (or super-renormalizable if some parameters in the interaction are cut-off independent). **Renormalizable theories are short distance insensitive** in the sense that even if a large mass or a microscopic scale in space provide a true physical cut-off, their long distance or low momentum properties can be described, without explicit knowledge of the short distance structure, in terms of a small number of *effective parameters.* However, and this is a deep issue, the ratio between the microscopic and macroscopic scales does not disappear from the description. We already emphasize these ideas here because they motivate the technical analysis which follows. Renormalization group will eventually provide us with the necessary tool to understand the renormalization procedure (see Chapters 23–37).

Note, that in the ϕ^3 theory, for $d \geq 8$, the method fails already at one-loop order. Indeed to generate in the tree approximation a term proportional to the divergence of the four-point function, for example, a ϕ^4 term is required in the action. We show below that such an interaction induces, in turn, worse divergences which cannot be reproduced. Power counting will show that no polynomial interaction can lead to a finite theory.

9.2 Divergences General Analysis and Power Counting

We first introduce the notion of canonical, or engineering, dimension of fields and interaction vertices.

9.2.1 Dimension of fields and vertices

We explicitly assume that the propagator $\Delta(p)$ of every field ϕ is $O(d)$ covariant in d dimensions, or at least has a uniform large momentum behaviour of the form

$$C_1\lambda^{-\sigma} < |\Delta(\lambda p)| < C_2\lambda^{-\sigma} \quad \forall\, p \neq 0 \quad \text{for } \lambda \to \infty\,, \tag{9.9}$$

in which C_1 is a strictly positive constant. Other cases require a special analysis.

The canonical, or engineering dimension $[\phi]$ of a field $\phi(x)$ is then defined in terms of the large momentum behaviour of the ϕ propagator by

$$[\phi] = \tfrac{1}{2}(d-\sigma). \tag{9.10}$$

Examples. In the scalar field theory (9.1),

$$\Delta(\lambda p) \propto 1/\lambda^2\,, \quad \text{for } \lambda \to \infty \quad \Rightarrow \quad [\phi] = \tfrac{1}{2}(d-2). \tag{9.11}$$

For the fermions considered in Section 8.3, the propagator of the field Fourier components reads

$$\widetilde{W}^{(2)}_{\alpha\beta}(p) = \langle \bar\psi_\alpha(-p)\psi_\beta(p)\rangle = (m + i\not{p})^{-1}_{\beta\alpha}\,,$$

and, thus,

$$\Delta(\lambda p) \propto 1/\lambda\,, \quad \text{for } \lambda \to \infty \quad \Rightarrow \quad [\psi] = [\bar\psi] = \tfrac{1}{2}(d-1). \tag{9.12}$$

In the simple case of scalar and spinor fermion fields, the definition (9.10) coincides with the natural mass dimension of the field as deduced from the quadratic part of the action by dimensional analysis. Assigning a dimension +1 to momenta p and masses m,

$$[p] = [m] = 1\,,$$

correspondingly a dimension −1 to length and position variables,

$$[x] = -1 \quad \Rightarrow \quad [\partial/\partial x] = +1\,,$$

and expressing that the action is dimensionless, one indeed finds

$$\left[\int \mathrm{d}^d x\, (\partial_\mu \phi(x))^2\right] = 0 \quad \Rightarrow \quad -d + 2 + 2[\phi] = 0\,,$$
$$\left[\int \mathrm{d}^d x\, \bar\psi(x) \not\partial \psi(x)\right] = 0 \quad \Rightarrow \quad -d + 1 + 2[\psi] = 0\,.$$

This property no longer holds for higher spin fields, in general. For example, the free action for a massive vector field A_μ is

$$\mathcal{S}(A) = \int \mathrm{d}^d x \left(\frac{1}{4}\sum_{\mu\nu} (\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \frac{1}{2} m^2 A_\mu A_\mu\right). \tag{9.13}$$

The A-field propagator $\Delta_{\mu\nu}(p)$,

$$\Delta_{\mu\nu} = \frac{\delta_{\mu\nu} + p_\mu p_\nu/m^2}{p^2 + m^2}\,, \tag{9.14}$$

is such that

$$\Delta_{\mu\nu}(\lambda p) \propto \lambda^0 \quad \text{for } \lambda \to \infty \Rightarrow [A_\mu] = d/2, \tag{9.15}$$

while from dimensional analysis one would have concluded that A_μ has the dimension of a scalar field. This property is directly related to the presence of negative powers of m^2 in the propagator: the quadratic form in the action is not invertible for $m = 0$. On the pole $p^2 = -m^2$ of the propagator, the numerator of $\Delta_{\mu\nu}$ becomes a projector orthogonal to p_μ.

The same phenomenon also occurs with higher spin fields. A spin s massive field propagator has the form

$$\Delta(p) = \frac{P_{2s}(p/m)}{p^2 + m^2}\,, \tag{9.16}$$

in which $P_{2s}(p)$ is a polynomial of degree $2s$ in p, which is a projector on "mass-shell", that is, for $p^2 = -m^2$. The dimension of the corresponding field ϕ_s is

$$[\phi_s] = \tfrac{1}{2}(d - 2 + 2s). \tag{9.17}$$

Equation (9.17) generalizes equations (9.11,9.12,9.15).

Dimension of vertices. We consider here only theories invariant under space translations. The interaction in the action is thus a linear combination of monomials which are space integrals of a product of fields and their derivatives. Let us write a monomial $V(\phi)$ symbolically:

$$V(\phi) \propto \int \mathrm{d}^d x \, (\nabla_x)^k \, \phi_1^{n_1}(x)\phi_2^{n_2}(x)\ldots\phi_s^{n_s}(x),$$

where the k derivatives act in an unspecified way on the fields ϕ_i. We call these elementary interaction terms vertices because they are represented by vertices in Feynman diagrams.

We now define the dimension $[V]$ of the vertex $V(\phi)$ by

$$[V] = -d + k + \sum_{i=1}^{s} n_i \, [\phi_i] \,, \tag{9.18}$$

in which k is the total number of differential operators and n_i the number of fields ϕ_i appearing in the vertex V.

In terms of the Fourier components $\tilde{\phi}_i(p)$ of the fields $\phi_i(x)$, and taking into account translation invariance, the vertex $V(\phi)$ can be written (in analogous symbolic notation) as

$$V(\phi) \propto \int \prod_{n=1}^{n_1+\cdots+n_s} \mathrm{d}^d p_n \; \delta^{(d)}(p_1 + p_2 + \cdots + p_{n_1+\cdots+n_s}) \, p^k \tilde{\phi}_1(p_1)\ldots\tilde{\phi}_s(p_{n_1+\cdots+n_s}).$$

9.2.2 Superficial degree of divergence: power counting

As mentioned before, we consider only 1PI diagrams. Each vertex yields a δ-function of momentum conservation. 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 7.7.

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 $\lambda \to \infty$ the diagram is scaled by a factor $\lambda^{\delta(\gamma)}$ with

$$\delta(\gamma) = dL - \sum_i I_i \sigma_i + \sum_\alpha v_\alpha k_\alpha \,, \tag{9.19}$$

in which v_α is the number of vertices of type α with k_α derivatives, and I_i the number of internal lines corresponding to propagators Δ_i joining the different vertices.

The number $\delta(\gamma)$ is called the superficial degree of divergence of the diagram γ. For a one-loop diagram regularized with a momentum cut-off, it characterizes the divergence of the diagram as a power of the cut-off.

More generally, if $\delta(\gamma)$ is positive a regularized diagram diverges at least like $\Lambda^{\delta(\gamma)}$. If $\delta(\gamma) = 0$, it diverges at least like a power of $\ln \Lambda$. If $\delta(\gamma)$ is negative, the diagram is superficially convergent, which means that divergences can come only from subdiagrams.

Example. In the example of the ϕ^3 field theory in $d = 6$ dimensions, since $\sigma = 2$ and $k = 0$, expression (9.19) yields $\delta(\gamma) = 6 - 2I$. To $I = 1, 2, 3$, respectively, correspond the values 4, 2, 0 in agreement with equations (9.6). For $I > 3$ the diagrams are convergent.

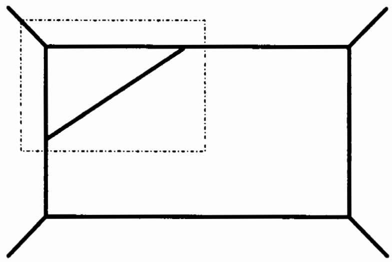


Fig. 9.3 Divergent subdiagram.

Figure 9.3 exhibits a superficially convergent diagram with a divergent subdiagram in the same theory: the superficial degree of divergence is -2, the diagram is superficially convergent, but the subdiagram inside the dotted box is divergent.

Other expression. Various topological relations on graphs allow to write $\delta(\gamma)$ in different forms.

We consider a diagram γ contributing to a 1PI correlation function with E_i (for external line) fields ϕ_i. Then if we call n_i^α the number of fields ϕ_i at a vertex α belonging to the diagram, we have the relation

$$E_i + 2I_i = \sum_\alpha n_i^\alpha v_\alpha \,. \tag{9.20}$$

The interpretation of the relation is simple: each internal line connects two vertices while an external line is only attached to one vertex.

Figure 9.4 gives an example.

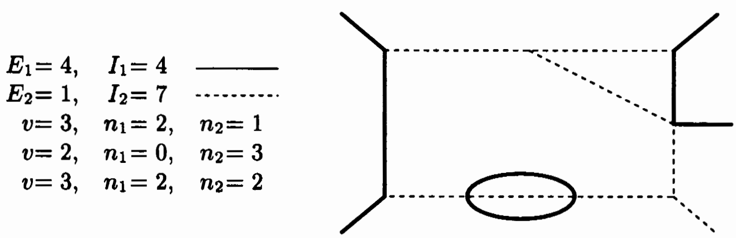


Fig. 9.4

Combining equation (9.20) with the relation (7.111) written in the form

$$L = \sum_i I_i - \sum_\alpha v_\alpha + 1\,, \tag{9.21}$$

we eliminate L and I_i in $\delta(\gamma)$ and obtain

$$\delta(\gamma) = d - \sum_i E_i[\phi_i] + \sum_\alpha v_\alpha [V_\alpha], \tag{9.22}$$

where $[V_\alpha]$ is the dimension of the vertex α (equation (9.18)):

$$[V_\alpha] = -d + k_\alpha + \sum_i n_i^\alpha [\phi_i],$$

and $[\phi_i]$ the dimension of ϕ_i (equation (9.10)):

$$[\phi_i] = \tfrac{1}{2}(d - \sigma_i).$$

Equation (9.22) directly leads to a classification of renormalizable theories.

9.3 Classification of Renormalizable Field Theories

The program outlined in Section 9.1 can only be realized if the superficial degree of divergence is bounded. When this condition is fulfilled we call the theory renormalizable by power counting. We now classify theories based on the form (9.22) of the superficial degree of divergence.

Non-renormalizable theories. If at least one vertex V has a positive dimension, $[V] > 0$, then the degree of divergence of diagrams contributing to any 1PI correlation 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 because in order to cancel divergences one would have to add an infinite number of new interactions to the action, and the final theory would depend on an infinite number of parameters.

The ϕ^3 theory in $d > 6$ dimensions provides an example.

Super-renormalizable theories. When only a finite number of Feynman diagrams are superficially divergent the corresponding field theory is called super-renormalizable. This happens when all vertices have strictly negative dimensions.

Example. In the ϕ^4 field theory in $d = 3$ dimensions,

$$\delta(\gamma) = 3 - \tfrac{1}{2}E - v\,.$$

The superficially divergent diagrams are listed in figure 9.5.

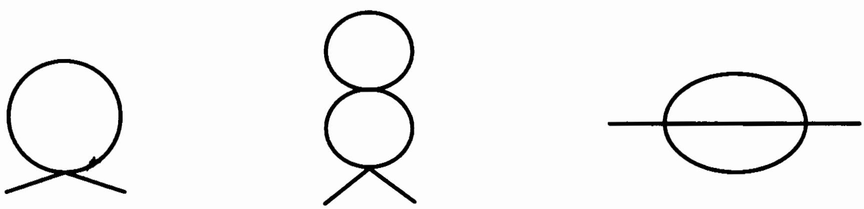


Fig. 9.5 Superficially divergent diagrams in $\phi^4_{d=3}$.

Renormalizable theories. These theories are characterized by the property that at least one vertex has dimension zero, and no vertex has a positive dimension. Then, an infinite number of diagrams have a positive superficial degree of divergence; however, the maximal degree of divergence at E_i fixed does not increase with the number of loops, and is independent of the number of insertions of the vertices of dimension zero.

In addition, if all dimensions of fields $[\phi_i]$ are strictly positive, only a finite number of correlation functions are superficially divergent.

If at least one field has dimension zero, the situation is more complicated: the degree of divergence is bounded; however, an infinite number of correlation functions are superficially divergent. Generically this leads to field theories depending on an infinite number of parameters although, in contrast to the case of non-renormalizable theories, only a subclass of all possible interactions is generated by renormalization.

In addition, in some cases symmetries relate all these parameters so that only a finite number are really independent. Examples will be met in Chapters 14,15 in the discussion of models defined on homogeneous spaces like the non-linear σ-model.

Classification of renormalizable theories. In what follows we restrict ourselves to the most frequent situation: $[\phi_i] \geq 0$ for all fields. Other cases require a special analysis.

We consider simple field theories containing fields of spin s with dimensions $[\phi_s]$ given by equation (9.17):

$$[\phi_s] = \tfrac{1}{2}(d-2+2s).$$

The condition $[\phi_s] > 0$ is satisfied for $d \geq 2$ except for the $s=0$, $d=2$, case which must be examined separately.

All vertices should satisfy

$$-d+k+\tfrac{1}{2}\sum_s n_s(d-2+2s) \leq 0\,. \tag{9.23}$$

This condition bounds k, the number of derivatives, n_s the number of fields of spin s at the vertex, s the spin and the dimension d.

For $k=s=0$, the condition (9.23) implies for all vertices

$$n \leq 2d/(d-2)\,.$$

The corresponding renormalizable interactions are

ϕ^3 in $d=6$ dimensions, ϕ^4 in $d=4$ dimensions, ϕ^6 in $d=3$ dimensions (ϕ^5 is either non- or super-renormalizable).

Note that any polynomial in ϕ is super-renormalizable in $d=2$ dimensions.

With two derivatives, $k=2$, the only solution is $d=2$, but then $[\phi]=0$.

We now consider vertices with one spinor fermion pair $\bar\psi\psi$ and n scalar fields. The condition (9.23) then becomes

$$n \leq 2/(d-2)\,.$$

Renormalizable interactions are

$\psi\bar\psi\phi$ in $d=4$ dimensions, $\bar\psi\psi\phi^2$ in $d=3$ dimensions.

In addition, $P(\phi)\bar\psi\psi$, in which $P(\phi)$ is a polynomial in ϕ, is super-renormalizable in two dimensions.

Finally, the vertex $(\bar\psi\psi)^2$ is renormalizable in two dimensions.

The vertices $P(\phi)(\bar\psi\psi)^2$ and $P(\phi)\bar\psi\not{\partial}\psi$ also have dimension zero in two dimensions but again the dimension of ϕ vanishes.

For spin one vector fields, general $O(d)$ invariance leaves only dimension two as a possibility. The only candidate with only fermions is the vertex

$$\bar\psi A_\mu \gamma_\mu \psi \equiv \bar\psi \not{A} \psi\,,$$

which is renormalizable in two dimensions. In addition the vertices $\phi\partial_\mu\phi A_\mu$ and $\phi^2 A_\mu^2$, which appear in gauge theories, are dimensionless. However, they again lead to a non-trivial renormalization problem because scalar fields are dimensionless ($[\phi]=0$).

No higher spin field leads to renormalizable theories.

Note, however, that spin one vector fields associated with gauge symmetries do not enter into this classification because their propagator has in some gauges the behaviour of a scalar field propagator (for a discussion of gauge theories see Chapters 18–21).

Finally, we note that no physically acceptable, from the point of view of Particle Physics, and renormalizable theory exists above dimension 4. It is not known whether this property has a physics relation with the empirical fact that space–time has just four dimensions, or is a mere coincidence.

9.4 Operator Insertions: Generating Functionals, Power Counting

Up to now we have analysed only the divergences of field correlation functions. However, various physical problems involve correlation functions of local polynomials of the field, called hereafter composite fields or for historical reasons composite operators (this terminology comes from the operator formulation of quantum field theory). Typical examples are

$$\mathcal{O}(\phi;x) \equiv \phi^2(x),\ \phi^4(x),\ [\partial_\mu\phi(x)]^2 \ldots .$$

One insertion of an operator $\mathcal{O}(\phi)$ yields the correlation functions

$$\langle \mathcal{O}(\phi;y)\phi(x_1)\ldots\phi(x_n)\rangle .$$

Such correlation functions can, in principle, be obtained from the field correlation functions by letting various points coincide. However, this procedure corresponds in momentum space to additional integrations, and, therefore, generates new divergences. It is, therefore, necessary to analyse the 1PI correlation functions with operator insertions, from the point of view of power counting, separately.

Generating functional. We again call $\mathcal{S}(\phi)$ the euclidean action and add to it a source (which is a space-dependent coupling constant) for the operator $\mathcal{O}(\phi;x)$:

$$\mathcal{S}_g(\phi) = \mathcal{S}(\phi) + \int \mathrm{d}^d x\, g(x)\mathcal{O}(\phi;x). \tag{9.24}$$

To this new action $\mathcal{S}_g$ corresponds a generating functional $\mathcal{Z}(J,g)$:

$$\mathcal{Z}(J,g) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}_g(\phi) + \int J(x)\phi(x)\mathrm{d}^d x\right]. \tag{9.25}$$

The correlation functions with one operator, $\mathcal{O}(\phi;x)$, insertion can be obtained from the generating functional $\delta\mathcal{Z}/\delta g(x)$ taken at $g=0$:

$$\langle \mathcal{O}(\phi;y)\phi(x_1)\ldots\phi(x_n)\rangle = -\left[\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)}\frac{\delta}{\delta g(y)}\mathcal{Z}(J,g)\right]\Bigg\|_{J=g=0}. \tag{9.26}$$

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 $J(x)$ we obtain the 1PI functional $\Gamma(\varphi,g)$. The generating functional of 1PI correlation functions with one $\mathcal{O}(\phi;y)$ insertion, $\Gamma^{(n)}_{\mathcal{O}}$, is then:

$$\left.\frac{\delta\Gamma(\varphi,g)}{\delta g(y)}\right|_{g=0} = \sum_{n=1}^{\infty}\frac{1}{n!}\int \mathrm{d}x_1\ldots\mathrm{d}x_n\,\varphi(x_1)\ldots\varphi(x_n)\Gamma^{(n)}_{\mathcal{O}}(y;x_1,\ldots,x_n).$$

The corresponding Feynman diagrams have the structure displayed in figure 9.6.

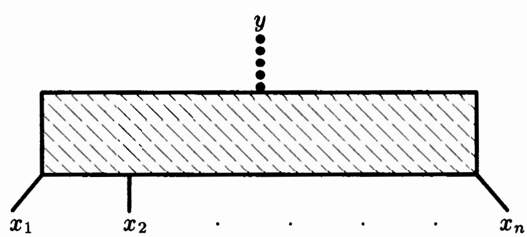


Fig. 9.6

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 not conserved.

Power counting. When all integration momenta are scaled by a factor λ, in the limit $\lambda \to \infty$ all external momenta become negligible. Therefore, asymptotically momentum is conserved even at the vertex corresponding to the operator insertion: the power counting of 1PI correlation functions with one operator insertion is the same as with one vertex insertion. We assign to an operator $\mathcal{O}(\phi)$ the dimension $[\mathcal{O}]$:

$$[\mathcal{O}] = k + \sum_i n_i[\phi_i], \tag{9.27}$$

in which k is the number of derivatives in the operator, and n_i the number of fields of type i. This definition differs from the definition of the dimension of the corresponding vertex (equation (9.18)) by d.

The expression (9.22) of the superficial degree of divergence of a 1PI diagram γ is then modified in the case of the insertion of the product of operators $\mathcal{O}_1(x_1)\ldots\mathcal{O}_r(x_r)$ and becomes

$$\delta_\gamma(\mathcal{O}_1\ldots\mathcal{O}_r) = d - \sum_i E_i[\phi_i] + \sum_\alpha v_\alpha[V_\alpha] + [\mathcal{O}_1] + \cdots + [\mathcal{O}_r] - rd\,. \tag{9.28}$$

For example, for $d=4$ one insertion of $\phi^m(x)$ in the ϕ^4 field theory yields

$$[\phi^m] = m \;\Rightarrow\; \delta_\gamma = 4 - E + m - 4\,. \tag{9.29}$$

In the same theory the 1PI n-point function with $l\,\phi^2$ insertions,

$$\Gamma^{(n,l)}(x_1,\ldots,x_n;y_1,\ldots,y_l) = \left\langle \phi(x_1)\ldots\phi(x_n)\phi^2(y_1)\ldots\phi^2(y_l)\right\rangle_{\mathrm{1PI}},$$

has a degree of divergence δ:

$$\delta = 4 - n - 2l\,. \tag{9.30}$$

The new divergent correlation functions are displayed in figure 9.7.

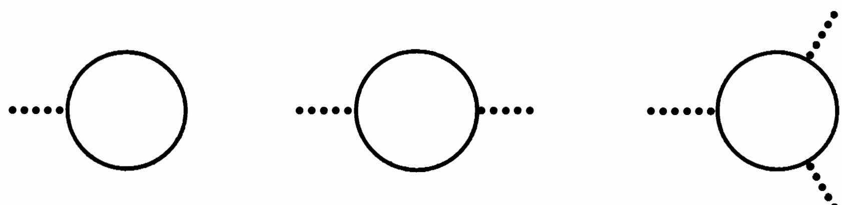


Fig. 9.7

9.5 Momentum Cut-Off and Regulator Fields

Having shown that quantum field theories are in general affected by UV divergences, we now exhibit various modifications of the initial theory that render perturbation theory finite, and which are such that in some limit the initial theory is recovered.

We first discuss methods that work in the continuum (compared to lattice methods) and at fixed dimension (unlike dimensional regularization). The idea is then to modify the field propagators beyond a large momentum cut-off to render all Feynman diagrams convergent. However, any regularization of this type must satisfy one important condition: the inverse of the regularized propagator must remain a *smooth* function of the momentum $\mathbf{p}$. Indeed, singularities in momentum variables generate, after Fourier transformation, contributions to the large distance behaviour of the propagator, and regularization should modify the theory only at short distance.

Momentum cut-off deals with divergences caused by the infinite number of degrees of freedom of fields, a property that will also lead to the renormalization group. With some care it preserves, beyond space–time symmetries, all linear symmetries of the initial action. However, it does not regularize divergences related to quantization problems and order of quantum operators.

9.5.1 Scalars and fermions

Scalar fields. A simple modification of the propagator improves the convergence of Feynman diagrams at large momentum. For example, in the action

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\phi(-\nabla^2 + m^2)\phi + V(\phi)\right], \tag{9.31}$$

one changes the quadratic term

$$\mathcal{S}_{\mathrm{reg}}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\phi(-\nabla^2 + m^2)\prod_{i=1}^{n}(1 - \nabla^2/M_i^2)\phi + V(\phi)\right], \tag{9.32}$$

where the masses M_i grow with the momentum cut-off Λ, for instance (consistent with representation (9.34)),

$$M_i^2 = m^2 + \alpha_i \Lambda^2, \quad \alpha_i > 0\,.$$

The propagator $1/(m^2 + p^2)$ of the scalar field is then replaced by

$$\Delta_{\mathrm{B}}(p) = \left(\frac{1}{p^2 + m^2}\right)_{\mathrm{reg.}}, \quad \Delta_{\mathrm{B}}^{-1}(p) = (p^2 + m^2)\prod_{i=1}^{n}(1 + p^2/M_i^2). \tag{9.33}$$

The degree n is chosen large enough to render all diagrams convergent. In the large cut-off limit, at parameters α fixed, the initial propagator is recovered. This is the spirit of Pauli–Villars's regularization scheme.

Note that such a propagator cannot be derived from a hermitian hamiltonian. Indeed, the hermiticity of the hamiltonian leads to the representation (6.90) 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 $1/p^2$.

While this modification can be implemented also in Minkowski space because the regularized propagators decrease in all complex p^2 directions, in the euclidean form more general modifications are possible. Schwinger's proper time representation (Appendix A9.1) suggests

$$\Delta_{\mathrm{B}}(p) = \int_0^{\infty} \mathrm{d}t\, \rho(t\Lambda^2)\, \mathrm{e}^{-t(p^2+m^2)}, \tag{9.34}$$

in which the function $\rho(t)$ is positive and satisfies the condition

$$|1-\rho(t)| < C\,\mathrm{e}^{-\sigma t}\ (\sigma > 0) \text{ for } t \to +\infty .$$

By choosing a function $\rho(t)$ that decreases fast enough for $t \to 0$, the behaviour of the propagator can be arbitrarily improved. If $\rho(t) = O(t^n)$ the behaviour (9.33) is recovered. Another example is

$$\rho(t) = \theta(t-1), \tag{9.35}$$

$\theta(t)$ being the step function, which leads to exponential decrease:

$$\Delta_{\mathrm{B}}(p) = \frac{\mathrm{e}^{-(p^2+m^2)/\Lambda^2}}{p^2+m^2}. \tag{9.36}$$

As the example (9.36) shows, it is possible to find in this more general class, propagators without unphysical singularities, but they do not follow from a hamiltonian formalism because continuation to real time becomes impossible.

Fermions. For relativistic fermions similar methods are applicable. For euclidean fermion conventions and hermitian matrices γ_μ (Section 8.2) the quadratic term of the fermion action can be written as

$$\mathcal{S}_0(\bar{\psi},\psi) = \int \mathrm{d}^d x\, \bar{\psi}(x)(\not{\partial} + m)\psi(x). \tag{9.37}$$

It can be replaced by

$$\mathcal{S}_{0\,\mathrm{reg}}(\bar{\psi},\psi) = \int \mathrm{d}^d x\, \bar{\psi}(x)(\not{\partial} + m)\prod_{i=1}^{n}(1-\nabla^2/M_i^2)\psi(x). \tag{9.38}$$

The propagator $1/(m+i\not{p})$ is then replaced by $\Delta_{\mathrm{F}}(p)$ with

$$\Delta_{\mathrm{F}}^{-1}(p) = (m+i\not{p})\prod_{i=1}^{n}(1+p^2/M_i^2). \tag{9.39}$$

Note that for $m=0$, this modification preserves, in even dimensions, chiral symmetry (defined by (8.27)).

Power counting. Let us verify that these methods work for simple scalar field theories. We assume that the regularized propagator $\Delta_{\mathrm{B}}(p)$ behaves like $1/p^{2n}$. Equation (9.19) gives the degree of divergence of a regularized diagram γ:

$$\delta(\gamma) = dL - 2nI + \sum_\alpha v_\alpha k_\alpha . \tag{9.40}$$

Using the topological relation (9.21),

$$L = I - \sum_\alpha v_\alpha + 1,$$

general modifications are possible. Schwinger's proper time representation (Appendix A9.1) suggests

$$\Delta_{\mathrm{B}}(p) = \int_0^{\infty} \mathrm{d}t\, \rho(t\Lambda^2)\, \mathrm{e}^{-t(p^2+m^2)}, \tag{9.34}$$

in which the function $\rho(t)$ is positive and satisfies the condition

$$|1-\rho(t)| < C\,\mathrm{e}^{-\sigma t} \ (\sigma > 0) \text{ for } t \to +\infty\,.$$

By choosing a function $\rho(t)$ that decreases fast enough for $t \to 0$, the behaviour of the propagator can be arbitrarily improved. If $\rho(t) = O(t^n)$ the behaviour (9.33) is recovered. Another example is

$$\rho(t) = \theta(t-1), \tag{9.35}$$

$\theta(t)$ being the step function, which leads to exponential decrease:

$$\Delta_{\mathrm{B}}(p) = \frac{\mathrm{e}^{-(p^2+m^2)/\Lambda^2}}{p^2+m^2}. \tag{9.36}$$

As the example (9.36) shows, it is possible to find in this more general class, propagators without unphysical singularities, but they do not follow from a hamiltonian formalism because continuation to real time becomes impossible.

Fermions. For relativistic fermions similar methods are applicable. For euclidean fermion conventions and hermitian matrices γ_μ (Section 8.2) the quadratic term of the fermion action can be written as

$$\mathcal{S}_0(\bar\psi,\psi) = \int \mathrm{d}^d x\, \bar\psi(x)(\not\partial + m)\psi(x). \tag{9.37}$$

It can be replaced by

$$\mathcal{S}_{0\,\mathrm{reg}}(\bar\psi,\psi) = \int \mathrm{d}^d x\, \bar\psi(x)(\not\partial + m)\prod_{i=1}^{n}(1-\nabla^2/M_i^2)\psi(x). \tag{9.38}$$

The propagator $1/(m+i\not p)$ is then replaced by $\Delta_{\mathrm{F}}(p)$ with

$$\Delta_{\mathrm{F}}^{-1}(p) = (m+i\not p)\prod_{i=1}^{n}(1+p^2/M_i^2). \tag{9.39}$$

Note that for $m=0$, this modification preserves, in even dimensions, chiral symmetry (defined by (8.27)).

Power counting. Let us verify that these methods work for simple scalar field theories. We assume that the regularized propagator $\Delta_{\mathrm{B}}(p)$ behaves like $1/p^{2n}$. Equation (9.19) gives the degree of divergence of a regularized diagram γ:

$$\delta(\gamma) = dL - 2nI + \sum_\alpha v_\alpha k_\alpha\,. \tag{9.40}$$

Using the topological relation (9.21),

$$L = I - \sum_\alpha v_\alpha + 1,$$

we can rewrite equation (9.40) as

$$\delta(\gamma) = (d - 2n)L + \sum_{\alpha} v_\alpha (k_\alpha - 2n) + 2n \,.$$

It is thus necessary to choose $2n > d$ and $2n > \sup_\alpha k_\alpha$, since both L and v_α can increase indefinitely. Moreover, since $L \geq 1$ and one at least of the v_α is positive one can then always satisfy $\delta(\gamma) < 0$ for all diagrams.

Examples.
(i) In the ϕ^3 field theory in six dimensions, $2n = 8$ renders all diagrams finite.
(ii) In the ϕ^4 field theory in four dimensions, $2n = 6$ suffices.

9.5.2 Regulator fields

Momentum regularization has another, sometimes equivalent, implementation based on the introduction of regulator fields, as we now show.

Scalar fields. In the case of scalar fields, to regularize the action (9.31) one introduces additional dynamical fields ϕ_k, $k = 1, \ldots, n$, and considers the modified action $\mathcal{S}_{\text{reg}}(\phi, \phi_k)$:

$$\mathcal{S}_{\text{reg}}(\phi, \phi_k) = \int \mathrm{d}^d x \left[\frac{1}{2}\phi\left(-\nabla^2 + m^2\right)\phi + \sum \frac{1}{2z_k}\phi_k\left(-\nabla^2 + M_k^2\right)\phi_k + V(\phi + \textstyle\sum_k \phi_k)\right]. \tag{9.41}$$

With the action (9.41) any internal ϕ propagator is replaced by the sum of the ϕ propagator and all the ϕ_k propagators $z_k/(p^2 + M_k^2)$. For an appropriate choice of the constants z_k, after integration over the regulator fields, the form (9.33) is recovered, as a simple calculation shows. In the functional integral

$$\int [\mathrm{d}\phi] \prod_1^n [\mathrm{d}\phi_k] \exp\left[-\mathcal{S}_{\text{reg}}(\phi, \phi_k)\right], \tag{9.42}$$

we first change variables:

$$\phi = \phi' - \sum_1^n \phi_k \,. \tag{9.43}$$

The regularized action then becomes

$$\begin{aligned}\mathcal{S}_{\text{reg}}(\phi', \phi_k) = \int \mathrm{d}^d x \Bigg[&\frac{1}{2}\left(\phi' - \textstyle\sum_k \phi_k\right)\left(-\nabla^2 + m^2\right)\left(\phi' - \textstyle\sum_k \phi_k\right) \\ &+ \sum_1^n \frac{1}{2z_k}\phi_k\left(-\nabla^2 + M_k^2\right)\phi_k + V(\phi')\Bigg].\end{aligned}$$

The gaussian integration over the fields ϕ_k can now be performed explicitly. A straightforward calculation leads, as expected, to

$$\mathcal{S}_{\text{reg}}(\phi) = \int [\mathrm{d}\phi] \exp\left[-\int \mathrm{d}^d x \left(\tfrac{1}{2}\phi\, S^{(2)}\, \phi + V(\phi)\right)\right], \tag{9.44}$$

$$\left[S^{(2)}\right]^{-1} = \Delta_{\text{B}} = \left(-\nabla^2 + m^2\right)^{-1} + \sum_k z_k \left(-\nabla^2 + M_k^2\right)^{-1}. \tag{9.45}$$

It is possible to choose the coefficients z_k in such a way that

$$\left[\frac{1}{-\nabla^2+m^2}+\sum_k \frac{z_k}{-\nabla^2+M_k^2}\right]^{-1}=(-\nabla^2+m^2)\prod_k \frac{(-\nabla^2+M_k^2)}{(-m^2+M_k^2)}, \tag{9.46}$$

and the quadratic part of the actions (9.32) and (9.41) become proportional.

Fermions. Note that the fermion inverse propagator (9.39) can be written as

$$\Delta_{\mathrm{F}}^{-1}(p)=(m+i\not{p})\prod_{i=1}^{n}(1+i\not{p}/M_i)(1-i\not{p}/M_i).$$

This indicates that again the same form can be obtained by a set of regulator fields $\bar\psi_{i\pm}\psi_{i\pm}$. One replaces the action (9.37) by

$$\mathcal{S}_{0\,\mathrm{reg}}=\int \mathrm{d}^d x\,\bar\psi(x)(\not{\partial}+m)\psi(x)+\sum_{k,\epsilon=\pm}\frac{1}{z_{k\epsilon}}\int \mathrm{d}^d x\,\bar\psi_{k\epsilon}(x)(\not{\partial}+\epsilon M_k)\psi_{k\epsilon}(x). \tag{9.47}$$

In the same way, in the interaction the fields ψ and $\bar\psi$ are replaced by the sums

$$\psi\mapsto\psi+\sum_{k,\epsilon}\psi_{k\epsilon}\,,\quad \bar\psi\mapsto\bar\psi+\sum_{k,\epsilon}\bar\psi_{k\epsilon}\,.$$

Again, for a proper choice of the constants z_k, after integration over the regulator fields the form (9.39) is recovered. Note, in particular, 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

$$\int \mathrm{d}^d x\,\bar\psi(x)\,(\not{\partial}\otimes\mathbf{1}+M\mathbf{1}\otimes\sigma_3)\,\psi(x),$$

where the first matrix $\mathbf{1}$ and the Pauli matrix σ_3 act in $\pm$ space. The spinors then transform like

$$\psi_\theta(x)=e^{i\theta\gamma_S\otimes\sigma_1}\psi(x),\qquad \bar\psi_\theta(x)=\bar\psi(x)\,e^{i\theta\gamma_S\otimes\sigma_1}\,,$$

because σ_1 anticommutes with σ_3.

Remarks. A potential weakness of these momentum cut-off methods has 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 (7.22)), as

$$\mathcal{Z}(J)=\det{}^{1/2}(\Delta_{\mathrm{B}})\exp\left[-\mathcal{V}_{\mathrm{I}}\left(\delta/\delta J\right)\right]\exp\left(\frac{1}{2}\int \mathrm{d}^d x\,\mathrm{d}^d y\,J(x)\Delta_{\mathrm{B}}(x-y)J(y)\right),$$

where the determinant is generated by the gaussian integration. None of these methods can regularize 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 3.2): a class of Feynman diagrams then cannot be regularized by this method. Quantum field theories where the problem occurs include models with non-linear (see Chapter 14) or gauge symmetries (Chapter 19).

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.

9.6 Lattice Regularization

We shall meet examples where Pauli–Villars's regularization does not work: field theories in which the action has a definite geometric character like models on homogeneous spaces (for example, the non-linear σ-model) or gauge theories. In these theories some divergences are related to the problem of quantization and order of operators. As we have shown in Chapter 3, they already appear at the level of simple quantum mechanics. In particular, the forms of the propagator and of the interaction terms are not independent. When the propagator is regularized, new more singular interactions have to be added to the action to preserve some symmetry and, as we shall show in examples, some one-loop diagrams cannot be regularized. Other regularization methods are needed. In many cases lattice regularization, which we have introduced in Section 7.1.2 to define precisely the functional integral, can be used. 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. For instance the regularized functional integral can be calculated by numerical methods, like stochastic methods (Monte Carlo type simulations) or strong coupling expansions.

(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 will be argued in Chapters 24 and 25) and fermion chirality, which turns out to be a more difficult problem.

One obvious disadvantage is that it leads to very cumbersome perturbative calculations.

9.6.1 Scalar field theories

To regularize the scalar field theory with field $\phi(x)$ and action (9.31) 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,

$$\partial_\mu \phi \mapsto \nabla_\mu^{\mathrm{lat}} \phi = \frac{1}{a}\left[\phi(x+an_\mu) - \phi(x)\right],$$

where x is a lattice site, a the lattice spacing and n_μ the unit vector in the μ direction.

The regularized lattice action then takes the form

$$\mathcal{S}_{\mathrm{reg}}(\phi) = a^d \sum_{x\in(a\mathbb{Z})^d} \left[\frac{1}{2}\sum_{\mu=1}^{d} \left[\nabla_\mu^{\mathrm{lat}}\phi(x)\right]^2 + \frac{1}{2}m^2\phi^2(x) + V\left(\phi(x)\right)\right]. \tag{9.48}$$

The Fourier transform $\tilde{\phi}(p)$ of the field is now a periodic function of cyclic momentum variables p_μ:

$$\tilde{\phi}(p) = \left(\frac{a}{2\pi}\right)^d \sum_{\mathbf{x}\in a\mathbb{Z}^d} \phi(x)\,\mathrm{e}^{-i\mathbf{p}\cdot\mathbf{x}} \Leftrightarrow \phi(x) = \int \mathrm{d}^d p\,\tilde{\phi}(p)\,\mathrm{e}^{i\mathbf{p}\cdot\mathbf{x}}, \tag{9.49}$$

which can, therefore, be restricted to a Brillouin zone, for example,

$$-\pi/a \le p_\mu < \pi/a; \quad \mu = 1,\ldots,d.$$

The corresponding propagator $\Delta_{\mathrm{B}}(p)$ is then given by

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + \frac{2}{a^2}\sum_{\mu=1}^{d}\bigl(1 - \cos(ap_\mu)\bigr). \tag{9.50}$$

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:

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + p^2 - \frac{1}{12}\sum_{\mu} a^2 p_\mu^4 + O\left(p_\mu^6\right). \tag{9.51}$$

We note that hypercubic symmetry implies $O(d)$ symmetry at order p^2.

9.6.2 Fermions and the doubling problem

Let us briefly indicate a few problems arising when field theories also involve fermions.

To regularize the free fermion action (9.37) by the lattice method we can replace $\bar\psi(x)\not\partial\psi(x)$ by

$$\bar\psi(x)\gamma_\mu\nabla_\mu^{\mathrm{lat}}\psi(x) \equiv \bar\psi(x)\gamma_\mu\left[\psi(x+an_\mu) - \psi(x-an_\mu)\right]/2a\,.$$

Then, in the momentum representation, the fermion propagator becomes

$$\Delta_{\mathrm{F}}^{-1}(p) = m + i\sum_{\mu}\gamma_\mu\frac{\sin ap_\mu}{a}. \tag{9.52}$$

A problem then arises: the equations relevant to the small lattice spacing limit,

$$\sin(a\,p_\mu) = 0\,,$$

have each two solutions $p_\mu = 0$ and $p_\mu = \pi/a$ within one period, that is, within the Brillouin zone. Therefore, the propagator (9.52) propagates 2^d fermions.

This degeneracy can be removed by the addition to the regularized action of a term δS involving second derivatives:

$$\delta S(\bar\psi,\psi) = \frac{1}{2}M\sum_{x,\mu}\left[2\bar\psi(x)\psi(x) - \bar\psi(x+an_\mu)\psi(x) - \bar\psi(x)\psi(x+an_\mu)\right]. \tag{9.53}$$

After Fourier transformation the modified Dirac operator D_W reads

$$D_W(p) = m + M\sum_{\mu}(1-\cos ap_\mu) + \frac{i}{a}\sum_{\mu}\gamma_\mu\sin ap_\mu\,. \tag{9.54}$$

The fermion propagator becomes

$$\Delta_{\mathrm{F}}(p) = D_W^{\dagger}(p)\left[D_W(p)D_W^{\dagger}(p)\right]^{-1}$$

The corresponding propagator $\Delta_{\mathrm{B}}(p)$ is then given by

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + \frac{2}{a^2}\sum_{\mu=1}^{d}\left(1 - \cos(ap_\mu)\right). \tag{9.50}$$

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:

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + p^2 - \frac{1}{12}\sum_{\mu} a^2 p_\mu^4 + O\left(p_\mu^6\right). \tag{9.51}$$

We note that hypercubic symmetry implies $O(d)$ symmetry at order p^2.

9.6.2 Fermions and the doubling problem

Let us briefly indicate a few problems arising when field theories also involve fermions.

To regularize the free fermion action (9.37) by the lattice method we can replace $\bar{\psi}(x)\not{\partial}\psi(x)$ by

$$\bar{\psi}(x)\gamma_\mu\nabla_\mu^{\mathrm{lat}}\psi(x) \equiv \bar{\psi}(x)\gamma_\mu\left[\psi(x+an_\mu) - \psi(x-an_\mu)\right]/2a\,.$$

Then, in the momentum representation, the fermion propagator becomes

$$\Delta_{\mathrm{F}}^{-1}(p) = m + i\sum_{\mu}\gamma_\mu\frac{\sin ap_\mu}{a}. \tag{9.52}$$

A problem then arises: the equations relevant to the small lattice spacing limit,

$$\sin(a\,p_\mu) = 0\,,$$

have each two solutions $p_\mu = 0$ and $p_\mu = \pi/a$ within one period, that is, within the Brillouin zone. Therefore, the propagator (9.52) propagates 2^d fermions.

This degeneracy can be removed by the addition to the regularized action of a term δS involving second derivatives:

$$\delta S(\bar{\psi},\psi) = \frac{1}{2}M\sum_{x,\mu}\left[2\bar{\psi}(x)\psi(x) - \bar{\psi}(x+an_\mu)\psi(x) - \bar{\psi}(x)\psi(x+an_\mu)\right]. \tag{9.53}$$

After Fourier transformation the modified Dirac operator D_W reads

$$D_W(p) = m + M\sum_{\mu}(1 - \cos ap_\mu) + \frac{i}{a}\sum_{\mu}\gamma_\mu\sin ap_\mu\,. \tag{9.54}$$

The fermion propagator becomes

$$\Delta_{\mathrm{F}}(p) = D_W^\dagger(p)\left[D_W(p)D_W^\dagger(p)\right]^{-1}$$

The corresponding propagator $\Delta_{\mathrm{B}}(p)$ is then given by

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + \frac{2}{a^2}\sum_{\mu=1}^{d}\bigl(1-\cos(ap_\mu)\bigr). \tag{9.50}$$

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:

$$\Delta_{\mathrm{B}}^{-1}(p) = m^2 + p^2 - \frac{1}{12}\sum_\mu a^2 p_\mu^4 + O\left(p_\mu^6\right). \tag{9.51}$$

We note that hypercubic symmetry implies $O(d)$ symmetry at order p^2.

9.6.2 Fermions and the doubling problem

Let us briefly indicate a few problems arising when field theories also involve fermions.

To regularize the free fermion action (9.37) by the lattice method we can replace $\bar\psi(x)\not\partial\psi(x)$ by

$$\bar\psi(x)\gamma_\mu\nabla_\mu^{\mathrm{lat}}\psi(x) \equiv \bar\psi(x)\gamma_\mu\left[\psi(x+an_\mu)-\psi(x-an_\mu)\right]/2a\,.$$

Then, in the momentum representation, the fermion propagator becomes

$$\Delta_{\mathrm{F}}^{-1}(p) = m + i\sum_\mu \gamma_\mu\frac{\sin ap_\mu}{a}. \tag{9.52}$$

A problem then arises: the equations relevant to the small lattice spacing limit,

$$\sin(a\,p_\mu)=0\,,$$

have each two solutions $p_\mu = 0$ and $p_\mu = \pi/a$ within one period, that is, within the Brillouin zone. Therefore, the propagator (9.52) propagates 2^d fermions.

This degeneracy can be removed by the addition to the regularized action of a term δS involving second derivatives:

$$\delta S(\bar\psi,\psi) = \frac{1}{2}M\sum_{x,\mu}\left[2\bar\psi(x)\psi(x)-\bar\psi(x+an_\mu)\psi(x)-\bar\psi(x)\psi(x+an_\mu)\right]. \tag{9.53}$$

After Fourier transformation the modified Dirac operator D_W reads

$$D_W(p) = m + M\sum_\mu(1-\cos ap_\mu) + \frac{i}{a}\sum_\mu\gamma_\mu\sin ap_\mu\,. \tag{9.54}$$

The fermion propagator becomes

$$\Delta_{\mathrm{F}}(p) = D_W^\dagger(p)\left[D_W(p)D_W^\dagger(p)\right]^{-1}$$

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 lifted. 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 of Wilson's fermions.

However, a serious problem arises if one wants to construct a theory with massless fermions and preserve chiral symmetry (equation (8.27)):

$$\psi'(x) = e^{i\theta\gamma_S}\psi(x), \qquad \bar\psi'(x) = \bar\psi(x)\,e^{i\theta\gamma_S}\,.$$

This implies that the Dirac operator $\mathbf{D}$ must anticommute with γ_S:

$$\{\mathbf{D},\gamma_S\} = 0\,. \tag{9.55}$$

Then both the mass term and the term (9.53) are excluded. It remains possible to add various counter-terms and try to adjust them to recover chiral symmetry in the continuum limit. But there is no *a priori* guarantee that this is indeed possible and moreover calculations are plagued by fine tuning problems or cancellations of unnecessary UV divergences.

One can 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 form

$$\Delta_{\rm F}^{-1}(p) = \sum_\mu \gamma_\mu f_\mu(p),$$

in which $f_\mu(p)$ is a smooth function (singularities introduce violation of locality), periodic with period $2\pi/a$, which vanishes linearly for $|\mathbf{p}|$ small:

$$f_\mu(\mathbf{p}) \sim p_\mu, \quad \text{for } |\mathbf{p}| \to 0\,.$$

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. This doubling of the number of fermion degrees of freedom is related to the problem of anomalies (see Chapter 20).

Since the most naive form of the propagator yields 2^d fermion states, one tries in practical calculations to reduce this number to a smaller multiple of two. The idea of staggered fermions introduced by Kogut and Susskind is often used: first, by modifying the action one is able to decrease the multiplication factor from 2^d to $2^{d/2}$. Then the remaining degeneracy is 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.

9.6.3 Ginsparg–Wilson relation and overlap fermion

Notation. For convenience we now set the lattice spacing $a=1$ and use for the fields the notation $\psi(x) \equiv \psi_x$.

Ginsparg–Wilson relation and overlap fermions. Recently an important advance on the problem of chiral fermions has been achieved. It had been noted, many years ago, that a potential way to avoid the doubling problem, while still retaining chiral properties in the continuum limit, was to construct a lattice Dirac operator $\mathbf{D}$ satisfying, instead of (9.55), the Ginsparg–Wilson relation,

$$\{\mathbf{D}^{-1}, \gamma_S\} = \gamma_S\,, \tag{9.56}$$

where γ_S in the r.h.s. means $\gamma_S\delta_{xy}$, and δ_{xy} is the identity for lattice sites.

To find consistent solutions to the relation (9.56) however is not easy because the demands that both $\mathbf{D}$ and the anticommutator $\{\mathbf{D}^{-1}, \gamma_S\}$ should be local, are difficult to satisfy, specially in the most interesting case of a covariant operator in gauge theories.

Let us briefly explain the main idea. Using the relation, consequence of euclidean hermiticity and reflection symmetry,

$$\mathbf{D}^\dagger = \gamma_S\mathbf{D}\gamma_S\,,$$

one can rewrite the relation (9.56):

$$\mathbf{D}^{-1} + \left(\mathbf{D}^{-1}\right)^\dagger = \mathbf{1} \ \Rightarrow\ \mathbf{D} + \mathbf{D}^\dagger = \mathbf{D}\mathbf{D}^\dagger = \mathbf{D}^\dagger\mathbf{D}\,.$$

This implies that the lattice operator $\mathbf{D}$ has an index, and, in addition, the operator

$$\mathbf{S} = \mathbf{1} - \mathbf{D} \tag{9.57}$$

is unitary:

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{1}\,. \tag{9.58}$$

The eigenvalues of $\mathbf{S}$ lie on the unit circle, the eigenvalue one corresponding to the pole of the Dirac propagator.

Moreover,

$$\gamma_S\mathbf{S} = \mathbf{S}^\dagger\gamma_S\,, \quad (\gamma_S\mathbf{S})^2 = \mathbf{1}\,. \tag{9.59}$$

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.

An explicit solution can be derived from a Wilson–Dirac operator without doublers like $\mathbf{D}_\mathrm{W}$ in equation (9.54). Setting

$$\mathbf{A} = \mathbf{1} - \mathbf{D}_\mathrm{W}\,,$$

one takes

$$\mathbf{S} = \mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2} \ \Rightarrow\ \mathbf{D} = \mathbf{1} - \mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2}. \tag{9.60}$$

With this ansatz $\mathbf{D}$ has a zero eigenmode when $\mathbf{A}\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1/2}$ has the eigenvalue one. 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 (9.54) 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 wanted $p_\mu = 0$ mode, but eliminates all doublers which then correspond to the eigenvalue two for $\mathbf{D}$, and the doubling problem is, in principle, solved.

that a potential way to avoid the doubling problem, while still retaining chiral properties in the continuum limit, was to construct a lattice Dirac operator $\mathbf{D}$ satisfying, instead of (9.55), the Ginsparg–Wilson relation,

$$\{\mathbf{D}^{-1}, \gamma_S\} = \gamma_S \,, \tag{9.56}$$

where γ_S in the r.h.s. means $\gamma_S \delta_{xy}$, and δ_{xy} is the identity for lattice sites.

To find consistent solutions to the relation (9.56) however is not easy because the demands that both $\mathbf{D}$ and the anticommutator $\{\mathbf{D}^{-1}, \gamma_S\}$ should be local, are difficult to satisfy, specially in the most interesting case of a covariant operator in gauge theories.

Let us briefly explain the main idea. Using the relation, consequence of euclidean hermiticity and reflection symmetry,

$$\mathbf{D}^\dagger = \gamma_S \mathbf{D} \gamma_S \,,$$

one can rewrite the relation (9.56):

$$\mathbf{D}^{-1} + \left(\mathbf{D}^{-1}\right)^\dagger = \mathbf{1} \;\Rightarrow\; \mathbf{D} + \mathbf{D}^\dagger = \mathbf{D}\mathbf{D}^\dagger = \mathbf{D}^\dagger \mathbf{D} \,.$$

This implies that the lattice operator $\mathbf{D}$ has an index, and, in addition, the operator

$$\mathbf{S} = \mathbf{1} - \mathbf{D} \tag{9.57}$$

is unitary:

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{1} \,. \tag{9.58}$$

The eigenvalues of $\mathbf{S}$ lie on the unit circle, the eigenvalue one corresponding to the pole of the Dirac propagator.

Moreover,

$$\gamma_S \mathbf{S} = \mathbf{S}^\dagger \gamma_S \,, \quad (\gamma_S \mathbf{S})^2 = \mathbf{1} \,. \tag{9.59}$$

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.

An explicit solution can be derived from a Wilson–Dirac operator without doublers like $\mathbf{D}_\mathrm{W}$ in equation (9.54). Setting

$$\mathbf{A} = \mathbf{1} - \mathbf{D}_\mathrm{W} \,,$$

one takes

$$\mathbf{S} = \mathbf{A}\left(\mathbf{A}^\dagger \mathbf{A}\right)^{-1/2} \;\Rightarrow\; \mathbf{D} = \mathbf{1} - \mathbf{A}\left(\mathbf{A}^\dagger \mathbf{A}\right)^{-1/2} . \tag{9.60}$$

With this ansatz $\mathbf{D}$ has a zero eigenmode when $\mathbf{A}\left(\mathbf{A}^\dagger \mathbf{A}\right)^{-1/2}$ has the eigenvalue one. 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 (9.54) 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 wanted $p_\mu = 0$ mode, but eliminates all doublers which then correspond to the eigenvalue two for $\mathbf{D}$, and the doubling problem is, in principle, solved.

It is then possible to construct lattice actions that have a chiral symmetry that corresponds to local but non-point-like transformations of the infinitesimal form

$$\psi'_x = \psi_x + \sum_y (\boldsymbol{\theta}\gamma_S \mathbf{S})_{xy}\,\psi_y\,, \quad \bar\psi'_x = \bar\psi_x + \bar\psi_x \boldsymbol{\theta}\gamma_S\,, \tag{9.61}$$

where $\boldsymbol{\theta}$ commutes with $\mathbf{D}$. Note that $\gamma_S \mathbf{S}$ plays for ψ the role of γ_S for $\bar\psi$, $\frac{1}{2}(\mathbf{1} \pm \gamma_S \mathbf{S})\psi$ 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 $\prod_x \mathrm{d}\psi_x \mathrm{d}\bar\psi_x$ automatically invariant. Indeed, the infinitesimal change of variables $\psi, \bar\psi \mapsto \psi', \bar\psi'$

$$\frac{\partial \psi'_x}{\partial \psi_y} = \delta_{xy} + \boldsymbol{\theta}\,\gamma_S\,\mathbf{S}_{xy}\,, \quad \frac{\partial \bar\psi'_x}{\partial \bar\psi_y} = \delta_{xy}\,(\mathbf{1} + \boldsymbol{\theta}\gamma_S)\,,$$

leads to the jacobian J:

$$\ln J \sim \operatorname{tr} \gamma_S\, \boldsymbol{\theta} \sum_x (2 - \mathbf{D}_{xx})\,.$$

This leaves the possibility of generating the expected anomalies (see Section 20.3), when the Wilson–Dirac operator is a covariant operator in the background of a gauge field.

Finally, let us stress that, if it seems that the doubling problem has been solved from the formal point of view, from the numerical point of view the calculation of the operator $(\mathbf{A}^\dagger \mathbf{A})^{-1/2}$ in a gauge background represents a major challenge.

9.6 Dimensional Regularization

Dimensional regularization involves continuation of Feynman diagrams in the parameter d (d is the space dimension) to arbitrary complex values and, therefore, seems to have no meaning outside perturbation theory. It is very often used because it leads to simpler perturbative calculations than other methods. In addition, it solves with the problem of commutation of quantum operators in local field theories. Indeed commutators, for example, in the case of a scalar field take the form (6.55)

$$[\hat\phi(x), \hat\pi(y)] = i\hbar\,\delta^{d-1}(x-y) = i\hbar(2\pi)^{1-d} \int \mathrm{d}^{d-1}p\; \mathrm{e}^{ip(x-y)}\,,$$

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 $d > 1$) proportional to

$$\delta^{d-1}(0) = (2\pi)^{1-d} \int \mathrm{d}^{d-1}p\,.$$

The rules of dimensional regularization imply the consistency of the change of variables $p \mapsto \lambda p$ and thus $\int \mathrm{d}^d p = 0$, in contrast with momentum regularization where it is proportional to a power of the cut-off. Therefore, the order between quantum operators becomes irrelevant because the commutator vanishes. Dimensional regularization is thus specially useful for geometric models where these problems of quantization occur, like non-linear σ-models whose hamiltonians have the generic form (3.20) (see Chapters 14, 15), or gauge theories.

It is not applicable when some essential property of the field theory is specific to the initial dimension. An example is provided by violations of parity symmetry involving the complete antisymmetric tensor $\epsilon_{\mu_1\cdots\mu_d}$, for example through the relation between γ_S (identical to γ_5 in four dimensions) and the other γ matrices.

Its use also requires some care in massless theories because its rules may lead to unwanted cancellations between UV and IR logarithmic divergences.

We now define an integral in d dimensions by a set of conditions which, if d is an integer, lead to the usual integral.

9.6.1 Defining properties of d-dimensional integrals

We define the dimensional continuation of integrals by the three conditions:

$$\begin{array}{llr} \text{(i)} & \int \mathrm{d}^d p\, F(p+q) = \int \mathrm{d}^d p\, F(p) & \text{translation} \\ \text{(ii)} & \int \mathrm{d}^d p\, F(\lambda p) = |\lambda|^{-d} \int \mathrm{d}^d p\, F(p) & \text{dilatation} \\ \text{(iii)} & \int \mathrm{d}^d p\, \mathrm{d}^{d'} q\, f(p) g(q) = \int \mathrm{d}^d p\, f(p) \int \mathrm{d}^{d'} q\, g(q) & \text{factorization} \end{array}$$

We now show that these simple rules provide a dimensional continuation to Feynman diagrams.

From the property (iii) we derive

$$\int \mathrm{d}^d p\, \mathrm{e}^{-tp^2} = \left(\int \mathrm{d}p_1\, \mathrm{e}^{-tp_1^2}\right)^d = \left(\frac{\pi}{t}\right)^{d/2}. \tag{9.62}$$

We can then write any scalar propagator $\Delta(\mathbf{p})$, including regularized propagators (9.34), as a Laplace transform:

$$\Delta(\mathbf{p}) = \int_0^\infty \mathrm{d}t\, \rho(t)\, \mathrm{e}^{-(\mathbf{p}^2+m^2)t}. \tag{9.63}$$

In particular for $\rho(t) = 1$, one recovers the un-regularized propagator $1/(\mathbf{p}^2 + m^2)$.

Using the representation (9.63) and the property that the momentum associated with an internal line is a linear combination of loop and external momenta, we can write a scalar Feynman diagram γ with constant vertices in the form

$$\begin{aligned} I_\gamma(\mathbf{p}) = (2\pi)^{-Ld} \int \prod_{i=1}^{I} \mathrm{d}t_i\, \rho_i(t_i) \prod_{\ell=1}^{L} \mathrm{d}^d q_\ell\, \exp\Bigg[&-\sum_1^L \mathbf{q}_\ell \cdot \mathbf{q}_{\ell'} M_{\ell\ell'}(t_i) \\ &-2\sum_1^L \mathbf{q}_\ell \cdot \mathbf{k}_\ell(\mathbf{p}, t_i) - S(\mathbf{p}, t_i) \Bigg]. \end{aligned} \tag{9.64}$$

The properties (i), (ii) and equation (9.62) allow an integration over all loop momenta q_ℓ and we find

$$I_\gamma(\mathbf{p}) = (4\pi)^{-Ld/2} \int \prod_{i=1}^{I} \mathrm{d}t_i\, \rho_i(t_i)\, (\det \mathbf{M})^{-d/2} \exp\left[\sum_1^L \mathbf{k}_\ell \left(M^{-1}\right)_{\ell\ell'} \mathbf{k}_{\ell'} - S(\mathbf{p}, t_i)\right]. \tag{9.65}$$

In this expression, the dependence in d is now explicit, and, therefore, continuation in d (at generic momenta if the theory is massless) can be achieved.

Examples: one-loop contributions.

(i) The one-loop contribution to the two-point function in the massive ϕ^4 field theory is proportional to

$$\begin{aligned}
I_\gamma(\mathbf{p}) &= \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{\mathbf{q}^2+m^2} \\
&= \int_0^\infty \mathrm{d}t \frac{\mathrm{d}^d q}{(2\pi)^d}\, \mathrm{e}^{-t(\mathbf{q}^2+m^2)} = \frac{1}{(4\pi)^{d/2}} \int_0^\infty \mathrm{d}t\, t^{-d/2}\, \mathrm{e}^{-m^2 t} \\
&= \frac{1}{(4\pi)^{d/2}} \Gamma(1-d/2) m^{d-2}.
\end{aligned}$$

This expression has poles for $d = 2+2n$, $n \ge 0$, 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 $d = 2$.

The one-loop contribution to the two-point function in the massless ($m = 0$) ϕ^3 field theory is proportional to

$$\begin{aligned}
I_\gamma(\mathbf{p}) &= \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{\mathbf{q}^2\,(\mathbf{p}+\mathbf{q})^2} \\
&= \int_0^\infty \mathrm{d}t_1 \mathrm{d}t_2 \frac{\mathrm{d}^d q}{(2\pi)^d}\, \mathrm{e}^{-t_1\mathbf{q}^2 - t_2(\mathbf{p}+\mathbf{q})^2} \\
&= \frac{\pi^{d/2}}{(2\pi)^d} \int_0^\infty \mathrm{d}t_1 \mathrm{d}t_2\, (t_1+t_2)^{-d/2}\, \mathrm{e}^{-t_1 t_2 \mathbf{p}^2/(t_1+t_2)} .
\end{aligned}$$

For $\mathbf{p} \ne 0$, the integral converges for $2 < d < 4$.

We can complete the calculation by setting

$$t_1 = ts\,, \qquad t_2 = (1-t)\,s\,.$$

The integral becomes

$$I_\gamma = \frac{1}{(4\pi)^{d/2}} \int_0^1 \mathrm{d}t \int_0^\infty \mathrm{d}s\, s^{1-d/2}\, \mathrm{e}^{-st(1-t)\mathbf{p}^2} .$$

We integrate over s:

$$I_\gamma = \frac{1}{(4\pi)^{d/2}} \left(\mathbf{p}^2\right)^{(d/2)-2} \Gamma(2-d/2) \int_0^1 \mathrm{d}t\, [t\,(1-t)]^{(d/2)-2} ,$$

and thus, finally,

$$I_\gamma = \frac{1}{(4\pi)^{d/2}} \Gamma(2-d/2) \frac{\Gamma^2((d/2)-1)}{\Gamma(d-2)} \left(\mathbf{p}^2\right)^{(d/2)-2} . \tag{9.66}$$

This expression has a pole at $d = 2$ corresponding to IR (low momentum) singularities because the theory is massless and has poles at $d = 4, 6, \ldots$ which clearly are consequences of the UV (large momentum) divergences of the Feynman diagram.

It is interesting to explain in this example the interplay between dimensional continuation and cut-off regularization. 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

$$I_\gamma \sim \frac{1}{8\pi^2 (4-d)} \left[\left(p^2\right)^{(d/2)-2} - \Lambda^{d-4} \right] \quad \text{for } d \to 4\,.$$

If, at d fixed, $d < 4$, we sent the cut-off to infinity we obtain the continuation of the initial diagram with a pole at $d = 4$. If at cut-off fixed we take the limit $d = 4$, we get a finite result in which $\ln\Lambda$ has replaced the pole at $d = 4$.

Important remarks

The property (ii) implies

$$\int \frac{\mathrm{d}^d p}{p^{2n}} = 0 \tag{9.67}$$

for these integrals that exist for no value of d, and leads, for $n = 0$, to the commutation of operators in local products.

Dimensional regularization in fact is not only a regularization since it performs a partial renormalization, suppressing divergences that in momentum regularization are powers of the cut-off.

Let us also point out one dangerous consequence. The result (9.67) can be interpreted for $d \to 2n$ as a cancellation between a UV and IR divergence:

$$\int \frac{\mathrm{d}^d p}{p^{2n}} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \left[\int_1^\infty p^{d-1-2n} \mathrm{d}p + \int_0^1 p^{d-1-2n} \mathrm{d}p \right].$$

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.

A simple example allows to verify the consistency (guaranteed *a priori* by the consistency of the defining rules) of the property (9.67) of dimensional regularization, which is at first sight somewhat strange. The integral

$$I = \int \frac{\mathrm{d}^d p}{p^2 \left(p^2 + 1\right)},$$

can be calculated in two ways: first, one notes

$$I = \int \mathrm{d}^d p \left(\frac{1}{p^2} - \frac{1}{p^2+1} \right) = -\int \frac{\mathrm{d}^d p}{p^2+1}.$$

This yields

$$I = -\pi^{d/2}\Gamma(1-d/2).$$

Second, one uses the transformation (9.63) on the initial expression. After a simple calculation one obtains the same result.

Continuation of tensor structures. Up to now we have 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,

$$\int \mathrm{d}^d q\, q_\mu q_\nu f(q^2, p^2, p\cdot q) = A(p^2) p_\mu p_\nu + B(p^2)\delta_{\mu\nu}\,. \tag{9.68}$$

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

$$I_\gamma \sim \frac{1}{8\pi^2 (4-d)} \left[\left(p^2\right)^{(d/2)-2} - \Lambda^{d-4} \right] \quad \text{for } d \to 4\,.$$

If, at d fixed, $d < 4$, we sent the cut-off to infinity we obtain the continuation of the initial diagram with a pole at $d = 4$. If at cut-off fixed we take the limit $d = 4$, we get a finite result in which $\ln \Lambda$ has replaced the pole at $d = 4$.

Important remarks

The property (ii) implies

$$\int \frac{\mathrm{d}^d p}{p^{2n}} = 0 \tag{9.67}$$

for these integrals that exist for no value of d, and leads, for $n = 0$, to the commutation of operators in local products.

Dimensional regularization in fact is not only a regularization since it performs a partial renormalization, suppressing divergences that in momentum regularization are powers of the cut-off.

Let us also point out one dangerous consequence. The result (9.67) can be interpreted for $d \to 2n$ as a cancellation between a UV and IR divergence:

$$\int \frac{\mathrm{d}^d p}{p^{2n}} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \left[\int_1^\infty p^{d-1-2n} \mathrm{d}p + \int_0^1 p^{d-1-2n} \mathrm{d}p \right].$$

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.

A simple example allows to verify the consistency (guaranteed *a priori* by the consistency of the defining rules) of the property (9.67) of dimensional regularization, which is at first sight somewhat strange. The integral

$$I = \int \frac{\mathrm{d}^d p}{p^2 \left(p^2 + 1\right)},$$

can be calculated in two ways: first, one notes

$$I = \int \mathrm{d}^d p \left(\frac{1}{p^2} - \frac{1}{p^2+1} \right) = - \int \frac{\mathrm{d}^d p}{p^2 + 1}.$$

This yields

$$I = -\pi^{d/2} \Gamma(1 - d/2).$$

Second, one uses the transformation (9.63) on the initial expression. After a simple calculation one obtains the same result.

Continuation of tensor structures. Up to now we have 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,

$$\int \mathrm{d}^d q \, q_\mu q_\nu f(q^2, p^2, p \cdot q) = A(p^2) p_\mu p_\nu + B(p^2) \delta_{\mu\nu}\,. \tag{9.68}$$

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:

$$\begin{aligned} A(p^2) &= \frac{1}{d-1}\frac{1}{(p^2)^2}\int \mathrm{d}^d q\, \left[d(p\cdot q)^2 - p^2 q^2\right] f(q^2,p^2,p\cdot q), \\ B(p^2) &= \frac{1}{d-1}\frac{1}{p^2}\int \mathrm{d}^d q\, \left[-(p\cdot q)^2 + p^2 q^2\right] f(q^2,p^2,p\cdot q). \end{aligned} \tag{9.69}$$

We have reduced the problem to the calculation of integrals of the form (9.64) with additional factors polynomial in momenta. The integration over momenta can then also be performed to yield the continuation in d.

9.6.2 Fermions

For fermions belonging to the fundamental representation of the spin group Spin(d) the strategy is similar. 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 have shown in Appendix A8.3 that no algebraic manipulation depends on the explicit value of the trace. Thus any smooth continuation in the neighbourhood of the relevant dimension will be satisfactory. A convenient choice, which we shall 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 Section 9.6.2)

$$d!\gamma_S = i^{-d/2}\epsilon_{\mu_1\cdots\mu_d}\gamma_{\mu_1}\cdots\gamma_{\mu_d}\,.$$

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

$$4!\gamma_5 = -e_{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\,.$$

It is easy then to verify that, with this definition, γ_5 anticommutes with all other γ_μ matrices only in 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 traces 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 $SO(d)$ invariance.

9.6.3 Dimensional regularization and UV divergences

When the dimension d approaches the initial dimension, Feynman diagrams become singular as a consequence of the original UV divergences. The singular contributions take the form of poles and can be isolated by performing a Laurent expansion of the diagram. For example, the expression (9.66) is the value of a Feynman diagram of the massless ϕ^4 field theory which is renormalizable in $d=4$. The Laurent expansion is

$$I_\gamma = N_d \left[\frac{1}{4-d} + \frac{1}{2} - \frac{1}{2}\ln p^2 + O\left(d-4\right)\right].$$

As we have implicitly done above, we shall in general include a factor

$$N_d = \frac{2\pi^{d/2}}{(2\pi)^d \Gamma(d/2)},$$

product of the surface of the sphere S_{d-1} by $(2\pi)^{-d}$, 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 would appear in a cut-off regularization in the large Λ limit, are replaced by powers of $1/(d-4)$. However, as a consequence of identity (9.67), no divergent contribution equivalent to a power of Λ can appear, and in this sense dimensional regularization already performs a partial renormalization.

Bibliographical Notes

For an early reference on power counting see

N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience, New York 1959).

For the Yang–Lee edge singularity see

C.N. Yang and T.D. Lee, *Phys. Rev.* 87 (1952) 404; T.D. Lee and C.N. Yang, *Phys. Rev.* 87 (1952) 410; M.E. Fisher, *Phys. Rev. Lett.* 40 (1978) 1610.

An early description of regularization methods involving modified propagators can be found in

W. Pauli and F. Villars, *Rev. Mod. Phys.* 21 (1949) 434.

For the proper time representation see

J. Schwinger, *Phys. Rev.* 82 (1951) 664.

The consistency of the lattice regularization is proven (except for theories with chiral fermions) in

T. Reisz, *Commun. Math. Phys.* 117 (1988) 79, 639.

The doubling phenomenon for lattice fermions has been proven quite generally by

H.B. Nielsen and M. Ninomiya, *Nucl. Phys.* B185 (1981) 20.

Wilson's solution to the fermion doubling problem is described in

K.G. Wilson in *New Phenomena in Subnuclear Physics*, Erice 1975, A. Zichichi ed. (Plenum, New York 1977).

Staggered fermions have been proposed in

T. Banks, L. Susskind and J. Kogut, *Phys. Rev.* D13 (1976) 1043.

Discussion of solutions of the Ginsparg–Wilson relation

P.H. Ginsparg and K.G. Wilson, *Phys. Rev.* D25 (1982) 2649,

and consequences can be found in

H. Neuberger, *Phys. Lett.* B417 (1998) 141, hep-lat/9707022, *ibidem* B427 (1998) 353, hep-lat/9801031; P. Hasenfratz, V. Laliena, F. Niedermayer, *Phys. Lett.* B427 (1998) 125, hep-lat/9801021; M. Lüscher, *Phys. Lett.* B428 (1998) 342, hep-lat/9802011, hep-lat/9811032, hep-th/0102028; J. Zinn-Justin, hep-th/0201220.

Dimensional regularization has been introduced by

J. Ashmore, *Lett. Nuovo Cimento* 4 (1972) 289; G. 't Hooft and M. Veltman, *Nucl. Phys.* B44 (1972) 189; C.G. Bollini and J.J. Giambiaggi, *Phys. Lett.* 40B (1972) 566, *Nuovo Cimento* 12B (1972).

Its use in problems with chiral anomalies has been proposed in

D.A. Akyeampong and R. Delbourgo, *Nuovo Cimento* 17A (1973) 578.

For an early review see

G. Leibbrandt, *Rev. Mod. Phys.* 47 (1975) 849.

For other schemes see also

E.R. Speer in *Renormalization Theory*, Erice 1975, G. Velo and A.S. Wightman eds. (D. Reidel, Dordrecht, Holland 1976).

H. Neuberger, *Phys. Lett.* B417 (1998) 141, hep-lat/9707022, *ibidem* B427 (1998) 353, hep-lat/9801031; P. Hasenfratz, V. Laliena, F. Niedermayer, *Phys. Lett.* B427 (1998) 125, hep-lat/9801021; M. Lüscher, *Phys. Lett.* B428 (1998) 342, hep-lat/9802011, hep-lat/9811032, hep-th/0102028; J. Zinn-Justin, hep-th/0201220.

Dimensional regularization has been introduced by

J. Ashmore, *Lett. Nuovo Cimento* 4 (1972) 289; G. 't Hooft and M. Veltman, *Nucl. Phys.* B44 (1972) 189; C.G. Bollini and J.J. Giambiaggi, *Phys. Lett.* 40B (1972) 566, *Nuovo Cimento* 12B (1972).

Its use in problems with chiral anomalies has been proposed in

D.A. Akyeampong and R. Delbourgo, *Nuovo Cimento* 17A (1973) 578.

For an early review see

G. Leibbrandt, *Rev. Mod. Phys.* 47 (1975) 849.

For other schemes see also

E.R. Speer in *Renormalization Theory*, Erice 1975, G. Velo and A.S. Wightman eds. (D. Reidel, Dordrecht, Holland 1976).

APPENDIX A9
SCHWINGER'S PROPER TIME REPRESENTATION

A9.1 Schwinger's Proper Time Representation

We first establish a formal representation of the one-loop contribution to the generating functional of proper vertices $\Gamma(\varphi)$. In Section 7.7 we have shown that $\Gamma_{1\,\text{loop}}(\varphi)$ is given by

$$\Gamma_{1\,\text{loop}}(\varphi) = \frac{1}{2} \operatorname{tr} \left[\ln \frac{\delta^2 \mathcal{S}}{\delta\varphi(x_1)\delta\varphi(x_2)} - \ln \left.\frac{\delta^2 \mathcal{S}}{\delta\varphi(x_1)\delta\varphi(x_2)}\right|_{\varphi=0} \right]. \tag{A9.1}$$

For example, if $\mathcal{S}(\phi)$ is

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2} \left(\partial_\mu \phi\right)^2 + \tfrac{1}{2} m^2 \phi^2 + U\left(\phi(x)\right) \right], \tag{A9.2}$$

where $U(\phi)$ is a polynomial, the second derivative $\delta^2 \mathcal{S}/\delta\varphi(x_1)\delta\varphi(x_2)$ takes the form of a quantum hamiltonian:

$$M(x_1, x_2) \equiv \frac{\delta^2 \mathcal{S}}{\delta\varphi(x_1)\delta\varphi(x_2)} = \left[-\nabla^2 + m^2 + U''\left(\varphi(x_1)\right)\right] \delta(x_1 - x_2). \tag{A9.3}$$

We also define M_0:

$$M_0(x_1, x_2) \equiv \left.\frac{\delta^2 \mathcal{S}}{\delta\varphi\delta\varphi}\right|_{\varphi=0} = \left(-\nabla^2 + m^2\right) \delta(x_1 - x_2). \tag{A9.4}$$

The general identity,

$$\operatorname{tr}\left(\ln M - \ln M_0\right) = -\int_0^\infty \frac{\mathrm{d}t}{t} \operatorname{tr}\left(\mathrm{e}^{-tM} - \mathrm{e}^{-tM_0}\right), \tag{A9.5}$$

then leads to a compact representation of the one-loop functional $\Gamma_{1\,\text{loop}}(\varphi)$ as an integral over Schwinger's proper time:

$$\Gamma_{1\,\text{loop}}(\varphi) = -\frac{1}{2} \int_0^\infty \frac{\mathrm{d}t}{t} \operatorname{tr}\left(\mathrm{e}^{-tM} - \mathrm{e}^{-tM_0}\right). \tag{A9.6}$$

The methods which are used in quantum mechanics to calculate the statistical operator $\mathrm{e}^{-\beta H}$ can again be used here.

A9.2 One-Loop Divergences

In the representation ($A9.6$) 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 $\langle x|\, \mathrm{e}^{-tM}\, |x\rangle$ for a Schrödinger-like operator M. This is a problem we have already faced in Section 2.2, and which can be solved for instance with Schrödinger's equation.

The expression ($A9.6$) 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 of equation (9.34).

Schwinger's proper time regularization. Schwinger's proper time regularization, which we have mentioned in Section 9.5, 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

$$\Gamma^{\mathrm{reg}}_{1-\mathrm{loop}}(\varphi) = -\frac{1}{2}\int_{\varepsilon}^{\infty} \frac{\mathrm{d}t}{t}\, \mathrm{e}^{-m^2 t}\, \mathrm{tr}\left(\mathrm{e}^{-tH} - \mathrm{e}^{-tH_0}\right). \tag{A9.7}$$

Let us, for illustration purposes, expand up to order t^3 in the case of the action ($A9.2$). Setting

$$U''\left(\phi(x)\right) = V(x),$$

we have to expand the solution of the Schrödinger equation:

$$\left[-\nabla_x^2 + V(x)\right] \langle x|\, \mathrm{e}^{-tH}\, |x'\rangle = -\frac{\partial}{\partial t} \langle x|\, \mathrm{e}^{-tH}\, |x'\rangle\,. \tag{A9.8}$$

We set

$$\langle x|\, \mathrm{e}^{-tH}\, |x'\rangle = \mathrm{e}^{-\sigma(x,x';t)}\,. \tag{A9.9}$$

The Schrödinger equation then takes the form

$$\nabla^2\sigma - \left(\partial_\mu\sigma\right)^2 + V(x) = \frac{\partial\sigma}{\partial t}. \tag{A9.10}$$

The function σ has for $t \to 0$ an expansion of the form

$$\sigma = \frac{1}{4t}(x-x')^2 + \frac{d}{2}\ln 4\pi t + At + Bt^2 + Ct^3 + O\left(t^4\right). \tag{A9.11}$$

We obtain equations for the coefficients A, B and C:

$$\begin{aligned} A + (x-x')_\mu \partial_\mu A &= V(x), \\ 2B + (x-x')_\mu \partial_\mu B &= \nabla^2 A\,, \\ 3C + (x-x')_\mu \partial_\mu C &= \nabla^2 B - \left(\partial_\mu A\right)^2. \end{aligned} \tag{A9.12}$$

The solutions A, B are

$$\begin{aligned} A(x,x') &= \int_0^1 \mathrm{d}s\, V\left(x' + s(x-x')\right), \\ B(x,x') &= \int_0^1 \mathrm{d}s\, s(1-s)\nabla^2 V\left(x' + s(x-x')\right). \end{aligned} \tag{A9.13}$$

It follows that

$$A(x,x) = V(x), \quad B(x,x) = \tfrac{1}{6}\nabla^2 V(x), \quad C(x,x) = -\tfrac{1}{2}\left(\partial_\mu V(x)\right)^2 + \tfrac{1}{20}\nabla^4 V(x).$$

The divergent part of the regularized expression ($A9.7$) 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 we find

$$I_\varepsilon = -\frac{1}{2}\frac{1}{(4\pi)^{d/2}}\int_\varepsilon \frac{\mathrm{d}t\,\mathrm{e}^{-m^2 t}}{t^{1+d/2}}\int \mathrm{d}^d x\left\{-V(x)t + \tfrac{1}{2}V^2(x)t^2 - \tfrac{1}{6}\left[V^3(x) + \tfrac{1}{2}\left(\partial_\mu V(x)\right)^2\right]t^3\right\} + O\left(t^4\right). \qquad (A9.14)$$

Keeping the contribution of the lower bound of the t integration, we finally obtain

$$I_\varepsilon = \frac{1}{2}\frac{1}{(4\pi)^{d/2}}\left\{-\frac{\varepsilon^{1-d/2}}{1-d/2}\int \mathrm{d}^d x\, V(x) + \frac{1}{2}\frac{\varepsilon^{2-d/2}}{2-d/2}\int \mathrm{d}^d x\left(V^2(x) + 2m^2 V(x)\right) - \frac{1}{6}\frac{\varepsilon^{3-d/2}}{3-d/2}\int \mathrm{d}^d x\left[V^3(x) + 3m^2 V^2(x) + 3m^4 V(x) + \tfrac{1}{2}\left(\partial_\mu V(x)\right)^2\right]\right\} + \cdots \qquad (A9.15)$$

When d is an even integer, $\varepsilon^0/0$ has to be replaced by $\ln(1/\varepsilon)$. This expression gives all divergences for $d \le 6$.

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,

$$\Gamma^{\text{div}}_{1\,\text{loop}}(\varphi) = \frac{1}{2^7\pi^3}\int \mathrm{d}^6 x\left\{\frac{\Lambda^4}{2}g\varphi(x) - \frac{\Lambda^2}{2}\left[g^2\varphi^2(x) + gm^2\varphi(x)\right] + \frac{1}{3}\ln\frac{\Lambda}{m}\left[g^3\varphi^3(x) + 3g^2m^2\varphi^2(x) + 3gm^4\varphi(x) + \frac{g^2}{2}\left(\partial_\mu\varphi(x)\right)^2\right]\right\}. \qquad (A9.16)$$

This leads to the results of equations (9.7,9.8).

For the interaction $U(\phi) = g\phi^4/4!$ in four dimensions we find

$$\Gamma^{\text{div}}_{1\,\text{loop}} = \frac{1}{32\pi^2}\left\{\frac{\Lambda^2}{2}g\int \mathrm{d}^4 x\,\varphi^2(x) - \ln\frac{\Lambda}{m}\int \mathrm{d}^4 x\left[\frac{g^2}{4}\varphi^4(x) + gm^2\varphi^2(x)\right]\right\}. \qquad (A9.17)$$

An identical expression will be recovered in Section 10.3, equation (10.24). Both in equations ($A9.16$) and ($A9.17$), we have defined the divergent part of $\Gamma(\varphi)$ as the sum of the divergent terms in the asymptotic expansion for Λ/m large.

Finally, we can also apply equation ($A9.15$) to $d = 2$ and a general interaction $U(\phi)$. We find

$$\Gamma^{\text{div}}_{1\,\text{loop}} = \frac{1}{4\pi}\ln\frac{\Lambda}{m}\int \mathrm{d}^2 x\, U''\left(\varphi(x)\right). \qquad (A9.18)$$

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 previous regularization method is to replace expression ($A9.6$) by

$$\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi) = -\frac{1}{2\Gamma(1+\sigma)}\int_0^\infty \mathrm{d}t\, t^{\sigma-1}\,\mathrm{tr}\left(\mathrm{e}^{-tM} - \mathrm{e}^{-tM_0}\right), \qquad (A9.19)$$

It follows that

$$A(x,x)=V(x),\quad B(x,x)=\tfrac{1}{6}\nabla^2 V(x),\quad C(x,x)=-\tfrac{1}{2}\left(\partial_\mu V(x)\right)^2+\tfrac{1}{20}\nabla^4 V(x).$$

The divergent part of the regularized expression ($A9.7$) 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 we find

$$I_\varepsilon=-\frac{1}{2}\frac{1}{(4\pi)^{d/2}}\int_\varepsilon\frac{\mathrm{d}t\,\mathrm{e}^{-m^2t}}{t^{1+d/2}}\int\mathrm{d}^dx\left\{-V(x)t+\tfrac{1}{2}V^2(x)t^2-\tfrac{1}{6}\left[V^3(x)+\tfrac{1}{2}\left(\partial_\mu V(x)\right)^2\right]t^3\right\}$$
$$+O\left(t^4\right). \qquad (A9.14)$$

Keeping the contribution of the lower bound of the t integration, we finally obtain

$$I_\varepsilon=\frac{1}{2}\frac{1}{(4\pi)^{d/2}}\left\{-\frac{\varepsilon^{1-d/2}}{1-d/2}\int\mathrm{d}^dx\,V(x)+\frac{1}{2}\frac{\varepsilon^{2-d/2}}{2-d/2}\int\mathrm{d}^dx\left(V^2(x)+2m^2V(x)\right)\right.$$
$$\left.-\frac{1}{6}\frac{\varepsilon^{3-d/2}}{3-d/2}\int\mathrm{d}^dx\left[V^3(x)+3m^2V^2(x)+3m^4V(x)+\tfrac{1}{2}\left(\partial_\mu V(x)\right)^2\right]\right\}+\cdots \qquad (A9.15)$$

When d is an even integer, $\varepsilon^0/0$ has to be replaced by $\ln(1/\varepsilon)$. This expression gives all divergences for $d\le 6$.

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,

$$\Gamma^{\text{div}}_{1\,\text{loop}}(\varphi)=\frac{1}{2^7\pi^3}\int\mathrm{d}^6x\left\{\frac{\Lambda^4}{2}g\varphi(x)-\frac{\Lambda^2}{2}\left[g^2\varphi^2(x)+gm^2\varphi(x)\right]\right.$$
$$\left.+\frac{1}{3}\ln\frac{\Lambda}{m}\left[g^3\varphi^3(x)+3g^2m^2\varphi^2(x)+3gm^4\varphi(x)+\frac{g^2}{2}\left(\partial_\mu\varphi(x)\right)^2\right]\right\}. \qquad (A9.16)$$

This leads to the results of equations (9.7,9.8).

For the interaction $U(\phi)=g\phi^4/4!$ in four dimensions we find

$$\Gamma^{\text{div}}_{1\,\text{loop}}=\frac{1}{32\pi^2}\left\{\frac{\Lambda^2}{2}g\int\mathrm{d}^4x\,\varphi^2(x)-\ln\frac{\Lambda}{m}\int\mathrm{d}^4x\left[\frac{g^2}{4}\varphi^4(x)+gm^2\varphi^2(x)\right]\right\}. \qquad (A9.17)$$

An identical expression will be recovered in Section 10.3, equation (10.24). Both in equations ($A9.16$) and ($A9.17$), we have defined the divergent part of $\Gamma(\varphi)$ as the sum of the divergent terms in the asymptotic expansion for Λ/m large.

Finally, we can also apply equation ($A9.15$) to $d=2$ and a general interaction $U(\phi)$. We find

$$\Gamma^{\text{div}}_{1\,\text{loop}}=\frac{1}{4\pi}\ln\frac{\Lambda}{m}\int\mathrm{d}^2x\,U''\left(\varphi(x)\right). \qquad (A9.18)$$

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 previous regularization method is to replace expression ($A9.6$) by

$$\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi)=-\frac{1}{2\Gamma(1+\sigma)}\int_0^\infty\mathrm{d}t\,t^{\sigma-1}\operatorname{tr}\left(\mathrm{e}^{-tM}-\mathrm{e}^{-tM_0}\right), \qquad (A9.19)$$

and to take, after analytic continuation in σ, the limit $\sigma = 0$ in the spirit of the dimensional regularization.

Let us 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:

$$\frac{1}{V}\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi) = -\frac{1}{2\Gamma(1+\sigma)}\int_0^\infty \mathrm{d}t\, t^{\sigma-1}\int\frac{\mathrm{d}^4p}{(2\pi)^4}\left[\mathrm{e}^{-t(p^2+m^2+g\varphi^2/2)} - (\varphi=0)\right].$$

The integration over the momentum p yields

$$\frac{1}{V}\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi) = -\frac{1}{32\pi^2\Gamma(1+\sigma)}\int_0^\infty \mathrm{d}t\, t^{\sigma-3}\left[\mathrm{e}^{-t(m^2+g\varphi^2/2)} - (\varphi=0)\right].$$

The integration over t can then also be performed:

$$\frac{1}{V}\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi) = -\frac{1}{32\pi^2\sigma(1-\sigma)(2-\sigma)}\left[\left(m^2+g\varphi^2/2\right)^{(2-\sigma)} - (\varphi=0)\right].$$

Expanding this expression for σ small and keeping only the divergent and finite parts, we finally obtain

$$\frac{1}{V}\Gamma^{\text{reg}}_{1\,\text{loop}}(\varphi) = -\frac{1}{64\pi^2}\left(m^2+\frac{g}{2}\varphi^2\right)^2\left[-\frac{1}{\sigma} + \ln\left(m^2+\frac{g}{2}\varphi^2\right) - \frac{3}{2}\right] - (\varphi=0).$$

The coefficient of the divergent part differs by a factor 2 from the one obtained in dimensional regularization. This is due to the choice of the normalization of σ.

INTRODUCTION TO RENORMALIZATION THEORY. RENORMALIZATION GROUP EQUATIONS

We will not enter into an extensive and general discussion of renormalization theory, but, instead, present the essential steps of the proof of the renormalizability of a simple scalar field theory: the ϕ^4 field theory in $d = 4$ dimensions. However, all the fundamental difficulties of renormalization theory are already present in this particular example and it will eventually become clear that the extension to other theories is not difficult. We have followed the elegant presentation of Callan (Les Houches 1975) which allows renormalizability and renormalization group (Callan–Symanzik) equations to be proved at once. This presentation is specially suited to our general purpose since a large part of this work is devoted to applications of renormalization group (RG). Moreover, it emphasizes already at this technical level the equivalence between renormalizability and the existence of a renormalization group. As a technical tool we shall define the initial unrenormalized theory by *dimensional regularization*, which leads to a derivation marginally simpler than momentum cut-off regularization.

One drawback of our proof of renormalizability is that it applies directly only to massive theories and the existence of a massless theory requires a specific discussion. Section 10.9 is devoted to this problem. A different form (homogeneous) of RG equations follows.

Finally, Section 10.11 contains a few remarks about the covariance of RG functions.

In the appendix, we briefly outline another method which more cleanly separates the small and large momentum region and has been employed to give another proof of renormalizability. It relies on a partial integration of large momentum modes. We also discuss some properties of super-renormalizable theories.

10.1 Power Counting. Dimensional Analysis

We consider the local action $\mathcal{S}(\phi)$, for a scalar field $\phi(x)$ in d dimensions:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\partial_\mu \phi(x)\partial_\mu \phi(x) + \frac{1}{2}m^2\phi^2(x) + \frac{1}{4!}g\phi^4(x)\right]. \tag{10.1}$$

We have used the traditional notation m^2 for the coefficient of ϕ^2, because for $g = 0$ it is the physical mass squared, though its physically relevant values are in general negative.

In Section 9.3, we have shown that the ϕ^4 vertex has dimension zero in four dimensions and this action is thus renormalizable in the sense of power counting: the superficial degree of divergence of correlation functions is independent of the order in perturbation theory. We are thus interested in the $d = 4$ limit.

Power counting in four dimensions. For a proper vertex with n external lines (1PI n-point correlation function) the degree of divergence δ is

$$\delta = 4 - n\,.$$

We also need functions with insertion of the $\phi^2(x)$ operator. The degree of divergence of the function $\Gamma^{(l,n)}$,

$$\Gamma^{(l,n)}(y_1,\ldots,y_l;x_1,\ldots,x_n) = 2^{-l}\left\langle \phi^2(y_1)\ldots\phi^2(y_l)\phi(x_1)\ldots\phi(x_n)\right\rangle_{1\mathrm{PI}},$$

is (equation (9.30))

$$\delta = 4 - n - 2l\,.$$

INTRODUCTION TO RENORMALIZATION THEORY. RENORMALIZATION GROUP EQUATIONS

We will not enter into an extensive and general discussion of renormalization theory, but, instead, present the essential steps of the proof of the renormalizability of a simple scalar field theory: the ϕ^4 field theory in $d = 4$ dimensions. However, all the fundamental difficulties of renormalization theory are already present in this particular example and it will eventually become clear that the extension to other theories is not difficult. We have followed the elegant presentation of Callan (Les Houches 1975) which allows renormalizability and renormalization group (Callan–Symanzik) equations to be proved at once. This presentation is specially suited to our general purpose since a large part of this work is devoted to applications of renormalization group (RG). Moreover, it emphasizes already at this technical level the equivalence between renormalizability and the existence of a renormalization group. As a technical tool we shall define the initial unrenormalized theory by *dimensional regularization*, which leads to a derivation marginally simpler than momentum cut-off regularization.

One drawback of our proof of renormalizability is that it applies directly only to massive theories and the existence of a massless theory requires a specific discussion. Section 10.9 is devoted to this problem. A different form (homogeneous) of RG equations follows.

Finally, Section 10.11 contains a few remarks about the covariance of RG functions.

In the appendix, we briefly outline another method which more cleanly separates the small and large momentum region and has been employed to give another proof of renormalizability. It relies on a partial integration of large momentum modes. We also discuss some properties of super-renormalizable theories.

10.1 Power Counting. Dimensional Analysis

We consider the local action $\mathcal{S}(\phi)$, for a scalar field $\phi(x)$ in d dimensions:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\partial_\mu \phi(x) \partial_\mu \phi(x) + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{4!} g \phi^4(x)\right]. \tag{10.1}$$

We have used the traditional notation m^2 for the coefficient of ϕ^2, because for $g = 0$ it is the physical mass squared, though its physically relevant values are in general negative.

In Section 9.3, we have shown that the ϕ^4 vertex has dimension zero in four dimensions and this action is thus renormalizable in the sense of power counting: the superficial degree of divergence of correlation functions is independent of the order in perturbation theory. We are thus interested in the $d = 4$ limit.

Power counting in four dimensions. For a proper vertex with n external lines (1PI n-point correlation function) the degree of divergence δ is

$$\delta = 4 - n\,.$$

We also need functions with insertion of the $\phi^2(x)$ operator. The degree of divergence of the function $\Gamma^{(l,n)}$,

$$\Gamma^{(l,n)}(y_1, \ldots, y_l; x_1, \ldots, x_n) = 2^{-l} \left\langle \phi^2(y_1) \ldots \phi^2(y_l) \phi(x_1) \ldots \phi(x_n) \right\rangle_{1\mathrm{PI}},$$

is (equation (9.30))

$$\delta = 4 - n - 2l\,.$$

INTRODUCTION TO RENORMALIZATION THEORY. RENORMALIZATION GROUP EQUATIONS

We will not enter into an extensive and general discussion of renormalization theory, but, instead, present the essential steps of the proof of the renormalizability of a simple scalar field theory: the ϕ^4 field theory in $d = 4$ dimensions. However, all the fundamental difficulties of renormalization theory are already present in this particular example and it will eventually become clear that the extension to other theories is not difficult. We have followed the elegant presentation of Callan (Les Houches 1975) which allows renormalizability and renormalization group (Callan–Symanzik) equations to be proved at once. This presentation is specially suited to our general purpose since a large part of this work is devoted to applications of renormalization group (RG). Moreover, it emphasizes already at this technical level the equivalence between renormalizability and the existence of a renormalization group. As a technical tool we shall define the initial unrenormalized theory by *dimensional regularization*, which leads to a derivation marginally simpler than momentum cut-off regularization.

One drawback of our proof of renormalizability is that it applies directly only to massive theories and the existence of a massless theory requires a specific discussion. Section 10.9 is devoted to this problem. A different form (homogeneous) of RG equations follows.

Finally, Section 10.11 contains a few remarks about the covariance of RG functions.

In the appendix, we briefly outline another method which more cleanly separates the small and large momentum region and has been employed to give another proof of renormalizability. It relies on a partial integration of large momentum modes. We also discuss some properties of super-renormalizable theories.

10.1 Power Counting. Dimensional Analysis

We consider the local action $\mathcal{S}(\phi)$, for a scalar field $\phi(x)$ in d dimensions:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\partial_\mu \phi(x)\partial_\mu \phi(x) + \frac{1}{2}m^2\phi^2(x) + \frac{1}{4!}g\phi^4(x)\right]. \tag{10.1}$$

We have used the traditional notation m^2 for the coefficient of ϕ^2, because for $g = 0$ it is the physical mass squared, though its physically relevant values are in general negative.

In Section 9.3, we have shown that the ϕ^4 vertex has dimension zero in four dimensions and this action is thus renormalizable in the sense of power counting: the superficial degree of divergence of correlation functions is independent of the order in perturbation theory. We are thus interested in the $d = 4$ limit.

Power counting in four dimensions. For a proper vertex with n external lines (1PI n-point correlation function) the degree of divergence δ is

$$\delta = 4 - n\,.$$

We also need functions with insertion of the $\phi^2(x)$ operator. The degree of divergence of the function $\Gamma^{(l,n)}$,

$$\Gamma^{(l,n)}(y_1,\ldots,y_l;x_1,\ldots,x_n) = 2^{-l}\left\langle \phi^2(y_1)\ldots\phi^2(y_l)\phi(x_1)\ldots\phi(x_n)\right\rangle_{1\mathrm{PI}},$$

is (equation (9.30))

$$\delta = 4 - n - 2l\,.$$

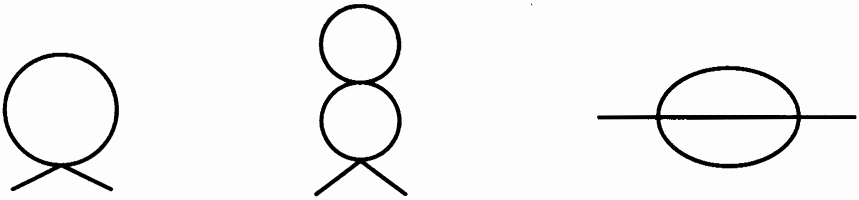


Fig. 10.1 The ϕ-field two-point function: $l = 0$, $n = 2$, $\delta = 2$.

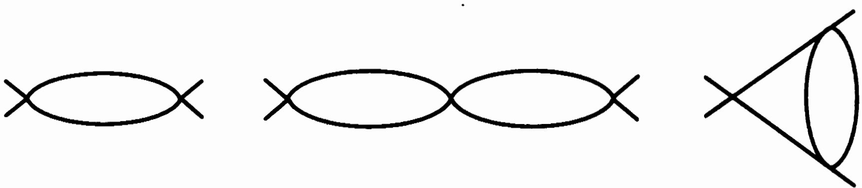


Fig. 10.2 The ϕ-field four-point function: $l = 0$, $n = 4$, $\delta = 0$.

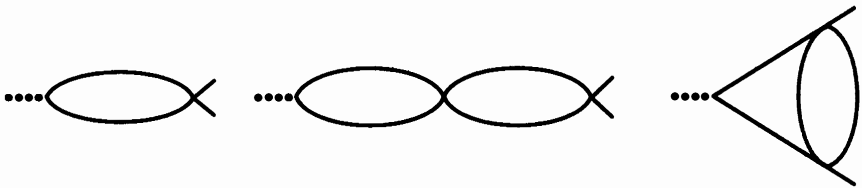


Fig. 10.3 The $\langle \phi^2 \phi\phi \rangle$ correlation function: $l = 1$, $n = 2$, $\delta = 0$.

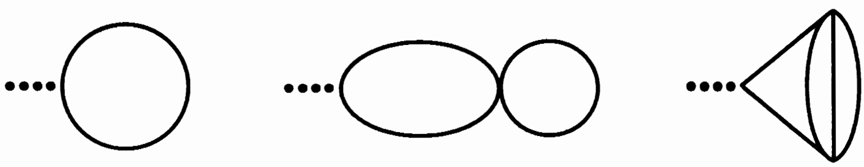


Fig. 10.4 The $\langle \phi^2 \rangle$ expectation value which is a constant $l = 1$, $n = 0$, $\delta = 2$.

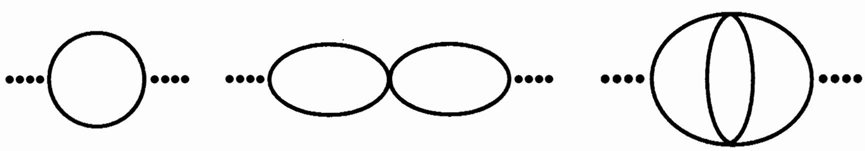


Fig. 10.5 The $\langle \phi^2 \phi^2 \rangle$ correlation function $l = 2$, $n = 0$, $\delta = 0$.

In figures 10.1–10.5, are listed the superficially divergent functions and the corresponding first few diagrams.

The diagrams with $n = 0$ never arise as subdiagrams since a subdiagram has necessarily external ϕ lines. Their renormalization can be discussed separately.

To give a meaning to perturbation theory we now replace the action $\mathcal{S}(\phi)$ by a regularized action $\mathcal{S}_\varepsilon(\phi)$, using the dimensional regularization scheme presented in Section 9.6. Some arguments would have to be slightly modified in the case of cut-off regularization.

Dimensional analysis. In what follows, mass dimensional analysis which, as we have shown in Section 9.2.1, is equivalent for a scalar field to power counting, will be useful. When dimensional regularization is used, the only dimensional parameters are the bare parameters $\{m, g\}$. In $d = 4 - \varepsilon$ dimension the field has dimension $[\phi] = (d-2)/2$, and thus the coupling constant g has a dimension $[g]$:

$$[g] = d - 4[\phi] = d - 2(d-2) = 4 - d = \varepsilon\,. \tag{10.2}$$

Connected correlation functions then satisfy

$$W^{(n)}(x_1/\lambda, \ldots, x_n/\lambda; \lambda m, \lambda^{4-d} g) = \lambda^{n(d-2)/2} W^{(n)}(x_1, \ldots, x_n; m, g). \tag{10.3}$$

After Fourier transformation and factorization of the δ-function of momentum conservation one finds

$$\widetilde{W}^{(n)}(\lambda p_1, \ldots, \lambda p_n; \lambda m, \lambda^{4-d} g) = \lambda^{d-n(d+2)/2} \widetilde{W}^{(n)}(p_1, \ldots, p_n; m, g). \tag{10.4}$$

After Legendre transformation one obtains the corresponding relations for vertex functions:

$$\begin{aligned}\Gamma^{(n)}(x_1/\lambda, \ldots, x_n/\lambda; \lambda m, g\lambda^{4-d}) &= \lambda^{n(d/2+1)} \Gamma^{(n)}(x_1, \ldots, x_n; m, g),\\ \tilde{\Gamma}^{(n)}(\lambda p_1, \ldots, \lambda p_n; \lambda m, g\lambda^{4-d}) &= \lambda^{d-n(d-2)/2} \tilde{\Gamma}^{(n)}(p_1, \ldots, p_n; m, g).\end{aligned}$$

The function $\tilde{\Gamma}^{(n)}$ has a mass dimension $d-n(d-2)/2$ which coincides with its dimension in the sense of power counting. The argument generalizes to $\tilde{\Gamma}^{(l,n)}$:

$$\tilde{\Gamma}^{(l,n)}(\lambda q_i, \lambda p_j; \lambda m, g\lambda^{4-d}) = \lambda^{d-n(d-2)/2-2l} \tilde{\Gamma}^{(n)}(q_i, p_j; m, g). \tag{10.5}$$

Loop expansion. We note that if we rescale the field ϕ,

$$\phi \mapsto \phi/\sqrt{g}\,,$$

the dependence on g is factorized in front of the action (10.1):

$$\mathcal{S}(\phi) = \frac{1}{g} \int \mathrm{d}^d x \left[\frac{1}{2}\left(\sqrt{g}\partial_\mu \phi\right)^2 + \frac{1}{2} m^2 \left(\sqrt{g}\phi\right)^2 + \frac{1}{4!}\left(\sqrt{g}\phi\right)^4\right]. \tag{10.6}$$

In the ϕ^4 field theory, the loop expansion is an expansion in powers of g at $(\sqrt{g}\phi)$ fixed.

10.2 Bare and Renormalized Field Theory. Operator ϕ^2 Insertions

Before we begin the discussion of divergences and renormalization a few definitions and remarks are in order.

10.2.1 Bare and renormalized action: counter-terms

The deviation $\varepsilon = 4 - d$ from dimension 4 is our regularization parameter and we now write the action $\mathcal{S}_\varepsilon(\phi)$ to emphasize it. The parameters m and g which appear in the action (10.1) are called the ***bare parameters***, the correlation functions of the ***bare*** field ϕ are ***bare correlation functions***. We now introduce two ***renormalized parameters*** m_r and g_r: m_r characterizes the decay of correlation functions and is proportional to the physical mass of the theory; g_r will be the new expansion parameter, $m_\mathrm{r}^\varepsilon g_\mathrm{r} = g + O(g^2)$, defined to be dimensionless in d dimensions. We want to show that it is possible to rescale the field ϕ:

$$\phi = Z^{1/2}\phi_\mathrm{r}\,,$$

thus the coupling constant g has a dimension $[g]$:

$$[g] = d - 4[\phi] = d - 2(d-2) = 4 - d = \varepsilon \,. \tag{10.2}$$

Connected correlation functions then satisfy

$$W^{(n)}(x_1/\lambda, \ldots, x_n/\lambda; \lambda m, \lambda^{4-d} g) = \lambda^{n(d-2)/2} W^{(n)}(x_1, \ldots, x_n; m, g). \tag{10.3}$$

After Fourier transformation and factorization of the δ-function of momentum conservation one finds

$$\widetilde{W}^{(n)}(\lambda p_1, \ldots, \lambda p_n; \lambda m, \lambda^{4-d} g) = \lambda^{d-n(d+2)/2} \widetilde{W}^{(n)}(p_1, \ldots, p_n; m, g). \tag{10.4}$$

After Legendre transformation one obtains the corresponding relations for vertex functions:

$$\begin{aligned} \Gamma^{(n)}(x_1/\lambda, \ldots, x_n/\lambda; \lambda m, g\lambda^{4-d}) &= \lambda^{n(d/2+1)} \Gamma^{(n)}(x_1, \ldots, x_n; m, g), \\ \tilde{\Gamma}^{(n)}(\lambda p_1, \ldots, \lambda p_n; \lambda m, g\lambda^{4-d}) &= \lambda^{d-n(d-2)/2} \tilde{\Gamma}^{(n)}(p_1, \ldots, p_n; m, g). \end{aligned}$$

The function $\tilde{\Gamma}^{(n)}$ has a mass dimension $d - n(d-2)/2$ which coincides with its dimension in the sense of power counting. The argument generalizes to $\tilde{\Gamma}^{(l,n)}$:

$$\tilde{\Gamma}^{(l,n)}(\lambda q_i, \lambda p_j; \lambda m, g\lambda^{4-d}) = \lambda^{d-n(d-2)/2-2l} \tilde{\Gamma}^{(n)}(q_i, p_j; m, g). \tag{10.5}$$

Loop expansion. We note that if we rescale the field ϕ,

$$\phi \mapsto \phi/\sqrt{g} \,,$$

the dependence on g is factorized in front of the action (10.1):

$$\mathcal{S}(\phi) = \frac{1}{g} \int \mathrm{d}^d x \left[\frac{1}{2} (\sqrt{g}\partial_\mu \phi)^2 + \frac{1}{2} m^2 (\sqrt{g}\phi)^2 + \frac{1}{4!} (\sqrt{g}\phi)^4 \right]. \tag{10.6}$$

In the ϕ^4 field theory, the loop expansion is an expansion in powers of g at $(\sqrt{g}\phi)$ fixed.

10.2 Bare and Renormalized Field Theory. Operator ϕ^2 Insertions

Before we begin the discussion of divergences and renormalization a few definitions and remarks are in order.

10.2.1 Bare and renormalized action: counter-terms

The deviation $\varepsilon = 4 - d$ from dimension 4 is our regularization parameter and we now write the action $\mathcal{S}_\varepsilon(\phi)$ to emphasize it. The parameters m and g which appear in the action (10.1) are called the ***bare parameters***, the correlation functions of the ***bare*** field ϕ are ***bare correlation functions***. We now introduce two ***renormalized parameters*** m_r and g_r: m_r characterizes the decay of correlation functions and is proportional to the physical mass of the theory; g_r will be the new expansion parameter, $m_\mathrm{r}^\varepsilon g_\mathrm{r} = g + O(g^2)$, defined to be dimensionless in d dimensions. We want to show that it is possible to rescale the field ϕ:

$$\phi = Z^{1/2} \phi_\mathrm{r} \,,$$

and to choose the bare parameters m and g as functions of m_{r}, g_{r} and ε in such a way that all ϕ_{r} correlation functions, the *renormalized correlation functions*, have a finite limit, order by order in the loop expansion, when ε goes to zero at m_{r} and g_{r} fixed. The field ϕ_{r} is called the renormalized field.

Moreover, we introduce the notion of *renormalized action* $\mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}})$ which is the initial action $\mathcal{S}_\varepsilon(\phi)$ expressed in terms of renormalized field and parameters:

$$\begin{aligned}\mathcal{S}_\varepsilon(\phi) \equiv \mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}}) = \int \mathrm{d}^d x \Big[&\frac{1}{2}\partial_\mu\phi_{\mathrm{r}}\partial_\mu\phi_{\mathrm{r}} + \frac{1}{2}m_{\mathrm{r}}^2\phi_{\mathrm{r}}^2 + \frac{1}{4!}m_{\mathrm{r}}^\varepsilon g_{\mathrm{r}}\phi_{\mathrm{r}}^4 \\ &+ \frac{1}{2}(Z-1)\partial_\mu\phi_{\mathrm{r}}\partial_\mu\phi_{\mathrm{r}} + \frac{1}{2}\delta m^2\phi_{\mathrm{r}}^2 + \frac{1}{4!}m_{\mathrm{r}}^\varepsilon g_{\mathrm{r}}(Z_g-1)\phi_{\mathrm{r}}^4\Big] . \end{aligned} \tag{10.7}$$

In the action (10.7) we have explicitly separated a *tree order action*,

$$\mathcal{S}_{\mathrm{r},0}(\phi_{\mathrm{r}}) = \int \mathrm{d}^d x \left(\frac{1}{2}\partial_\mu\phi_{\mathrm{r}}\partial_\mu\phi_{\mathrm{r}} + \frac{1}{2}m_{\mathrm{r}}^2\phi_{\mathrm{r}}^2 + \frac{1}{4!}m_{\mathrm{r}}^\varepsilon g_{\mathrm{r}}\phi_{\mathrm{r}}^4\right), \tag{10.8}$$

from the set of *counter-terms*.

The identity between the renormalized action (10.7) and the regularized action (10.1), called the *bare action*, is expressed by the set of relations between renormalized and bare quantities:

$$\begin{aligned}\phi &= Z^{1/2}\phi_{\mathrm{r}}\,, \\ g &= m_{\mathrm{r}}^\varepsilon g_{\mathrm{r}} Z_g/Z^2, \\ m^2 &= (m_{\mathrm{r}}^2 + \delta m^2)/Z\,. \end{aligned} \tag{10.9}$$

Z and Z_g/Z^2 are, respectively, called the field amplitude and coupling constant renormalization constants, δm^2 characterizes the mass renormalization.

The renormalization constants δm^2, Z_g, and Z are formal series in g_{r}:

$$\begin{aligned}\delta m^2 &= m_{\mathrm{r}}^2\left[a_1(\varepsilon)g_{\mathrm{r}} + a_2(\varepsilon)g_{\mathrm{r}}^2 + \cdots\right] \\ Z_g &= 1 + b_1(\varepsilon)g_{\mathrm{r}} + b_2(\varepsilon)g_{\mathrm{r}}^2 + \cdots \\ Z &= 1 + c_1(\varepsilon)g_{\mathrm{r}} + c_2(\varepsilon)g_{\mathrm{r}}^2 + \cdots\,. \end{aligned} \tag{10.10}$$

For dimensional reasons the coefficients a_i, b_i, c_i are independent of m_{r}. Note a property which is specific to dimensional regularization: δm^2 is proportional to m_{r}^2.

We want to prove that the coefficients $a_n(\varepsilon)$, $b_n(\varepsilon)$ and $c_n(\varepsilon)$ can be chosen in such a way that all renormalized correlation functions have a finite $\varepsilon \to 0$ limit, order by order in g_{r}.

Renormalized loop expansion. Because the renormalization constants are series in g_{r}, the expansion in powers of g_{r} at $(\sqrt{g_{\mathrm{r}}}\phi_{\mathrm{r}})$ fixed is no longer a loop expansion in the diagrammatic sense, in contrast with the expansion in powers of g (equation (10.6)). At order g_{r}^{L-1}, contribute L loop diagrams and diagrams with less than L loops multiplied by renormalization contributions. Below, when no confusion is possible we will, nevertheless, call loop expansion the expansion in powers of g_{r} at $(\sqrt{g_{\mathrm{r}}}\phi_{\mathrm{r}})$ fixed.

10.2.2 Bare and renormalized correlation functions

The relation between bare and renormalized fields $\phi = Z^{1/2}\phi_{\mathrm{r}}$ implies directly the relations between connected renormalized and bare correlation functions $W_{\mathrm{r}}^{(n)} = Z^{-n/2}W^{(n)}$, which can be summarized by

$$\mathcal{W}(J/\sqrt{Z}) = \mathcal{W}_{\mathrm{r}}(J), \tag{10.11}$$

in which $\mathcal{W}(J)$ and $\mathcal{W}_{\mathrm{r}}(J)$ are, respectively, the generating functionals of connected bare and renormalized correlation functions. One verifies, after Legendre transformation (J and φ are dual), that the corresponding 1PI functionals are related by

$$\Gamma_{\mathrm{r}}(\varphi) = \Gamma(\varphi\sqrt{Z}), \tag{10.12}$$

a relation which for the renormalized and bare proper vertices translates into

$$\Gamma_{\mathrm{r}}^{(n)} = Z^{n/2}\Gamma^{(n)}. \tag{10.13}$$

Operator ϕ^2 insertions. Actually, we also need the bare and renormalized ϕ^2 insertions. We, therefore, introduce a source $K(x)$ for $\phi^2(x)$ and add to the bare action (10.1) a source term:

$$\mathcal{S}(\phi, K) = \mathcal{S}(\phi) + \frac{1}{2}\int \mathrm{d}x\, K(x)\phi^2(x). \tag{10.14}$$

We now consider the functional integral:

$$\mathcal{Z}(J,K) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi,K) + \int \mathrm{d}x\, J(x)\phi(x)\right]. \tag{10.15}$$

Functional differentiation with respect to $K(x)$ generates insertions of the operator $-\frac{1}{2}\phi^2(x)$ (Section 9.4). In the same way, if we consider the action $\mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}}, K)$,

$$\mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}}, K) = \mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}}) + \frac{1}{2}Z_2 \int K(x)\phi_{\mathrm{r}}^2(x)\mathrm{d}x\,, \tag{10.16}$$

in which Z_2 is a new renormalization constant, the functional integral

$$\mathcal{Z}_{\mathrm{r}}(J,K) = \int [\mathrm{d}\phi_{\mathrm{r}}] \exp\left[-\mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}},K) + \int \mathrm{d}x\, J(x)\phi_{\mathrm{r}}(x)\right], \tag{10.17}$$

generates the renormalized correlation functions with $-\frac{1}{2}\phi^2$ insertions (we temporarily normalize $\mathcal{Z}(J,K)$ to $\mathcal{Z}(0,K) = 1$ to eliminate the pure $\phi^2(x)$ correlation functions).

The relation between renormalized and bare functionals is then:

$$\mathcal{W}_{\mathrm{r}}(J,K) = \mathcal{W}(J/\sqrt{Z}, KZ_2/Z). \tag{10.18}$$

After Legendre transformation this relation implies

$$\Gamma_{\mathrm{r}}(\varphi, K) = \Gamma(\varphi\sqrt{Z}, KZ_2/Z), \tag{10.19}$$

or in terms of proper vertices:

$$\Gamma_{\mathrm{r}}^{(l,n)}(y_1,\ldots,y_l;x_1,\ldots,x_n) = Z^{(n/2)-l}Z_2^l\Gamma^{(l,n)}(y_1,\ldots,y_l;x_1,\ldots,x_n). \tag{10.20}$$

We note here that if the source K is a constant, then, according to equation (10.14), it generates just a shift of m^2, the bare mass squared:

$$\mathcal{S}(\phi, K) = \int \mathrm{d}^d x \left[\frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}(m^2 + K)\phi^2 + \frac{1}{4!}g\phi^4\right]. \tag{10.21}$$

In this limit, differentiation with respect to K is equivalent to differentiation with respect to m^2. On the other hand, if $K(x)$ is a constant, its Fourier transform is proportional to $\delta(p)$, which means that it generates correlation functions with insertions of the Fourier transform of $\phi^2(x)$ at zero momentum:

$$\left.\frac{\partial}{\partial m^2}\right|_g \tilde{\Gamma}^{(l,n)}(q_1, \ldots, q_l; p_1, \ldots, p_n) = \tilde{\Gamma}^{(l+1,n)}(0, q_1, \ldots, q_l; p_1, \ldots, p_n). \tag{10.22}$$

The equation has a diagrammatic interpretation: the diagrams contributing to the r.h.s. are obtained from the diagrams contributing to $\Gamma^{(l,n)}$ by doubling a propagator in all possible ways (up to a sign). In figure 10.6 we give the example of $\Gamma^{(4)}$ and $\Gamma^{(1,4)}$.

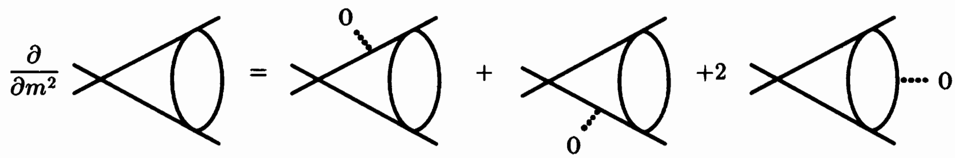


Fig. 10.6

10.3 One-Loop Divergences

It has already been shown in Section 9.1 that the ϕ^3_6 (the lower index is the space dimension) field theory can be renormalized at one-loop order. We repeat here the argument in the example of the ϕ^4_4 field theory.

We expand the generating functional of renormalized proper vertices $\Gamma_\mathrm{r}(\varphi)$ at one-loop order (really in powers of g_r at $\varphi\sqrt{g_\mathrm{r}}$ fixed).

At tree order the counter-terms, by definition, do not contribute and thus (equation (10.8))

$$\Gamma_\mathrm{r}(\varphi) = \Gamma_{\mathrm{r},0}(\varphi) = \lim_{\varepsilon \to 0} \mathcal{S}_{\mathrm{r},0}(\varphi) = \int \mathrm{d}^4 x \left[\frac{1}{2}(\partial_\mu \varphi)^2 + \frac{1}{2}m_\mathrm{r}^2 \varphi^2 + \frac{1}{4!}g_\mathrm{r}\varphi^4\right]. \tag{10.23}$$

At one-loop order we find the one-loop contributions generated by the tree order action and the counter-terms at leading order. The former are given by equation (7.101):

$$\begin{aligned}\Gamma_1(\varphi) &= \tfrac{1}{2}\operatorname{tr}\ln\left[1 + \left(m_\mathrm{r}^2 - \nabla^2\right)^{-1} m_\mathrm{r}^\varepsilon g_\mathrm{r}\varphi^2/2\right] \\ &= \tfrac{1}{4}\operatorname{tr}\left(m_\mathrm{r}^2 - \nabla^2\right)^{-1} m_\mathrm{r}^\varepsilon g_\mathrm{r}\varphi^2 - \tfrac{1}{16}\operatorname{tr}\left[\left(m_\mathrm{r}^2 - \nabla^2\right)^{-1} m_\mathrm{r}^\varepsilon g_\mathrm{r}\varphi^2\right]^2 + O\left(\varphi^6\right).\end{aligned}$$

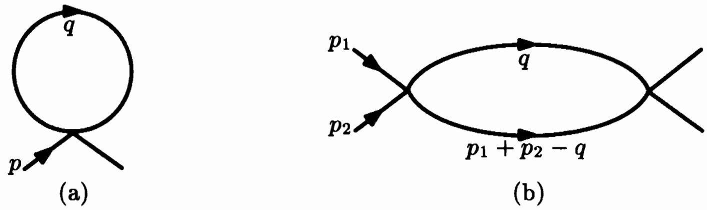


Fig. 10.7 One-loop divergent diagrams.

The first two terms in the expansion in powers of φ^2 correspond to the two divergent diagrams displayed in figure 10.7.

The divergent contributions, in dimensional regularization (see Section 11.5 for details), are

(i) $n = 2$: the coefficient of $g_r\varphi^2$ is

$$\frac{1}{4}\,(\mathrm{a}) = \frac{1}{4} m_r^{\varepsilon} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{(q^2 + m_r^2)} \sim -\frac{1}{32\pi^2}\left(\frac{m_r^2}{\varepsilon}\right).$$

(ii) $n = 4$: the coefficient of $g_r^2\varphi^4$ is

$$-\frac{1}{16}(\mathrm{b}) = -\frac{1}{16} m_r^{2\varepsilon} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{(q^2 + m_r^2)\left[(q - p_1 - p_2)^2 + m_r^2\right]} \sim -\frac{1}{128\pi^2\varepsilon} m_r^{\varepsilon}.$$

The one-loop divergent part $\Gamma_1^{\mathrm{div}}(\varphi)$ generated by the tree action (10.8) is thus:

$$\Gamma_1^{\mathrm{div}} = -\frac{1}{32\pi^2\varepsilon} \int \mathrm{d}^d x \left[m_r^2 g_r \left(\varphi(x)\right)^2 + \tfrac{1}{4} m_r^{\varepsilon} g_r^2 \left(\varphi(x)\right)^4 \right]. \tag{10.24}$$

Note the absence of a term proportional to $\int \mathrm{d}^d x (\partial_\mu \varphi)^2$. This is a peculiarity of the ϕ^4 field theory at one-loop order.

We now consider the modified action

$$\mathcal{S}_{r,1}(\phi_r) = \mathcal{S}_{r,0}(\phi_r) - \Gamma_1^{\mathrm{div}}(\phi_r).$$

At this order $\Gamma_1^{\mathrm{div}}(\phi_r)$ contributes additively to $\Gamma_r(\varphi)$. Its addition thus eliminates the divergences of the $\varepsilon \to 0$ limit and the ϕ^4 field theory can be renormalized at one-loop order

$$\Gamma_r(\varphi) = \Gamma_{r,0}(\varphi) + \Gamma_{r,1}(\varphi), \quad \Gamma_{r,1}(\varphi) \underset{\varepsilon\to 0}{=} \Gamma_1(\varphi) - \Gamma_1^{\mathrm{div}}(\varphi).$$

Identifying $\mathcal{S}_{r,1}(\phi_r)$ with the action (10.7) we infer the divergent part of the counter-terms expanded at one-loop order, that is, in the parametrization (10.10) of a_1, b_1 and c_1.

The condition of finiteness of correlation functions determines these coefficients only up to arbitrary finite constants. The difference between the 1PI correlation 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 1PI correlation functions

$\Gamma_{\rm r}^{(n)}$ which correspond to bare correlation functions with superficial divergences. In Fourier space they read

$$\begin{aligned}\tilde{\Gamma}_{\rm r}^{(2)}(p=0) &= m_{\rm r}^2\,,\\ \frac{\partial}{\partial p^2}\tilde{\Gamma}_{\rm r}^{(2)}\,(p)\big|_{p=0} &= 1\,,\\ \tilde{\Gamma}_{\rm r}^{(4)}(0,0,0,0) &= m_{\rm r}^{\varepsilon} g_{\rm r}\,.\end{aligned} \tag{10.25}$$

These conditions are consistent with the tree approximation. They completely determine the three renormalization constants. At one-loop order

$$\begin{aligned}a_1(\varepsilon) &= -\frac{1}{2}\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{m_{\rm r}^{2-d}}{(q^2+m_{\rm r}^2)} = -\frac{1}{2}\frac{1}{(4\pi)^{d/2}\Gamma(d/2)}\frac{\pi}{\sin(\pi d/2)} \sim \frac{1}{16\pi^2\varepsilon}\,,\\ b_1(\varepsilon) &= \frac{3}{2}\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{m_{\rm r}^{\varepsilon}}{(q^2+m_{\rm r}^2)^2} = \frac{3}{2-d}\frac{1}{(4\pi)^{d/2}\Gamma(d/2)}\frac{\pi}{\sin(\pi d/2)} \sim \frac{3}{16\pi^2\varepsilon}\,,\\ c_1(\varepsilon) &= 0\,.\end{aligned} \tag{10.26}$$

Operator ϕ^2 insertions. One obtains the one-loop contribution to the superficially divergent $\langle\phi^2\phi\phi\rangle$ 1PI correlation function by adding $K(x)$ to $\frac{1}{2}m_{\rm r}^{\varepsilon}g_r\varphi_{\rm r}^2$ in the tr ln, and looking for the $K\varphi^2$ term:

$$\tilde{\Gamma}^{(1,2)}(q;p_1,p_2) = 1 - \frac{1}{2}\frac{m_{\rm r}^{\varepsilon}g_{\rm r}}{(2\pi)^d}\int\frac{\mathrm{d}^d k}{(k^2+m_{\rm r}^2)[(k+q)^2+m_{\rm r}^2]} + O\left(g_{\rm r}^2\right). \tag{10.27}$$

The additional renormalization condition,

$$\tilde{\Gamma}_{\rm r}^{(1,2)}(q=0;p_1=p_2=0) = 1\,, \tag{10.28}$$

which again is consistent with the tree approximation, determines the new renormalization constant Z_2 (equation (10.20)). At one-loop order,

$$Z_2 - 1 = \frac{1}{2}\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{m_{\rm r}^{\varepsilon}}{(q^2+m_{\rm r}^2)^2}g_{\rm r} = \frac{1}{2-d}\frac{1}{(4\pi)^{d/2}\Gamma(d/2)}\frac{\pi}{\sin(\pi d/2)}g_{\rm r} \sim \frac{g_{\rm r}}{16\pi^2\varepsilon}\,. \tag{10.29}$$

10.4 Divergences Beyond One-Loop: Skeleton Diagrams

To all orders in the loop expansion, the two-point function has quadratic superficial divergences and the four-point function logarithmic divergences as power counting shows. However, at higher orders a new difficulty, which we have already mentioned in Section 9.2.1, arises: superficially convergent diagrams may have divergent subdiagrams. Let us take the example of the six-point function: at one-loop order it is given by a convergent diagram, but at two-loop order the diagrams displayed in figure 10.8 appear.

(c) (d)

Fig. 10.8

One recognizes inside the dashed boxes 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 (figure 10.9).

(c$'$) (d$'$)

Fig. 10.9 Two-loop contributions from one-loop counter-terms.

We want to show that this is generally true, that is, that counter-terms which render the divergent functions finite, at higher orders also cancel the divergence of subdiagrams of superficially convergent functions.

For this purpose, we introduce the notion of ***skeleton diagram***: a skeleton diagram is a really convergent diagram, that is, it is superficially convergent and has no divergent subdiagram.

For example the one-loop diagrams of the $2n$-point functions, $n > 2$, are all skeleton diagrams (see figure 10.10).

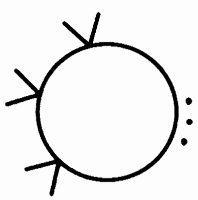


Fig. 10.10

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 figure 10.11.

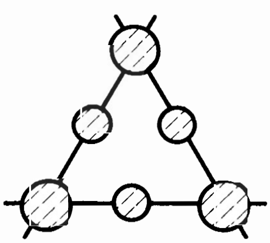


Fig. 10.11

An important property is the following: if in a dressed skeleton diagram, $\Gamma^{(4)}$ and $\Gamma^{(2)}$ are replaced by the renormalized functions $\Gamma_{\mathrm{r}}^{(4)}$ and $\Gamma_{\mathrm{r}}^{(2)}$ the dressed skeleton diagram is finite. This is a direct consequence of the following bounds on the large momentum behaviour:

$$\left.\begin{aligned} &\left|\tilde{\Gamma}_{\mathrm{r}}^{(2)}(\lambda p)\right| \le \lambda^2 \times \text{ power of } \ln \lambda, \\ &\left|\tilde{\Gamma}_{\mathrm{r}}^{(4)}(\lambda p_i)\right| \le \text{ power of } \ln \lambda, \\ &\left|\tilde{\Gamma}_{\mathrm{r}}^{(1,2)}(\lambda q; \lambda p_1, \lambda p_2)\right| \le \text{ power of } \ln \lambda, \end{aligned}\right\} \text{at any finite order for } \lambda \to \infty. \tag{10.30}$$

These bounds will not be derived here but a few comments can be found at the end of Section 10.6. 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.

Note that similar estimates exist for superficially convergent functions but are then valid only for generic momenta (see Sections 12.3–12.4).

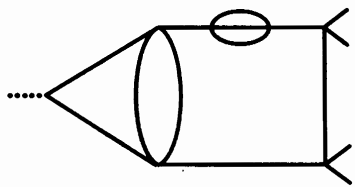


Fig. 10.12 Divergent contribution to $\Gamma^{(1,4)}$.

The bounds (10.30) together with the skeleton expansion completely reduce the problem of renormalization of superficially convergent proper vertices to the renormalization of the divergent proper vertices. The argument also applies to the proper vertices $\Gamma^{(l,n)}$ with ϕ^2 insertion. Let us, for example, consider the diagram of figure 10.12, which contributes to the superficially convergent function $\Gamma^{(1,4)}$: it 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 figure 10.13.

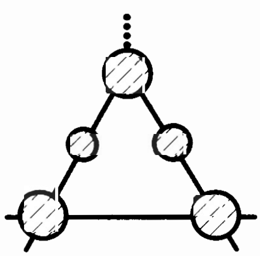


Fig. 10.13

In the next section we, therefore, 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. Figure 10.14 provides an example.

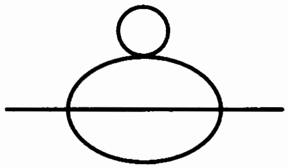


Fig. 10.14 Three-loop contribution to the two-point function.

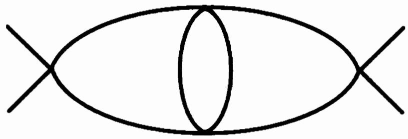


Fig. 10.15 Three-loop contribution to the four-point function.

However, another problem arises, the problem of overlapping divergences. Let us consider, for example, the diagram of figure 10.15.

Figure 10.16 shows the set of divergent subdiagrams. The three subdiagrams have a common part. The concept of insertion of divergent diagrams of lower order is, therefore, no longer well-defined. This is the problem of the so-called overlapping divergences.

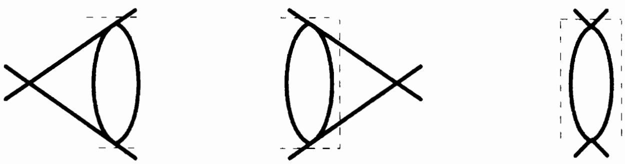


Fig. 10.16 Overlapping divergent subdiagrams.

In the next section, we develop a specific technique to deal with this problem, based on differentiating diagrams with respect to the mass.

10.5 Callan–Symanzik Equations

The starting point of the analysis is equation (10.22) 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 correlation 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 proper vertices. This introduces some additional difficulties which we will discover once we have transformed equation (10.22) into an equation for the renormalized proper vertices.

We first introduce a notation and apply the chain rule to transform differentiation with respect to m_r at g fixed, into differentiation at g_r fixed:

$$D_\mathrm{r} \equiv m_\mathrm{r} \left.\frac{\partial}{\partial m_\mathrm{r}}\right|_g = m_\mathrm{r} \left.\frac{\partial}{\partial m_\mathrm{r}}\right|_{g_\mathrm{r}} + (D_\mathrm{r}\, g_\mathrm{r}) \left.\frac{\partial}{\partial g_\mathrm{r}}\right|_{m_\mathrm{r}} . \tag{10.31}$$

We then define the quantities $\beta, \eta, \eta_2, \sigma$, taking immediately into account dimensional analysis:

$$D_\mathrm{r}\, g_\mathrm{r} = \beta(g_\mathrm{r}, \varepsilon), \tag{10.32}$$

$$D_{\mathrm{r}} \ln Z = \eta(g_{\mathrm{r}}, \varepsilon), \tag{10.33}$$
$$D_{\mathrm{r}} \ln(Z_2/Z) = \eta_2(g_{\mathrm{r}}, \varepsilon), \tag{10.34}$$
$$Z Z_2^{-1} D_{\mathrm{r}} m^2 = m_{\mathrm{r}}^2 \sigma(g_{\mathrm{r}}, \varepsilon). \tag{10.35}$$

With this notation and definitions, equation (10.22) implies

$$D_{\mathrm{r}} \tilde{\Gamma}^{(l,n)} = (D_{\mathrm{r}} m^2) \frac{\partial}{\partial m^2} \tilde{\Gamma}^{(l,n)} = m_{\mathrm{r}}^2 Z_2 Z^{-1} \sigma\, \tilde{\Gamma}^{(l+1,n)}(0, \ldots). \tag{10.36}$$

Using relation (10.20) to replace bare by renormalized functions, we then translate the equation into an equation for renormalized functions. After some simple algebra, in terms of the differential operator $\mathrm{D_{CS}}$,

$$\mathrm{D_{CS}} = m_{\mathrm{r}} \frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}}, \varepsilon) \frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2} \eta(g_{\mathrm{r}}, \varepsilon) - l \eta_2(g_{\mathrm{r}}, \varepsilon),$$

we find

$$\mathrm{D_{CS}} \tilde{\Gamma}_{\mathrm{r}}^{(l,n)}(q_1, \ldots, q_l; p_1, \ldots, p_n) = m_{\mathrm{r}}^2 \sigma(g_{\mathrm{r}}, \varepsilon) \tilde{\Gamma}_{\mathrm{r}}^{(l+1,n)}(0, q_1, \ldots, q_l; p_1, \ldots, p_n). \tag{10.37}$$

Equation (10.37), in the $\varepsilon \to 0$ limit, then yields an equation for the renormalized proper vertices, first derived by Callan and Symanzik, called, therefore, the Callan–Symanzik (CS) equation, which, in various forms, plays a central role in the part of this work devoted to phase transitions (Chapters 25–37).

To prove renormalizability, we shall prove inductively on the number of loops both the existence of the CS equation and the finiteness of correlation functions.

Renormalization conditions. The CS equation in the form (10.37) expresses only that we have rescaled the correlation functions and made an arbitrary change of parametrization. To be able to prove that the renormalized correlation functions have a finite $\varepsilon \to 0$ limit, it is necessary to determine the renormalization constants and, therefore, to impose on equation (10.37) the consequences of renormalization conditions (10.25) and (10.28):

(i) $n = 2, l = 0$

At zero momentum we obtain

$$\left(m_{\mathrm{r}} \frac{\partial}{\partial m_{\mathrm{r}}} - \eta(g_{\mathrm{r}}, \varepsilon) \right) m_{\mathrm{r}}^2 = m_{\mathrm{r}}^2 \sigma(g_{\mathrm{r}}, \varepsilon).$$

The function σ is thus related to η:

$$\sigma = 2 - \eta\,. \tag{10.38}$$

If we then differentiate with respect to momentum squared we find

$$-\eta = m_{\mathrm{r}}^2 (2 - \eta) \frac{\partial}{\partial p^2} \tilde{\Gamma}_{\mathrm{r}}^{(1,2)}(0; p, -p) \Big|_{p^2=0}. \tag{10.39}$$

(ii) $n = 4, l = 0$

At zero momentum we obtain

$$\varepsilon g_{\mathrm{r}} + \beta - 2 g_{\mathrm{r}} \eta = m_{\mathrm{r}}^{2-\varepsilon} (2 - \eta) \tilde{\Gamma}_{\mathrm{r}}^{(1,4)}(0; 0, 0, 0, 0). \tag{10.40}$$

(iii) $n = 2, l = 1$

Again at zero momentum we get

$$-\eta - \eta_2 = m_{\mathrm{r}}^2 (2 - \eta) \tilde{\Gamma}_{\mathrm{r}}^{(2,2)}(0, 0; 0, 0). \tag{10.41}$$

We have related all the coefficients of the partial differential equation (10.37) to values of proper vertices at zero momentum. From these relations it follows that if we can show that the renormalized proper vertices have a limit when ε vanishes, the functions β, η and η_2 will also have a limit.

Note also that if we know the coefficients of the CS equations, we can calculate the renormalization constants from the set of equations (10.32–10.35).

Leading order contributions.

(i) $\tilde{\Gamma}^{(1,2)}(0; p, -p)$ at order g does not depend on p (equation (10.27)). The equation (10.29) then implies

$$\tilde{\Gamma}_{\mathrm{r}}^{(1,2)}(0; p, -p) = 1 + O\left(g_{\mathrm{r}}^2\right).$$

We conclude that the expansion of $\eta(g_{\mathrm{r}})$, which can be calculated from equation (10.39), begins at order g_{r}^2.

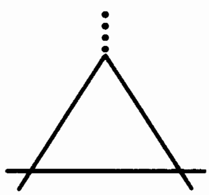


Fig. 10.17

(ii) The first diagram contributing to $\tilde{\Gamma}^{(1,4)}$ is of order g_{r}^2 (see figure 10.17). It then follows from equation (10.40) and the previous remark that the function

$$\beta_2(g_{\mathrm{r}}) = \beta(g_{\mathrm{r}}) + \varepsilon g_{\mathrm{r}}\,, \tag{10.42}$$

has an expansion which begins at order g_{r}^2. Thus the operator $\beta_2 \partial/\partial g_{\mathrm{r}}$ which contributes to the CS equation is of order g_{r}.

(iii) The function $\tilde{\Gamma}^{(2,2)}$ has a first contribution of order g_{r} (see figure 10.18). Equation (10.41) then shows that η_2 also begins at order g_{r}.

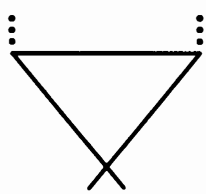


Fig. 10.18

Cluster properties and analyticity at low momentum. In Section 9.1 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 1PI correlation functions are analytic functions around $\mathbf{p} = 0$ 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 Appendix A7.3).

10.6 Inductive Proof of Renormalizability

We have shown in Section 10.3 that a finite theory could be constructed at one-loop order. We now assume that the correlation functions defined by equations (10.20) and renormalization conditions (10.25,10.28) have a finite limit up to loop order L, at m_r and g_r fixed, when ε goes to zero. This means that $\tilde{\Gamma}_\mathrm{r}^{(2)}$, $\tilde{\Gamma}_\mathrm{r}^{(4)}$, $\tilde{\Gamma}_\mathrm{r}^{(1,2)}$ have a limit when $\varepsilon \to 0$ up to order g_r^L, g_r^{L+1} and g_r^L, respectively. As we have shown in Section 10.5, from equations (10.20,10.25,10.28) follow the CS equations (10.37) and the relations (10.38–10.41).

We now use the CS equation (10.37) in the form

$$\left(m_\mathrm{r}\frac{\partial}{\partial m_\mathrm{r}} - \varepsilon g_\mathrm{r}\frac{\partial}{\partial g_\mathrm{r}}\right)\tilde{\Gamma}_\mathrm{r}^{(l,n)} = \left(-\beta_2\frac{\partial}{\partial g_\mathrm{r}} + \frac{n}{2}\eta + l\eta_2\right)\tilde{\Gamma}_\mathrm{r}^{(l,n)} + m_\mathrm{r}^2(2-\eta)\tilde{\Gamma}_\mathrm{r}^{(l+1,n)}, \quad (10.43)$$

and show that the r.h.s. is finite at loop order $L+1$. We note that $\Gamma^{(l,n)}$ in the r.h.s. 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 correlation functions let us examine the coefficient functions.

10.6.1 Coefficients of the CS equation

(i) Because $\partial\tilde{\Gamma}^{(1,2)}/\partial p^2$ is of order g_r^2, equation (10.39) then implies that η is finite up to order g_r^L.

(ii) The function $\tilde{\Gamma}^{(1,4)}$ is superficially convergent. It has, therefore, a skeleton expansion. The first dressed skeleton diagram contributing to $\tilde{\Gamma}^{(1,4)}$ is shown in figure 10.19.

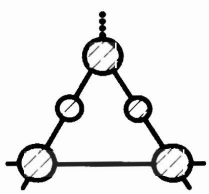


Fig. 10.19

If the functions $\tilde{\Gamma}_\mathrm{r}^{(2)}$, $\tilde{\Gamma}_\mathrm{r}^{(4)}$ and $\tilde{\Gamma}_\mathrm{r}^{(1,2)}$ are finite up to L loops, $\tilde{\Gamma}_\mathrm{r}^{(1,4)}$ is finite up to loop order $L+1$, which means up to order g_r^{L+2}. Equation (10.40) then shows that the combination $\beta - 2g_\mathrm{r}\eta$ 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 figure 10.20.

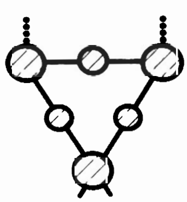


Fig. 10.20

The function $\tilde{\Gamma}_{\mathrm{r}}^{(2,2)}$ is, therefore, 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 (10.41) involves only η at order g_{r}^{L} and $\tilde{\Gamma}_{\mathrm{r}}^{(2,2)}$ 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 $\tilde{\Gamma}_{\mathrm{r}}^{(2)}$, $\tilde{\Gamma}_{\mathrm{r}}^{(4)}$ and $\tilde{\Gamma}_{\mathrm{r}}^{(1,2)}$ are, with the induction assumptions, finite up to loop order $L+1$.

10.6.2 The $\langle\phi\phi\phi\phi\rangle$ correlation function ($l=0$, $n=4$)

We consider the coefficient of order g_{r}^{L+2} in equation (10.43):

$$\left[\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}}-\varepsilon g_{\mathrm{r}}\frac{\partial}{\partial g_{\mathrm{r}}}\right)\tilde{\Gamma}_{\mathrm{r}}^{(4)}\right]_{L+2}=\left[\left(-\beta_2\frac{\partial}{\partial g_{\mathrm{r}}}+2\eta\right)\tilde{\Gamma}_{\mathrm{r}}^{(4)}\right]_{L+2}+m_{\mathrm{r}}^2\left[(2-\eta)\tilde{\Gamma}_{\mathrm{r}}^{(1,4)}\right]_{L+2}.$$

Since $\beta_2\partial/\partial g_{\mathrm{r}}$ is of order g_{r} and η of order g_{r}^2 for g_{r} small, in the first term of the r.h.s. we need $\tilde{\Gamma}_{\mathrm{r}}^{(4)}$ only up to order g_{r}^{L+1}, which is finite by assumption. We now separate in $\tilde{\Gamma}_{\mathrm{r}}^{(4)}$ the leading term $m_{\mathrm{r}}^{\varepsilon}g_{\mathrm{r}}$ 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 $\tilde{\Gamma}_{\mathrm{r}}^{(4)}$ then involves the combination

$$\left[\left(-\beta_2\frac{\partial}{\partial g_{\mathrm{r}}}+2\eta\right)g_{\mathrm{r}}\right]_{L+2}=\left[-\beta_2+2\eta g_{\mathrm{r}}\right]_{L+2},$$

and we have shown above that $\beta-2g_{\mathrm{r}}\eta$ is finite up to order g_{r}^{L+2}.

Finally, $\tilde{\Gamma}_{\mathrm{r}}^{(1,4)}$ 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 l.h.s. is finite at loop order $L+1$.

Perturbative integration of CS equations. We now integrate equation (10.43). The function $\tilde{\Gamma}_{\mathrm{r}}^{(4)}$ has dimension ε. We thus set

$$\tilde{\Gamma}_{\mathrm{r}}^{(4)}=m_{\mathrm{r}}^{\varepsilon}g_{\mathrm{r}}\gamma^{(4)}\ \Rightarrow\ \gamma^{(4)}(0,0,0,0)=1\,.$$

The function $\gamma^{(4)}$ is now invariant in a dilatation of parameter ρ:

$$(p_i,m_{\mathrm{r}})\mapsto(\rho p_i,\rho m_{\mathrm{r}})\,.$$

Therefore, we set

$$\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}}-\varepsilon g_{\mathrm{r}}\frac{\partial}{\partial g_{\mathrm{r}}}\right)\gamma^{(4)}(p_i/m_{\mathrm{r}},g_{\mathrm{r}},\varepsilon)=f^{(4)}(p_i/m_{\mathrm{r}},g_{\mathrm{r}},\varepsilon)\tag{10.44}$$

with

$$\lim_{\varepsilon\to0}f^{(4)}(p_i/m_{\mathrm{r}},g_{\mathrm{r}},\varepsilon)<\infty\,,$$
$$f^{(4)}(p_i=0,g_{\mathrm{r}},\varepsilon)=\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}}-\varepsilon g_{\mathrm{r}}\frac{\partial}{\partial g_{\mathrm{r}}}\right)1=0\,.$$

The regularity at low momentum in a massive theory then implies

$$f^{(4)}(p_i/m_{\mathrm{r}},g_{\mathrm{r}},\varepsilon)\underset{m_{\mathrm{r}}\to\infty}{=}O\left(m_{\mathrm{r}}^{-2}\right).$$

We now integrate equation (10.44) between m_r and infinity at $g_r m_r^\varepsilon$ fixed. When $m_r \to \infty$ at $g_r m_r^\varepsilon$ fixed, g_r goes to zero, which can only improve the situation. In this limit p_i/m_r goes to zero. We can then use for $m_r = \infty$ the renormalization condition (10.25) at $p_i = 0$ as boundary condition. Thus the integral converges for ρ large:

$$\gamma^{(4)}(p_i/m_r, g_r, \varepsilon) = 1 - \int_{m_r}^{\infty} \frac{d\rho}{\rho} f^{(4)}\left(\frac{p_i}{\rho}, \frac{m_r^\varepsilon g_r}{\rho^\varepsilon}, \varepsilon\right). \tag{10.45}$$

We conclude that $\tilde{\Gamma}_r^{(4)}$ has a limit for $\varepsilon \to 0$ at loop order $L+1$, given by equation (10.45), both sides being expanded up to g_r^{L+2}. We now repeat the argument for $\tilde{\Gamma}_r^{(1,2)}$.

10.6.3 The $\langle \phi^2 \phi\phi \rangle$ correlation function ($l = 1$, $n = 2$)

We consider the term of order g_r^{L+1} in equation (10.43):

$$\left[\left(m_r \frac{\partial}{\partial m_r} - \varepsilon g_r \frac{\partial}{\partial g_r}\right) \tilde{\Gamma}_r^{(1,2)}\right]_{L+1} = -\left[\beta_2 \frac{\partial}{\partial g_r} \tilde{\Gamma}_r^{(1,2)}\right]_{L+1} + \left[(\eta + \eta_2)\, \tilde{\Gamma}_r^{(1,2)}\right]_{L+1}$$
$$+ m_r^2 \left[(2-\eta)\, \tilde{\Gamma}_r^{(2,2)}\right]_{L+1}.$$

The first term in the r.h.s. involves $\tilde{\Gamma}_r^{(1,2)}$ up to order g_r^L, since $\beta_2 \partial/\partial g_r$ is of order g_r, and β up to order g_r^{L+1}. Both are finite. The second term again involves $\tilde{\Gamma}_r^{(1,2)}$ up to order g_r^L, since $(\eta+\eta_2)$ is of order g_r, and $(\eta+\eta_2)$ at order g_r^{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 g_r^{L+1}, and η up to order g_r^L since $\tilde{\Gamma}_r^{(2,2)}$ is of order g_r. We conclude that the l.h.s. is finite up to loop order $L+1$.

The function $\tilde{\Gamma}_r^{(1,2)}$ is dimensionless. Its value at zero momentum is fixed by the renormalization condition (10.28), therefore,

$$\left(m_r \frac{\partial}{\partial m_r} - \varepsilon g_r \frac{\partial}{\partial g_r}\right) \tilde{\Gamma}_r^{(1,2)}(0;0,0) = 0.$$

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 when ε vanishes at loop order $L+1$. We now use equation (10.39) 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}.

10.6.4 The $\langle \phi\phi \rangle$ correlation function ($l = 0$, $n = 2$)

The term of order g_r^{L+1} in equation (10.43) is

$$\left[\left(m_r \frac{\partial}{\partial m_r} - \varepsilon g_r \frac{\partial}{\partial g_r}\right) \tilde{\Gamma}_r^{(2)}\right]_{L+1} = \left[\left(-\beta_2 \frac{\partial}{\partial g_r} + \eta\right) \tilde{\Gamma}_r^{(2)}\right]_{L+1} + m_r^2 \left[(2-\eta)\, \tilde{\Gamma}_r^{(1,2)}\right]_{L+1}. \tag{10.46}$$

In the first term of the r.h.s. we need only $\tilde{\Gamma}_r^{(2)}$ up to order g_r^L since it is multiplied by terms of order g_r. We have also shown above 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 above that the two factors are finite up to order g_r^{L+1}. Let us consider the quantity $\tilde{\Gamma}_r^{(2)}(p) - m_r^2 - p^2$. Renormalization conditions imply that it vanishes as p^4 for $p \to 0$. It has mass dimension two. We set

$$\left(m_r \frac{\partial}{\partial m_r} - \varepsilon g_r \frac{\partial}{\partial g_r}\right) \left[\tilde{\Gamma}_r^{(2)}(p) - m_r^2 - p^2\right] = m_r^2 f^{(2)}(p/m_r, g_r, \varepsilon) \tag{10.47}$$

with

$$\lim_{\varepsilon \to 0} f^{(2)}(p/m_\mathrm{r}, g_\mathrm{r}, \varepsilon) < \infty$$
$$f^{(2)}(p/m_\mathrm{r}, g_\mathrm{r}, \varepsilon) = O\left(p^4\right) \text{ for } p \to 0 .$$

For the same regularity reasons $f^{(2)}(p/\rho, g_\mathrm{r})$ decreases as $1/\rho^4$ for $\rho \to \infty$. We then integrate equation (10.47):

$$\tilde{\Gamma}_\mathrm{r}^{(2)}(p, m_\mathrm{r}, g_\mathrm{r}, \varepsilon) - m_\mathrm{r}^2 - p^2 = -\int_{m_\mathrm{r}}^{\infty} \rho \, \mathrm{d}\rho \, f^{(2)}(p/\rho, m_\mathrm{r}^\varepsilon g_\mathrm{r}/\rho^\varepsilon). \tag{10.48}$$

The integral is convergent and, therefore, the ε limit can be taken in the integral.

This concludes the induction. The advantage of the method is that we have both proven renormalizability and derived the CS equations:

$$\left[m_\mathrm{r} \frac{\partial}{\partial m_\mathrm{r}} + \beta(g_\mathrm{r}) \frac{\partial}{\partial g_\mathrm{r}} - \frac{n}{2}\eta(g_\mathrm{r}) - l\eta_2(g_\mathrm{r})\right] \tilde{\Gamma}_\mathrm{r}^{(l,n)}(q_j; p_i) = \left(2 - \eta(g_\mathrm{r})\right) m_\mathrm{r}^2 \times \tilde{\Gamma}_\mathrm{r}^{(l+1,n)}(0, q_j; p_i). \tag{10.49}$$

10.6.5 Superficially divergent correlation functions at large momentum

We see from this derivation that induction can also be used to estimate the large momentum behaviour of correlation functions. If the bounds (10.30) are assumed to hold at loop order L, then $\Gamma^{(1,4)}$ and $\Gamma^{(2,2)}$ 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.

We now consider again the example of the four-point function. Once we have established the representation (10.45), the induction hypothesis tells us that the function $f^{(4)}$ in the r.h.s. is bound by

$$\left|f^{(4)}(\lambda p_i/m_\mathrm{r}, g_\mathrm{r})\right| < \text{const. } (\ln \lambda)^{k(L)} \quad \text{for } \lambda \to \infty .$$

We then separate the integral over ρ into the sum of two terms:

$$\int_1^\infty \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_\mathrm{r}}, g_\mathrm{r}\right) = \int_\lambda^\infty \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_\mathrm{r}}, g_\mathrm{r}\right) + \int_1^\lambda \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_\mathrm{r}}, g_\mathrm{r}\right).$$

The first integral can be bound by a constant. The second integral can be bound using the large momentum behaviour of $f^{(4)}$:

$$\int_1^\lambda \frac{\mathrm{d}\rho}{\rho} \left| f^{(4)}\left(\frac{\lambda p_i}{\rho m_\mathrm{r}}, g_\mathrm{r}\right)\right| < \text{const. } \int_1^\lambda \frac{\mathrm{d}\rho}{\rho} \left[\ln\left(\frac{\lambda}{\rho}\right)\right]^{k(L)} \sim (\ln \lambda)^{k(L)+1} .$$

The argument is the same for $\Gamma^{(1,2)}$. This last bound can then be used to bound $\Gamma^{(2)}$.

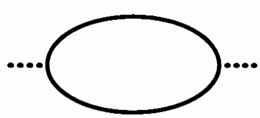


Fig. 10.21

with

$$\lim_{\varepsilon \to 0} f^{(2)}(p/m_{\mathrm{r}}, g_{\mathrm{r}}, \varepsilon) < \infty$$
$$f^{(2)}(p/m_{\mathrm{r}}, g_{\mathrm{r}}, \varepsilon) = O\left(p^4\right) \text{ for } p \to 0 .$$

For the same regularity reasons $f^{(2)}(p/\rho, g_{\mathrm{r}})$ decreases as $1/\rho^4$ for $\rho \to \infty$. We then integrate equation (10.47):

$$\tilde{\Gamma}^{(2)}_{\mathrm{r}}(p, m_{\mathrm{r}}, g_{\mathrm{r}}, \varepsilon) - m_{\mathrm{r}}^2 - p^2 = -\int_{m_{\mathrm{r}}}^{\infty} \rho\, \mathrm{d}\rho\, f^{(2)}(p/\rho, m_{\mathrm{r}}^{\varepsilon} g_{\mathrm{r}}/\rho^{\varepsilon}). \tag{10.48}$$

The integral is convergent and, therefore, the ε limit can be taken in the integral.

This concludes the induction. The advantage of the method is that we have both proven renormalizability and derived the CS equations:

$$\left[m_{\mathrm{r}} \frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}}) \frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2}\eta(g_{\mathrm{r}}) - l\eta_2(g_{\mathrm{r}})\right] \tilde{\Gamma}^{(l,n)}_{\mathrm{r}}(q_j; p_i) = \left(2 - \eta(g_{\mathrm{r}})\right) m_{\mathrm{r}}^2 \times \tilde{\Gamma}^{(l+1,n)}_{\mathrm{r}}(0, q_j; p_i). \tag{10.49}$$

10.6.5 Superficially divergent correlation functions at large momentum

We see from this derivation that induction can also be used to estimate the large momentum behaviour of correlation functions. If the bounds (10.30) are assumed to hold at loop order L, then $\Gamma^{(1,4)}$ and $\Gamma^{(2,2)}$ 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.

We now consider again the example of the four-point function. Once we have established the representation (10.45), the induction hypothesis tells us that the function $f^{(4)}$ in the r.h.s. is bound by

$$\left|f^{(4)}(\lambda p_i/m_{\mathrm{r}}, g_{\mathrm{r}})\right| < \text{const. } (\ln\lambda)^{k(L)} \quad \text{for } \lambda \to \infty .$$

We then separate the integral over ρ into the sum of two terms:

$$\int_1^{\infty} \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_{\mathrm{r}}}, g_{\mathrm{r}}\right) = \int_{\lambda}^{\infty} \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_{\mathrm{r}}}, g_{\mathrm{r}}\right) + \int_1^{\lambda} \frac{\mathrm{d}\rho}{\rho} f^{(4)}\left(\frac{\lambda p_i}{\rho m_{\mathrm{r}}}, g_{\mathrm{r}}\right).$$

The first integral can be bound by a constant. The second integral can be bound using the large momentum behaviour of $f^{(4)}$:

$$\int_1^{\lambda} \frac{\mathrm{d}\rho}{\rho} \left|f^{(4)}\left(\frac{\lambda p_i}{\rho m_{\mathrm{r}}}, g_{\mathrm{r}}\right)\right| < \text{const. } \int_1^{\lambda} \frac{\mathrm{d}\rho}{\rho} \left[\ln\left(\frac{\lambda}{\rho}\right)\right]^{k(L)} \sim (\ln\lambda)^{k(L)+1} .$$

The argument is the same for $\Gamma^{(1,2)}$. This last bound can then be used to bound $\Gamma^{(2)}$.

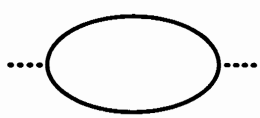


Fig. 10.21

10.7 The $\langle\phi^2\phi^2\rangle$ Correlation Function

We have seen that the $\langle\phi^2\phi^2\rangle$ proper vertex $\Gamma^{(2,0)}$ has superficial degree of divergence zero. Actually, even in free field theory ($g_\mathrm{r} = 0$) it is divergent (see figure 10.21). An additional renormalization is needed. We impose the renormalization condition:

$$\tilde{\Gamma}_\mathrm{r}^{(2,0)}(p = 0, m_\mathrm{r}, g_\mathrm{r}) = 0\,, \tag{10.50}$$

obviously consistent with the tree order approximation. We now expect the relation between bare and renormalized proper vertex to be

$$\tilde{\Gamma}_\mathrm{r}^{(2,0)}(p) = (Z_2/Z)^2\left[\tilde{\Gamma}^{(2,0)}(p) - \tilde{\Gamma}^{(2,0)}(0)\right]. \tag{10.51}$$

Differentiating with respect to m_r at g fixed we then obtain

$$\begin{aligned}&\left[m_\mathrm{r}\frac{\partial}{\partial m_\mathrm{r}} + \beta(g_\mathrm{r})\frac{\partial}{\partial g_\mathrm{r}} - 2\eta_2(g_\mathrm{r})\right]\tilde{\Gamma}_\mathrm{r}^{(2,0)}(p; m_\mathrm{r}, g_\mathrm{r})\\ &\quad = m_\mathrm{r}^2\left[2 - \eta(g_\mathrm{r})\right]\left[\tilde{\Gamma}_\mathrm{r}^{(3,0)}(p, -p, 0; m_\mathrm{r}, g_\mathrm{r}) - \tilde{\Gamma}_\mathrm{r}^{(3,0)}(0, 0, 0; m_\mathrm{r}, g_\mathrm{r})\right].\end{aligned} \tag{10.52}$$

The derivation of this equation follows the same lines as in previous cases. It uses the properties that $\Gamma^{(3,0)}$ has a skeleton expansion and that the L loop order of $\Gamma^{(2,0)}$ is of order g_r^{L-1}. Finally, in the integration with respect to m_r one takes into account the renormalization condition (10.50) and notes that $\tilde{\Gamma}^{(2,0)}$ is dimensionless in mass units.

It is possible then to summarize all CS equations by

$$\begin{aligned}&\left[m_\mathrm{r}\frac{\partial}{\partial m_\mathrm{r}} + \beta(g_\mathrm{r})\frac{\partial}{\partial g_\mathrm{r}} - \frac{n}{2}\eta(g_\mathrm{r}) - l\eta_2(g_\mathrm{r})\right]\tilde{\Gamma}_\mathrm{r}^{(l,n)}(q_j; p_i; m_\mathrm{r}, g_\mathrm{r})\\ &\quad = m_\mathrm{r}^2\left(2 - \eta(g_\mathrm{r})\right)\tilde{\Gamma}_\mathrm{r}^{(l+1,n)}(0, q_j; p_i; m_\mathrm{r}, g_\mathrm{r}) + \delta_{n0}\delta_{l2}m_\mathrm{r}^{-\varepsilon}B(g_\mathrm{r})\end{aligned} \tag{10.53}$$

with the notation

$$-m_\mathrm{r}^2\left(2 - \eta(g_\mathrm{r})\right)\tilde{\Gamma}_\mathrm{r}^{(3,0)}(0, 0, 0; m_\mathrm{r}, g_\mathrm{r}) = m_\mathrm{r}^{-\varepsilon}B(g_\mathrm{r}), \tag{10.54}$$

where the dimension of the l.h.s. has been used.

10.8 The Renormalized Action: General Construction

The proof of the renormalizability of arbitrary massive local field theories is a rather simple generalization of the proof we have given above for the ϕ^4 field theory in four dimensions. We, therefore, explain only the results.

We now consider an arbitrary field theory renormalizable by power counting. We assume that we have added to the tree order action $\mathcal{S}_{\mathrm{r},0}(\phi)$ all counter-terms needed to render the theory finite up to loop order L. If we then perform a loop expansion of the generating functional of proper vertices $\Gamma(\varphi)$,

$$\Gamma(\varphi) = \mathcal{S}_{\mathrm{r},0}(\varphi) + \sum_{l=1}^{\infty}\Gamma_l(\varphi),$$

the functionals $\Gamma_l(\varphi)$ have a finite limit for $l \le L$ when ε vanishes, and the diagrams contributing to $\Gamma_{L+1}(\varphi)$ have no divergent subdiagrams.

$\Gamma^{\mathrm{div}}_{L+1}(\varphi)$, the divergent part of $\Gamma_{L+1}(\varphi)$, is, therefore, a general local functional linear combination of all vertices of non-positive canonical dimensions (except if symmetries forbid some terms). By adding to the renormalized action the counter-terms $-\Gamma^{\mathrm{div}}_{L+1}(\varphi)$ we render the theory finite at loop order $L+1$. The counter-terms are of course only defined up to an arbitrary finite part, linear combination of the same vertices which appear in $\Gamma^{\mathrm{div}}_{L+1}$. It is sometimes convenient, in order to fix this arbitrariness, to impose specific renormalization conditions, as we have done in previous sections. *The resulting renormalized action is then also a general local functional of the fields, linear combination of all vertices of non-positive canonical dimension.*

Let us show in the case of the ϕ^4 field theory in four dimensions that this statement indeed summarizes the result derived in Section 10.6. The divergent part of $\Gamma(\varphi)$ at loop order L, after renormalization up to order $L-1$, should have the form

$$\Gamma^{\mathrm{div}}_L(\varphi) = -\int \mathrm{d}^4x \left[\frac{1}{2}\delta m^2_L \varphi^2 + \frac{1}{2}\delta Z_L \left(\partial_\mu \varphi\right)^2 + \frac{1}{4!} g_{\mathrm{r}} \delta Z_{g,L} \varphi^4\right], \tag{10.55}$$

because the vertex ϕ^2 has dimension -2, and the vertices $(\partial_\mu \phi)^2$ and ϕ^4 dimension 0. No odd powers of ϕ appear because the tree order action is symmetric in $\phi \mapsto -\phi$.

We now consider another example, a field theory containing a boson field ϕ, and fermion fields ψ, $\bar\psi$:

$$\mathcal{S}(\phi,\psi,\bar\psi) = \int \mathrm{d}^4x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} m^2 \phi^2 - \bar\psi \left(\not\partial + M\right)\psi - g\bar\psi\psi\phi\right]. \tag{10.56}$$

We have shown in Section 9.2.1 that the dimensions of fields in this theory are

$$[\phi] = 1, \quad [\psi] = [\bar\psi] = 3/2 \,,$$

and that the theory is renormalizable by power counting. Using power counting we can write the most general divergent term:

$$\begin{aligned} -\Gamma^{\mathrm{div}} = \int \mathrm{d}^4x \Big[&\frac{1}{2}\delta Z_\phi \left(\partial_\mu \varphi\right)^2 + \frac{1}{2}\delta m^2 \varphi^2 - \delta Z_\psi \bar\psi \not\partial \psi \\ &- \delta M \bar\psi\psi - g\delta Z_g \bar\psi\psi\varphi + \frac{1}{4!}\delta\lambda_4 \varphi^4 + \frac{1}{3!}\delta\lambda_3 \varphi^3 + \delta c \varphi\Big]. \end{aligned} \tag{10.57}$$

From these expressions, we see that terms not present in the tree order action (10.56) have been generated, proportional to $\int \phi^4 \mathrm{d}^4x$, $\int \phi^3 \mathrm{d}^4x$ and $\int \phi\, \mathrm{d}^4x$. We say that although action (10.56) is renormalizable, in contrast to the ϕ^4 field theory, it is not multiplicatively renormalizable, in the sense that not all coefficients of the renormalized action can be obtained by a rescaling of those of the tree order action. To avoid this difficulty and to always make the theory multiplicatively renormalizable, we shall in general try to include in the tree order action all terms which we expect renormalization to generate.

10.9 The Massless Theory

We have given a derivation of the renormalizability of a field theory which applies only to massive field theories, since the mass insertion operation has played an essential role in decreasing the degree of divergence of Feynman diagrams.

We now want to justify the existence of renormalized massless field theories. The correlation functions of a massless theory can be obtained by rescaling, at cut-off and momenta fixed, the mass

$$m_{\mathrm{r}} \mapsto m_{\mathrm{r}}/\rho, \qquad \rho \to \infty .$$

At fixed cut-off the limit exists, as will be discussed extensively in Chapters 24–27, 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 correlation function 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 field theory (at non-exceptional momenta) and we now examine what happens when the cut-off is removed. We use the CS equations combined with Weinberg's theorem. From now on, we omit the symbol $\tilde{}$, indicating a Fourier transform, when there is no ambiguity. The arguments x, y, z will indicate space variables and k, p, q, momenta, arguments of the Fourier transform.

10.9.1 Large momentum behaviour and massless theory

We write the CS equations:

$$\left[m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2}\eta(g_{\mathrm{r}})\right]\Gamma_{\mathrm{r}}^{(n)}(p_i) = m_{\mathrm{r}}^2(2-\eta)\Gamma_{\mathrm{r}}^{(1,n)}(0;p_i).$$

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 canonical dimension up to powers of logarithms. Thus,

$$\Gamma_{\mathrm{r}}^{(l,n)}(\rho p_i) \underset{\to\infty}{\sim} \rho^{4-n-2l} \times \text{power of } \ln\rho ,$$

$$\Gamma_{\mathrm{r}}^{(l+1,n)}(0;\rho p_i) \underset{\to\infty}{\sim} \frac{1}{\rho^2}\rho^{4-n-2l} \times \text{power of } \ln\rho .$$

If, therefore, in the asymptotic expansion of $\Gamma_{\mathrm{r}}^{(n)}(\rho p_i)$ for ρ large, we eliminate all terms subleading by a power ρ^{-2} up to powers of $\ln\rho$, we find a set of 1PI correlation functions $\Gamma_{\mathrm{r,as.}}^{(n)}$ which satisfy a homogeneous CS equation:

$$\left[m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2}\eta(g_{\mathrm{r}})\right]\Gamma_{\mathrm{r,as.}}^{(n)}(p_i) = 0 . \tag{10.58}$$

However, we know from dimensional analysis

$$\Gamma_{\mathrm{r}}^{(n)}(p_i, m_{\mathrm{r}}) = m_{\mathrm{r}}^{4-n}\Gamma_{\mathrm{r}}^{(n)}(p_i/m_{\mathrm{r}}, 1). \tag{10.59}$$

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 (10.58) are thus the correlation functions of the massless ϕ^4 field theory.

Perturbative solution of the homogeneous CS equations. It is actually interesting to study the structure of correlation functions implied by equation (10.58). We consider here, for illustration purposes, only the two-point function $\Gamma^{(2)}_{\mathrm{r,as.}}(p)$. It is convenient to introduce the function $\zeta(g_{\mathrm{r}})$:

$$\ln \zeta(g_{\mathrm{r}}) = \int_0^{g_{\mathrm{r}}} \frac{\mathrm{d}g'\,\eta(g')}{\beta(g')}.$$

Note that since both η and β are of order g_{r}^2 this function has a perturbative expansion.

We then set

$$\Gamma^{(2)}_{\mathrm{r,as.}}(p) = p^2 \zeta(g_{\mathrm{r}}) A(g_{\mathrm{r}}, p/m_{\mathrm{r}}).$$

The function A satisfies

$$\left[m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}}\right] A(g_{\mathrm{r}}, p/m_{\mathrm{r}}) = 0\,. \tag{10.60}$$

We expand A and $\beta(g_{\mathrm{r}})$ in powers of g_{r}, setting

$$A(g_{\mathrm{r}}, p/m_{\mathrm{r}}) = 1 + \sum_1^{\infty} g_{\mathrm{r}}^n a_n(p/m_{\mathrm{r}}), \qquad \beta(g_{\mathrm{r}}) = \sum_2^{\infty} \beta_n g_{\mathrm{r}}^n\,. \tag{10.61}$$

Introducing these expansions in equation (10.60) we find for $n \ge 2$,

$$-z a_n'(z) + \sum_{m=1}^{n-1} m a_m(z)\beta_{n-m+1} = 0\,. \tag{10.62}$$

The equation for $n = 1$ is special:

$$-z a_1'(z) = 0 \quad \Rightarrow \quad a_1(z) = C_1\,.$$

Let us examine the $n = 2$ case:

$$-z a_2'(z) + C_1\beta_2 = 0 \Rightarrow\; a_2(z) = C_1\beta_2 \ln z + C_2\,.$$

From this example, we understand the general structure: $a_n(z)$ is a polynomial of degree $(n-1)$ in $\ln z$:

$$a_n(z) = P_{n-1}(\ln z), \tag{10.63}$$

$$P_{n-1}'(x) = \sum_{m=1}^{n-1} m P_{m-1}(x)\beta_{n-m+1}\,. \tag{10.64}$$

Note, in particular, that 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 and so on. Finally, $\Gamma^{(2)}_{\mathrm{r,as.}}(p)$ is entirely determined by the functions $\beta(g_{\mathrm{r}})$ and $\eta(g_{\mathrm{r}})$ and, for example, $\Gamma^{(2)}_{\mathrm{r,as.}}(1, g_{\mathrm{r}})/m_{\mathrm{r}}^2$ which is a third function of g_{r}.

It also follows from these equations that $\Gamma^{(2)}_{\mathrm{r,as.}}(p)$ has a limit for $p = 0$:

$$\Gamma^{(2)}_{\mathrm{r,as.}}(p^2 = 0) = 0\,, \tag{10.65}$$

confirming, as expected, that the theory is massless. However, its derivative $\partial\Gamma^{(2)}_{\mathrm{r,as.}}/\partial p^2$ has no zero momentum limit. It is easy to verify that no other correlation function has a zero momentum limit either.

We have constructed, here, a massless theory by scaling a massive theory and shown that the corresponding 1PI functions satisfy a homogeneous CS equation, also called a renormalization group (RG) equation. We now show that such an equation can also be derived directly from the assumption of the existence of a renormalized massless field theory.

10.9.2 RG equations in a massless field theory

Renormalization conditions. We have shown that the renormalized massless ϕ^4 field theory exists in four dimensions. If we want to directly 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 above. We, therefore, introduce a mass scale μ and impose

$$\Gamma^{(2)}_{\mathrm{r}}(p^2 = 0) = 0\,, \tag{10.66}$$

$$\frac{\partial}{\partial p^2}\Gamma^{(2)}_{\mathrm{r}}(p^2 = \mu^2) = 1\,, \tag{10.67}$$

$$\Gamma^{(4)}_{\mathrm{r}}(p_i = \mu\theta_i) = \mu^{\varepsilon} g_{\mathrm{r}}\,, \tag{10.68}$$

in which the θ_i form a set of arbitrary non-exceptional numerical vectors.

The bare correlation functions in a massless theory depend only on the coupling constant (which provides the mass dimension) and momenta, since the bare mass parameter is fixed by imposing that the renormalized mass m_{r} vanishes. Note, and this will be discussed later, that at any finite order L in perturbation theory, the perturbation series exists only for $L\varepsilon < 2$, due to IR divergences. The domain shrinks with the order. The renormalized correlation functions depend on the arbitrary scale μ and momenta. They are related by

$$\Gamma^{(n)}_{\mathrm{r}}(p_i;\mu,g_{\mathrm{r}}) = Z^{n/2}(g_{\mathrm{r}},\varepsilon)\Gamma^{(n)}(p_i;g). \tag{10.69}$$

Remarks. While in a massive theory the value of the renormalized two-point function at zero momentum depends on the renormalization conditions, in a massless theory the bare two-point function and the renormalized function, independently of its normalization, vanish at zero momentum as equation (10.69) shows.

Renormalization group equations. The bare theory of course is completely independent of the parameter μ, which has just been introduced to fix the normalization of the renormalized functions:

$$\mu\frac{\partial}{\partial\mu}\bigg|_g \Gamma^{(n)}(p_i;g) = 0\,.$$

Therefore, if we differentiate equation (10.69) with respect to μ at g fixed, we find using the chain rule

$$\left[\mu\frac{\partial}{\partial\mu} + \tilde\beta(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2}\tilde\eta(g_{\mathrm{r}})\right]\Gamma^{(n)}_{\mathrm{r}} = 0 \tag{10.70}$$

with the definitions

$$\tilde{\beta}(g_{\rm r}) = \mu \frac{\partial}{\partial \mu}\bigg|_g g_{\rm r}, \qquad \tilde{\eta}(g_{\rm r}) = \mu \frac{\partial}{\partial \mu}\bigg|_g \ln Z(g_{\rm r}, \varepsilon). \tag{10.71}$$

A priori $\tilde{\beta}$ and $\tilde{\eta}$ could be singular for $\varepsilon \to 0$ but since they can be expressed in terms of the renormalized correlation functions themselves, they must have a finite limit.

Equation (10.70) is analogous to equation (10.58). They differ only in the definition of $g_{\rm r}$ and a finite field amplitude renormalization.

Both sets of equations will be an essential tool for the analysis of the large momentum behaviour of correlation functions. In addition, equation (10.70) will be used in the third part of this work to discuss the small momentum behaviour of massless theories.

In the massless theory, we can also define renormalized correlation functions with ϕ^2 insertions:

$$\Gamma_{\rm r}^{(l,n)}(y_1, \ldots, y_l; x_1, \ldots, x_n) = Z^{(n/2)-l} Z_2^l \Gamma^{(l,n)}(y_1, \ldots, y_l; x_1, \ldots, x_n). \tag{10.72}$$

The ϕ^2 renormalization constant Z_2 can be fixed by a renormalization condition of the form

$$\tilde{\Gamma}_{\rm r}^{(1,2)}(q; p_1, p_2)\Big|_{p_1^2 = p_2^2 = q^2 = \mu^2} = 1. \tag{10.73}$$

As before, the $\langle \phi^2 \phi^2 \rangle$ correlation function needs an additional renormalization, which can be fixed by a renormalization condition at non-zero momentum.

From equation (10.72) RG equations can be derived by differentiating with respect to μ. Introducing the RG function $\tilde{\eta}_2(g_{\rm r})$,

$$\tilde{\eta}_2(g_{\rm r}) = \mu \frac{\partial}{\partial \mu} \ln Z_2 \bigg|_{g}, \tag{10.74}$$

we obtain

$$\left[\mu \frac{\partial}{\partial \mu} + \tilde{\beta}(g_{\rm r}) \frac{\partial}{\partial g_{\rm r}} - \frac{n}{2} \tilde{\eta}(g_{\rm r}) - l \tilde{\eta}_2(g_{\rm r})\right] \Gamma_{\rm r}^{(l,n)}(q_1, \ldots, q_l; p_1, \ldots, p_n) = 0, \tag{10.75}$$

an equation valid except for $n = 0, l \le 2$ where the r.h.s. becomes a non-trivial function of $g_{\rm r}$ as in equation (10.53).

Dimensional and cut-off regularization. It is possible, with simple changes, to adapt the proof of existence of the renormalized theory to Paul–Villars's type regularizations. However, since we have now established this existence, we can argue more directly. We can define perturbatively a massless cut-off field theory in dimension d, in an infinitesimal neighbourhood of dimension 4. Unlike the theory defined by dimensional regularization, in $d = 4 - \varepsilon$ dimension with a cut-off the theory has divergences and requires a mass renormalization. After mass renormalization correlation functions converge at large cut-off towards those obtained by dimensional regularization. We can then introduce field and coupling constant renormalization, and we obtain the same renormalized theory by either taking first the large cut-off limit and then the $d = 4$ limit or the opposite.

10.10 Homogeneous RG Equations: Massive Theory

Once the renormalized correlation functions $\Gamma_{\mathrm{r}}^{(l,n)}$ of the massless theory have been constructed, a natural question arises: since a constant source term for ϕ^2 at zero momentum generates a mass shift (equation (10.21)), is it possible to express the correlation functions of a massive theory in terms of the correlation functions with ϕ^2 insertions of the massless theory?

An immediate difficulty arises: it is easy to verify that insertions at zero momentum in a massless theory are IR divergent. However, because the resulting theory is massive, this difficulty can be circumvented by first using a non-constant source for ϕ^2, then performing a partial summation of the two-point function and finally taking the constant source limit. After summation the propagator becomes massive and the limit is no longer IR divergent.

We, therefore, consider the renormalized action with a source $K(x)$ for renormalized ϕ^2 insertions:

$$\mathcal{S}_{\mathrm{r}}(\phi_{\mathrm{r}},K)=\int \mathrm{d}^4x\left[\frac{1}{2}Z\left(\partial_\mu\phi_{\mathrm{r}}\right)^2+\frac{1}{2}\left(\delta m^2+Z_2K(x)\right)\phi_{\mathrm{r}}^2+\frac{1}{4!}g_{\mathrm{r}}Z_g\phi_{\mathrm{r}}^4\right], \qquad (10.76)$$

where the renormalization constants are those of the massless theory. A correlation function in presence of the source $K(x)$ has the expansion

$$\begin{aligned}\Gamma_{\mathrm{r}}^{(n)}(p_1,p_2,\dots p_n;K)=\sum_{l=0}\frac{1}{2^l}\frac{1}{l!}\int \mathrm{d}q_1\mathrm{d}q_2\dots\mathrm{d}q_l\,\tilde{K}(q_1)\tilde{K}(q_2)\dots\tilde{K}(q_l)&\\ \times\Gamma_{\mathrm{r}}^{(l,n)}(q_1,q_2,\dots,q_l;p_1,p_2,\dots,p_n),&\qquad (10.77)\end{aligned}$$

in which $\tilde{K}$ is the Fourier transform of the source.

We apply the differential operator D:

$$\mathrm{D}\equiv\mu\frac{\partial}{\partial\mu}+\tilde{\beta}(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}}, \qquad (10.78)$$

on equation (10.77) and use the RG equations (10.75). Noting that

$$\int \mathrm{d}q\,\tilde{K}(q)\frac{\delta}{\delta\tilde{K}(q)}=l, \qquad (10.79)$$

we obtain

$$\left[\mu\frac{\partial}{\partial\mu}+\tilde{\beta}(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}}-\frac{n}{2}\tilde{\eta}(g_{\mathrm{r}})-\tilde{\eta}_2(g_{\mathrm{r}})\int\mathrm{d}q\,\tilde{K}(q)\frac{\delta}{\delta\tilde{K}(q)}\right]\Gamma_{\mathrm{r}}^{(n)}(K;p_1,\dots,p_n)=0. \qquad (10.80)$$

After the summation of the propagator, we set

$$K(x)=m_{\mathrm{r}}^2, \qquad (10.81)$$

and equation (10.80) becomes

$$\left[\mu\frac{\partial}{\partial\mu}+\tilde{\beta}(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}}-\frac{n}{2}\tilde{\eta}(g_{\mathrm{r}})-\tilde{\eta}_2(g_{\mathrm{r}})m_{\mathrm{r}}^2\frac{\partial}{\partial m_{\mathrm{r}}^2}\right]\Gamma_{\mathrm{r}}^{(n)}(p_1,\dots,p_n;\mu,g_{\mathrm{r}},m_{\mathrm{r}})=0. \qquad (10.82)$$

We have, therefore, derived a new RG equation for a massive theory which differs from the original CS equations in various respects:

(i) Correlation functions depend on two mass parameters while we know that only one is necessary. However, in this parametrization the massless limit of correlation functions is directly obtained by setting $m_r = 0$ and no asymptotic expansion at large momenta is needed.

(ii) In contrast to the CS equations, these RG equations are homogeneous. We shall exploit this property later when we solve the RG equations.

10.11 Covariance of RG Functions

In the example of the massless ϕ^4 we have been naturally led to consider two different renormalization schemes. It is thus necessary to understand how renormalization group functions of different schemes are related. Both theories differ by a redefinition of the coupling constant and a finite field amplitude renormalization. We call g and $\tilde{g}$ the renormalized coupling constants in the two schemes. Comparing the two renormalization group equations, we obtain

$$\beta(g)\frac{\partial}{\partial g} = \tilde{\beta}(\tilde{g})\frac{\partial}{\partial \tilde{g}},$$

and, therefore, using the chain rule,

$$\beta(g)\frac{\partial \tilde{g}}{\partial g} = \tilde{\beta}(\tilde{g}). \tag{10.83}$$

In a ϕ^4-like field theory in four dimensions, the function $\beta(g)$ has the expansion

$$\beta(g) = \beta_2 g^2 + \beta_3 g^3 + O\left(g^4\right). \tag{10.84}$$

Expanding $\tilde{g}$ in terms of g

$$\tilde{g} = g + \gamma_2 g^2 + O\left(g^3\right), \tag{10.85}$$

and using equation (10.84), after a short calculation, we find

$$\tilde{\beta}(\tilde{g}) = \beta_2 \tilde{g}^2 + \beta_3 \tilde{g}^3 + O\left(\tilde{g}^4\right). \tag{10.86}$$

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 result that the physical consequences derived from the form of the RG β-function are also to a large extent arbitrary.* Only change of variables $g \mapsto \tilde{g}$, which are regular mappings, are allowed. This implies in particular that $\partial\tilde{g}/\partial g$ in equation (10.83) must remain strictly positive. Therefore, the sign and zeros of the β-function are properties of the theory. Equation (10.83) shows also that if the β-function vanishes with a finite slope the slope is scheme-independent.

The difficulty arises from the fact that 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 $\zeta(\tilde{g})$ the additional finite field amplitude renormalization we find

$$\tilde{\eta}(\tilde{g}) = \eta(g) + \tilde{\beta}(\tilde{g})\frac{\partial}{\partial \tilde{g}}\ln\zeta(\tilde{g}). \tag{10.87}$$

Since the field amplitude renormalization appears at order g^2, $\ln \zeta(\tilde{g})$ is of order $\tilde{g}^2$ and, therefore, the modification to $\tilde{\eta}$ of order $\tilde{g}^3$. The coefficient of order g^2 is universal. The value of η at a zero of $\beta(g)$ is also universal.

Several coupling constants. We shall meet actions depending on several fields and coupling constants g_i. The transformation law under a change of parametrization of the coupling space of the RG β-functions then becomes

$$\tilde{\beta}_i(\tilde{g}) = \frac{\partial \tilde{g}_i}{\partial g_j} \beta_j(g), \tag{10.88}$$

where the mapping should satisfy $\det(T_{ij} \equiv \partial \tilde{g}_i / \partial g_j) \neq 0$. Then the existence of zeros of the β-function is universal and at a zero g^*

$$\frac{\partial \tilde{\beta}_i}{\partial \tilde{g}_j} = T_{ik} \frac{\partial \beta_k}{\partial g_l} T^{-1}_{lj},$$

and the eigenvalues of $\partial \beta_i / \partial g_j$ at a zero are thus parametrization independent.

Bibliographical Notes

The method we have adopted here to prove renormalizability of the ϕ^4 field theory follows the presentation of

C.G. Callan in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

The idea to combine proof of renormalizability and RG equations is due to

A.S. Blaer and K. Young, *Nucl. Phys.* B83 (1974) 493.

A systematic discussion of renormalization theory in a different language as well as historical references can be found in

W. Zimmermann, *Lectures on Elementary Particles and Quantum Field Theory, Brandeis 1970*, S. Deser, M. Grisaru and H. Pendleton eds. (MIT Press, Cambridge 1970); *Ann. of Phys. (NY)* 77 (1973) 536.

Among the early work see, for example,

F.J. Dyson, *Phys. Rev.* 75 (1949) 1736; N.N. Bogoliubov and O.S. Parasiuk, *Acta Math.* 97 (1957) 227; and the textbook N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, (Interscience, New York 1959).

The asymptotic behaviour of renormalized integrals was studied in

S. Weinberg, *Phys. Rev.* 118 (1960) 838; K. Hepp, *Commun. Math. Phys.* 2 (1966) 301.

For different aspects of this problem in parametric space see also

T. Appelquist, *Ann. of Phys. (NY)* 54 (1969) 27; M.C. Bergère and J.-B. Zuber, *Commun. Math. Phys.* 35 (1974) 113; M. Bergère and Y.M.P. Lam, *Commun. Math. Phys.* 39 (1974) 1.

Early references about RG equations are

E.C.G. Stueckelberg and A. Peterman, *Helv. Phys. Acta* 26 (1953) 499; M. Gell-Mann and F.E. Low, *Phys. Rev.* 95 (1954) 1300.

See also N.N. Bogoliubov and D.V. Shirkov as quoted above and

K.G. Wilson, *Phys. Rev.* 179 (1969) 1499.

In their modern form the RG or CS equations have been derived independently by

C.G. Callan, *Phys. Rev.* D2 (1970) 1541; K. Symanzik, *Commun. Math. Phys.* 18 (1970) 227.

A pedagogical presentation has been given in

S. Coleman, *Dilatations, Erice Lectures 1971* reprinted in *Aspects of Symmetry* (Cambridge University Press, Cambridge 1985).

Massless theories have been discussed in

K. Symanzik, *Commun. Math. Phys.* 7 (1973) 34.

Homogeneous RG equations have been proposed in

S. Weinberg, *Phys. Rev.* D8 (1973) 3497; G. 't Hooft, *Nucl. Phys.* B61 (1973) 455; J. Zinn-Justin, *Cargèse Lectures 1973* unpublished, later incorporated in the review: E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Field Theoretical Approach to Critical Phenomena* in *Phase Transitions and Critical Phenomena* vol. 6, C. Domb and M.S. Green eds. (Academic Press, New York 1976).

See also

J.C. Collins, *Nucl. Phys.* B80 (1974) 341.

In the appendix we briefly outline a, different very interesting approach which combines renormalization and renormalization group; see

J. Polchinski, *Nucl. Phys.* B231 (1984) 269.

Similar RG equations were introduced first in

F.J. Wegner and A. Houghton, *Phys. Rev.* A8 (1973) 401; F.J. Wegner, *J. Physics* C7 (1974) 2098.

See also the review

F.J. Wegner, in *Phase Transitions and Critical Phenomena* vol. 6, C. Domb and M.S. Green eds. (Academic Press, New York 1976).

For recent applications of such ideas see for example

M. Bonini, M. D'Attanasio and G. Marchesini, *Nucl. Phys.* B409 (1993) 441, *ibidem* B444 (1995) 602; J. Adams, J. Berges, S. Bornholdt, F. Freire, N. Tetradis and C. Wetterich, *Mod. Phys. Lett.* A10 (1995) 2367; T.R. Morris, *Prog. Theor. Phys. Suppl.* 131 (1998) 395, hep-th/9802039,

and references therein.

APPENDIX A10

A10.1 Large Momentum Mode Integration and General RG Equations

Here, we briefly describe a direct approach to renormalization and renormalization group, closer to Wilson's ideas (in a form also developed by Wegner), which was first proposed by Polchinski. This approach leads to another proof of renormalizability. The physical motivation for such a method will become more apparent when we discuss RG and critical phenomena starting in Chapter 25. The idea is to study the cut-off dependence of bare correlation functions by integrating out systematically the large momentum modes of the field. By such a procedure one constructs an exact renormalization group 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. This equivalence results into RG equations for correlation functions valid for large distance or low momentum (in the cut-off scale).

A10.1.1 A simple equivalence

In Section 9.5, we have shown by a gaussian integration that the two actions $\mathcal{S}(\phi)$,

$$\mathcal{S}(\phi)=\frac{1}{2}\int \mathrm{d}x\,\phi\Delta^{-1}\phi+V(\phi), \tag{A10.1}$$

and $\mathcal{S}(\phi_1,\phi_2)$,

$$\mathcal{S}(\phi_1,\phi_2)=\frac{1}{2}\int \mathrm{d}x\,\left(\phi_1\Delta_1^{-1}\phi_1+\phi_2\Delta_2^{-1}\phi_2\right)+V(\phi_1+\phi_2) \tag{A10.2}$$

with

$$\Delta=\Delta_1+\Delta_2\,, \tag{A10.3}$$

generate the same perturbation theory.

We now use this equivalence in the limit in which the propagator Δ_2 goes to zero. Then only small values of ϕ_2 contribute to the partition function. Expanding the interaction for ϕ_2 small

$$V(\phi_1+\phi_2)=V(\phi_1)+\int \mathrm{d}x\frac{\delta V(\phi_1)}{\delta\phi(x)}\phi_2(x)+\frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\frac{\delta^2 V}{\delta\phi(x)\delta\phi(y)}\phi_2(x)\phi_2(y)+\cdots\,,$$

we integrate over ϕ_2 to obtain the leading order correction,

$$\begin{aligned}\int [\mathrm{d}\phi_2]\exp\left\{-\left[\frac{1}{2}\int \mathrm{d}x\,\phi_2\Delta_2^{-1}\phi_2+V(\phi+\phi_2)-V(\phi)\right]\right\}&\sim 1\\ &+\frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\,\Delta_2(x,y)\left[\frac{\delta V}{\delta\phi(x)}\frac{\delta V}{\delta\phi(y)}-\frac{\delta^2 V}{\delta\phi(x)\delta\phi(y)}\right]+\cdots\,.\end{aligned} \tag{A10.4}$$

Taking the logarithm of expression (A10.4), we rewrite the partition function as a functional integral over a field ϕ (the index 1 is no longer useful) with an effective action $\mathcal{S}'(\phi)$:

$$\mathcal{S}'(\phi)=\frac{1}{2}\int \mathrm{d}x\,\phi\Delta_1^{-1}\phi+V(\phi)+\frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\,\Delta_2(x,y)\left[\frac{\delta^2 V}{\delta\phi(x)\delta\phi(y)}-\frac{\delta V}{\delta\phi(x)}\frac{\delta V}{\delta\phi(y)}\right]+\cdots\,. \tag{A10.5}$$

We have thus established that the actions ($A10.1$) and ($A10.5$) lead to the same partition function, up to a trivial renormalization. Note that the same identity can be proven by playing with integrations by part as in the case of the quantum equation of motion (7.32).

We show below that this equivalence can be used to partially integrate out the large momenta in a field theory with a cut-off Λ.

A10.1.2 Large momentum mode partial integration and RG equations

We take for propagator Δ a massless propagator of the form

$$\Delta(k) = \frac{C(k^2/\Lambda^2)}{k^2}, \tag{A10.6}$$

in which the function $C(t)$ is smooth, goes to 1 for t small and decreases faster than any power for t large (see Section 9.5). We then take for Δ_1 the propagator corresponding to an infinitesimal rescaling of the cut-off Λ:

$$\Delta_1(k) = \frac{C(k^2/\Lambda^2(1+\sigma)^2)}{k^2}. \tag{A10.7}$$

At leading order, for σ small, Δ_2 is then

$$\Delta_2(k) = \frac{2\sigma}{\Lambda^2} C'(k^2/\Lambda^2) = \sigma D(k). \tag{A10.8}$$

We see that the propagator Δ_2 *has no pole at* $k=0$. Moreover, if we choose a function $C(t)$ which is very close to 1 for t small

$$|C(t)-1|\,t^{-p} \to 0 \quad \forall p>0\,,$$

then Δ_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 ($A10.1$) and ($A10.5$), which is the starting point of a renormalization group, can be written as

$$\Lambda\frac{\mathrm{d}}{\mathrm{d}\Lambda} V(\phi,\Lambda) = \frac{1}{2}\int \mathrm{d}^d x\, \mathrm{d}^d y\, D(x,y)\left[\frac{\delta^2 V}{\delta\phi(x)\delta\phi(y)} - \frac{\delta V}{\delta\phi(x)}\frac{\delta V}{\delta\phi(y)}\right], \tag{A10.9}$$

or after Fourier transformation,

$$\Lambda\frac{\mathrm{d}}{\mathrm{d}\Lambda} V(\phi,\Lambda) = \frac{1}{2}\int \frac{\mathrm{d}^d k}{(2\pi)^d} D(k)\left[\frac{\delta^2 V}{\delta\phi(k)\delta\phi(-k)} - \frac{\delta V}{\delta\phi(k)}\frac{\delta V}{\delta\phi(-k)}\right]. \tag{A10.10}$$

Note that the equation can also be derived from the quantum field equations and, therefore, partial integration does not imply any loss of information, unlike what would happen on a lattice.

To study the existence of fixed points we start with a given interaction $V_0(\phi)$ at a scale Λ_0 and use equation ($A10.9$) 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 call $\tilde{V}^{(n)}(p_1, p_2, \dots, p_n)$ the coefficients of $V(\phi, \Lambda)$ in an expansion in powers of $\tilde{\phi}(p)$, the Fourier transform of the field. Equation (A10.10) can then be written in component form (assuming translation invariance):

$$\Lambda \frac{\mathrm{d}}{\mathrm{d}\Lambda} \tilde{V}^{(n)}(p_1, p_2, \dots, p_n) = \frac{1}{2} \int \frac{\mathrm{d}^d k}{(2\pi)^d} D(k) \tilde{V}^{(n+2)}(p_1, p_2, \dots, p_n, k, -k)$$
$$- \frac{1}{2} \sum_I D(p_0) \tilde{V}^{(l+1)}(p_{i_1}, \dots, p_{i_l}, p_0) \tilde{V}^{(n-l+1)}(p_{i_{l+1}}, \dots, p_{i_n}, -p_0), \quad (A10.11)$$

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 $\{1, 2, \dots, n\}$.

We see in these equations 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 (A10.9) 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 we have to add a source to the interaction $V(\phi)$:

$$V(\phi) \mapsto V(\phi) - \int \mathrm{d}x \, J(x)\phi(x).$$

However, equation (A10.5) 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 $\int \mathrm{d}x \, J(x)\phi_2(x)$ does not contribute in integral (A10.4) and $\mathcal{S}'(\phi) - \mathcal{S}(\phi)$ does not depend on $J(x)$.

We note, however, 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.

A10.2 Super-Renormalizable Field Theories: The Normal-Ordered Product

In most of this work we discuss strictly renormalizable field theories. Let us, however, make a few simple comments about the super-renormalizable case. We first take the example of the ϕ^4 field theory in three dimensions. We then examine the special properties of two-dimensional field theories.

The ϕ^4 theory in three dimensions. As derived in Chapter 9, the ϕ^4 field theory in three dimensions has only three superficially divergent diagrams, all contributing to the two-point function, which are shown in figure 10.22.

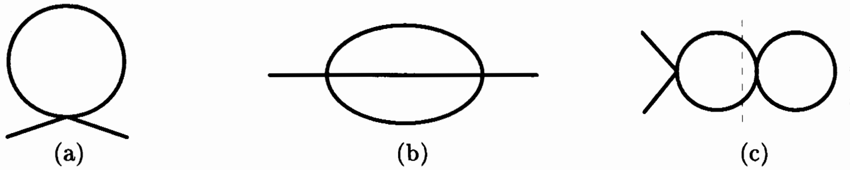


Fig. 10.22 The three divergent diagrams in the ϕ_3^4 theory.

$$(\mathrm{a}) = \frac{1}{(2\pi)^3}\int \frac{\mathrm{d}^3 q}{(q^2+m^2)_\Lambda}\,.$$

$$(\mathrm{b}) = \frac{1}{(2\pi)^6}\int \frac{\mathrm{d}^3 q_1 \mathrm{d}^3 q_2}{(q_1^2+m^2)_\Lambda\,(q_2^2+m^2)_\Lambda\left[(p-q_1-q_2)^2+m^2\right]_\Lambda}\,.$$

By adding a counter-term of the form

$$\tfrac{1}{2}\left[a_1(\Lambda)g + a_2(\Lambda)g^2\right]\phi^2,$$

we render the first two diagrams finite.

The diagram (c) is also superficially divergent, but it is clear that the counter-term which renders diagram (a) finite, also renormalizes diagram (c). Thus, the counter-terms which renormalize 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* $:\phi^4:$

$$:\phi^4:(x) = \phi^4(x) - 6\phi^2(x)\left\langle\phi^2(x)\right\rangle + 3\left(\left\langle\phi^2(x)\right\rangle\right)^2, \tag{A10.12}$$

in which the expectation value $\left\langle\phi^2(x)\right\rangle$ is calculated in a free field theory with a mass μ which may or may not be equal to the renormalized mass m:

$$\left\langle\phi^2(x)\right\rangle_\mu = \frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3 q}{(q^2+\mu^2)_\Lambda}. \tag{A10.13}$$

The denomination normal ordering comes from the operator language. The quantity $:\phi^4(x):$ is such that

$$\begin{aligned}\left\langle:\phi^4(x):\right\rangle &= 0\\ \left\langle:\phi^4(x):\phi(y_1)\phi(y_2)\right\rangle &= 0\,,\end{aligned} \tag{A10.14}$$

in which again the expectation values are calculated with the action $\mathcal{S}_\mu(\phi)$:

$$\mathcal{S}_\mu(\phi) = \frac{1}{2}\int \mathrm{d}^3 x\left[(\nabla\phi)^2 + \mu^2\phi^2\right]. \tag{A10.15}$$

The expectation values calculated with another mass are then finite.

Finally, we still have to add a counter-term for diagram (b).

Super-renormalizable scalar field theories in two dimensions. An action $\mathcal{S}(\phi)$ of the form

$$\mathcal{S}(\phi) = \int \mathrm{d}^2 x\left[\frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2 + V(\phi)\right], \tag{A10.16}$$

in which $V(\phi)$ is an arbitrary function of $\phi(x)$ (but not of its derivatives) 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 figure 10.23). Therefore, the operation $:V(\phi):$ removes all divergences.

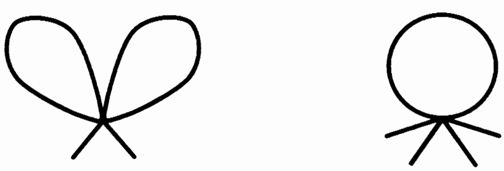


Fig. 10.23

To obtain an explicit expression for $: V(\phi) :$ we first consider the special interaction

$$V(\phi) = \mathrm{e}^{\lambda\phi} . \tag{A10.17}$$

The expectation value of $V(\phi)$ in the presence of a source term can then be calculated explicitly:

$$\begin{aligned}
&\int [\mathrm{d}\phi(y)] \exp\left\{-\int \mathrm{d}y \left[\frac{1}{2}\left((\nabla\phi)^2 + \mu^2\phi^2\right) - (J(y) + \lambda\delta(x-y))\,\phi(y)\right]\right\} \\
&\quad = \exp\left[\frac{1}{2}\int \mathrm{d}y\,\mathrm{d}y' J(y)\Delta(y,y')J(y') + \lambda\int \mathrm{d}y\,\Delta(x,y)J(y) + \frac{1}{2}\lambda^2\Delta(x,x)\right],
\end{aligned} \tag{A10.18}$$

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

$$: \exp\left[\lambda\phi(x)\right] : = \exp\left[\lambda\phi(x) - \lambda^2\left\langle\phi^2(x)\right\rangle/2\right], \tag{A10.19}$$

in which we have used

$$\left\langle\phi^2(x)\right\rangle = \Delta(x,x). \tag{A10.20}$$

We then write an arbitrary interaction term as a Laplace transform:

$$V\big(\phi(x)\big) = \int \mathrm{d}\rho(\lambda)\,\mathrm{e}^{\lambda\phi(x)} . \tag{A10.21}$$

The normal ordering is a linear operation. Thus,

$$: V(\phi) : = \int \mathrm{d}\rho(\lambda) : \mathrm{e}^{\lambda\phi(x)} : \ .$$

We then use the result ($A10.19$) for the exponential interaction and obtain

$$: V(\phi) : = \left\{\exp\left[-\tfrac{1}{2}\left\langle\phi^2\right\rangle(\partial/\partial\phi)^2\right]\right\} V(\phi), \tag{A10.22}$$

or more explicitly,

$$: V(\phi) : = V(\phi) + \left[\sum_{n=1}^{\infty}\frac{(-1)^n}{2^n n!}\left\langle\phi^2\right\rangle^n\left(\frac{\partial}{\partial\phi}\right)^{2n}\right] V(\phi).$$

The existence of the Laplace transform of a given interaction is irrelevant in this argument since the final identity is purely algebraic.

DIMENSIONAL REGULARIZATION, MINIMAL SUBTRACTION: RG FUNCTIONS

Dimensional regularization is a powerful regularization technique, and often the most convenient for practical perturbative calculations, when applicable. It has been defined and a few of its properties have been discussed in Section 9.6. It has then been used to prove the renormalizability of the ϕ^4 field theory in Chapter 10. We now introduce the concept of renormalization by minimal subtraction, within the framework of dimensional regularization.

We first discuss the structure of renormalization constants and renormalization group functions $\beta(g)$, $\eta(g)$, $\eta_2(g)$, and then show that it is specially simple in the minimal subtraction scheme. We perform explicit calculations at two-loop order first in the simple one-component ϕ^4 field and then in an N-component field theory with a general four-field interaction. These calculations will be useful for the theory of Critical Phenomena.

Finally, to give an example of a theory involving fermions, we calculate the RG functions at one-loop order in a theory containing fermions interacting through a Yukawa-like interaction with a scalar boson: the Gross–Neveu–Yukawa model.

11.1 Renormalization Group (RG) Functions

We again discuss the example of the ϕ^4 field theory with the bare action:

$$\mathcal{S}(\phi_0) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi_0\right)^2 + \frac{1}{2} m_0^2 \phi_0^2 + \frac{1}{4!} g_0 \phi_0^4\right], \tag{11.1}$$

within the framework of dimensional regularization. Following the discussion of Chapter 10, 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

$$\mathcal{S}_\mathrm{r}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2} Z(g)\left(\partial_\mu \phi\right)^2 + \frac{1}{2} m^2 Z_m(g) \phi^2 + \frac{1}{4!} m^{4-d} g Z_g(g) \phi^4\right]. \tag{11.2}$$

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.

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

$$m_0^2 = m^2 Z_m(g) \,/ Z(g), \tag{11.3}$$

$$g_0 = g\, m^{4-d} Z_g(g) \,/ Z^2(g). \tag{11.4}$$

Remember that in the dimensional regularization scheme renormalization constants depend on an additional hidden parameter, the dimension d.

From the relation (11.3) we can now calculate the RG β-function, the coefficient of the CS equation. Setting

$$g\, Z_g \,/ Z^2 = G(g), \tag{11.5}$$

and differentiating equation (11.4) with respect to m at g_0 fixed, we find (equation (10.32))

$$0 = (4-d)\, G(g) + \beta(g) \frac{\partial}{\partial g} G(g),$$

or,

$$\beta(g) = -\,(4-d) \left(\frac{\mathrm{d} \ln G(g)}{\mathrm{d} g} \right)^{-1} . \tag{11.6}$$

Using the notation of Section 10.2, we call $Z_2(g)$ the renormalization constant associated with the renormalization of ϕ^2. From equations (10.33,10.34) we then derive

$$\eta(g) = \beta(g) \frac{\mathrm{d}}{\mathrm{d}g} \ln Z(g), \tag{11.7}$$

$$\eta_2(g) = \beta(g) \frac{\mathrm{d}}{\mathrm{d}g} \ln \left[Z_2(g) \,/\, Z(g) \right], \tag{11.8}$$

and, finally, from equations (11.3) and (10.35)

$$\frac{Z_m}{Z_2} \left[2 + \beta(g) \frac{\mathrm{d}}{\mathrm{d}g} \ln(Z_m/Z) \right] = \sigma(g). \tag{11.9}$$

This equation shows that Z_m/Z_2 is a finite function of g, and, therefore, Z_m is not a new renormalization constant.

The renormalizability of the ϕ^4 field theory in four dimensions implies that the renormalized correlation 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

$$G(g) = g + O\left(g^2\right),$$

the function $\beta(g)$ can be written as

$$\beta(g) = -\varepsilon g + \beta_2(\varepsilon) g^2 + \beta_3(\varepsilon) g^3 + \cdots, \tag{11.10}$$

in which all the functions $\beta_n(\varepsilon)$ have regular Taylor series expansion at $\varepsilon = 0$:

$$\beta_n(\varepsilon) = \beta_n(0) + \varepsilon \beta_n'(0) + \cdots .$$

11.2 Dimensional Regularization: The Form of Renormalization Constants

Conversely, we now determine the form of $G(g)$ from the knowledge of $\beta(g)$:

$$g \frac{G'(g)}{G(g)} = -\frac{\varepsilon g}{\beta(g)} \equiv \left[1 - \frac{1}{\varepsilon} \beta_2(\varepsilon) g - \frac{1}{\varepsilon} \beta_3(\varepsilon) g^2 - \cdots \right]^{-1} .$$

If we expand the r.h.s. 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 $\beta(g)$:

$$\frac{G'(g)}{G(g)} = \frac{1}{g} + \left(\frac{\beta_2(0)}{\varepsilon} + O(1) \right) + g \left(\frac{\beta_2^2(0)}{\varepsilon^2} + O\left(\frac{1}{\varepsilon}\right) \right) + \cdots .$$

Integrating the expansion term by term we get

$$G(g) = g + \sum_{n=2} g^n \left[\left(\frac{\beta_2(0)}{\varepsilon} \right)^{n-1} + \text{ less singular terms} \right].$$

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

$$Z(g) = 1 + \sum_1^\infty \frac{\alpha^{(n)}(g)}{\varepsilon^n} + \text{ regular terms in } \varepsilon,$$

$$Z_2(g) = 1 + \sum_1^\infty \frac{\alpha_2^{(n)}(g)}{\varepsilon^n} + \text{ regular terms in } \varepsilon,$$

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 $\Gamma(\varphi)$ is a polynomial of degree L in $1/\varepsilon$.

11.3 Minimal Subtraction Scheme

Although the minimal subtraction idea can be used in any regularization scheme (see, for example, equation (9.8)), it is specially useful, in particular for critical phenomena, in dimensional regularization. Renormalization constants are determined in the following way: instead of imposing renormalization conditions to divergent 1PI correlation functions, one just subtracts, at each order in the loop expansion, the singular part of Laurent expansion in ε. Let us denote by $\Gamma_L^{\text{div}}(\varphi)$ the divergent part of the generating functional of proper vertices, renormalized up to $L-1$ loops:

$$\Gamma_L^{\text{div}}(\varphi) = \sum_{\ell=1}^{L} \frac{\gamma_{L,\ell}}{\varepsilon^\ell}. \tag{11.11}$$

We just add, as a counter-term to the action, $-\Gamma_L^{\text{div}}(\varphi)$ as defined above.

Example. We again consider the ϕ^4 field theory. At one-loop order only two diagrams contributing to $\Gamma^{(2)}$ and $\Gamma^{(4)}$ are divergent. The one-loop divergent part of $\Gamma(\varphi)$ is given by equation (10.24):

$$\Gamma_1^{\text{div}}(\varphi) = -\frac{1}{32\pi^2\varepsilon} \left[m^2 g \int \varphi^2(x) \mathrm{d}^d x + \frac{1}{4} g^2 m^\varepsilon \int \varphi^4(x) \mathrm{d}^d x \right]. \tag{11.12}$$

This expression should be compared to equations ($A9.17$) and (10.24). The functions Z, Z_m, Z_g at this order then are

$$Z = 1 + O\left(g^2\right), \tag{11.13}$$

$$Z_g = 1 + \frac{3}{16\pi^2}\frac{g}{\varepsilon} + O\left(g^2\right), \tag{11.14}$$

$$Z_m = 1 + \frac{1}{16\pi^2}\frac{g}{\varepsilon} + O\left(g^2\right). \tag{11.15}$$

The calculation of the function $\langle \phi^2 \phi\phi \rangle$ at one-loop also leads to

$$Z_2 = 1 + \frac{1}{16\pi^2}\frac{g}{\varepsilon} + O\left(g^2\right). \tag{11.16}$$

The RG functions are then

$$\beta(g) = -\varepsilon g + \frac{3}{16\pi^2}g^2 + O\left(g^3\right), \tag{11.17}$$

$$\eta(g) = O\left(g^2\right), \tag{11.18}$$

$$\eta_2(g) = -\frac{g}{16\pi^2} + O\left(g^2\right). \tag{11.19}$$

We note that at this order $Z_m = Z_2$. Since the minimal subtraction scheme eliminates any possible finite renormalization and Z_2/Z_m is finite, this relation remains true to all orders.

Let us explore more generally the consequences of this choice for the RG functions. Writing for instance

$$G(g) = g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n}, \quad G_n(g) = O\left(g^{n+1}\right), \tag{11.20}$$

we can calculate $\beta(g)$:

$$\beta(g) = -\varepsilon\left[g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n}\right]\left[1 + \sum_1^\infty \frac{G'_n(g)}{\varepsilon^n}\right]^{-1}.$$

Since $G'_n(g)$ is of order g^n, we can expand the denominator:

$$\beta(g) = -\varepsilon\left[g + \sum_1^\infty \frac{G_n(g)}{\varepsilon^n}\right]\left[1 - \frac{G'_1(g)}{\varepsilon} + \frac{\left[G'_1(g)\right]^2}{\varepsilon^2} - \frac{G'_2(g)}{\varepsilon^2} + \cdots\right].$$

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

$$\beta(g) = -\varepsilon g + gG'_1(g) - G_1(g), \tag{11.21}$$

$$b_n(g) = 0 \quad \forall n. \tag{11.22}$$

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.

Similar arguments apply for the other RG functions and renormalization constants.

In the expansion of equation (11.11) the whole new L-loop information about divergences is contained in $\gamma_{L,1}(\varphi)$. All other functions are determined by the counter-terms of previous orders.

In addition, the RG functions $\beta(g)$, $\eta(g)$ and $\eta_2(g)$ have a very simple dependence on ε. For $\beta(g)$ it is given by equation (11.21). Let us calculate η and η_2. The renormalization constants Z and Z_2 now have the form

$$Z(g) = 1 + \sum_1^{\infty} \frac{\alpha^{(n)}(g)}{\varepsilon^n}, \quad \alpha^{(n)}(g) = O\left(g^{n+1}\right), \tag{11.23}$$

$$Z_2(g) = 1 + \sum_1^{\infty} \frac{\alpha_2^{(n)}(g)}{\varepsilon^n}, \quad \alpha_2^{(n)}(g) = O\left(g^{n}\right). \tag{11.24}$$

Using relation (11.7) and the form (11.21) of $\beta(g)$ we obtain

$$\eta(g) = \left[-\varepsilon g + g G_1'(g) - G_1(g)\right] \left[\frac{1}{\varepsilon}\frac{\mathrm{d}}{\mathrm{d}g}\alpha^{(1)}(g) + O\left(\frac{1}{\varepsilon^2}\right)\right].$$

Since $\eta(g)$ has a finite limit, it is given by

$$\eta(g) = -g\frac{\mathrm{d}}{\mathrm{d}g}\alpha^{(1)}(g). \tag{11.25}$$

Similarly for $\eta_2(g)$ we find

$$\eta_2(g) = -g\frac{\mathrm{d}}{\mathrm{d}g}\alpha_2^{(1)}(g). \tag{11.26}$$

Finally, the explicit dependence of the renormalization constants on ε can be obtained by calculating them from β, η and η_2. For example,

$$G(g) = g \exp\left(-\varepsilon \int_0^g \left[\frac{1}{-\varepsilon g' + b(g')} + \frac{1}{\varepsilon g'}\right] \mathrm{d}g'\right), \tag{11.27}$$

in which we have set

$$\beta(g, \varepsilon) = -\varepsilon g + b(g). \tag{11.28}$$

The modified minimal subtraction scheme. In the calculation of low order Feynman diagrams, a factor $(N_d)^L$,

$$N_d = \frac{\text{area of the sphere } S_{d-1}}{(2\pi)^d} = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}, \tag{11.29}$$

L being the number of loops of the diagram, is generated naturally. It is, therefore, convenient to rescale the loop expansion parameter to suppress this factor. In the ϕ^4 field theory, for example, two choices are possible. Either one multiplies each Feynman diagram by a factor $(N_4/N_d)^L$, then the normalizations of field and coupling constant in four dimensions are only modified by the change in the renormalization constants. Or one completely absorbs the factor N_d in a coupling constant and field redefinition.

11.4 The Massless Theory

Dimensional regularization can also be used to define the massless ϕ^4 field theory. Unlike, however, the massive theory, due to small momentum (IR) divergences the regularized theory only exists in an infinitesimal neighbourhood of the dimension 4. This problem will be discussed at length in Chapters 25–26 devoted to critical phenomena. Then, since the massless theory is renormalizable, the minimal subtraction scheme is also applicable.

The bare action can be written as

$$\mathcal{S}(\phi_0) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi_0(x)\right)^2 + \frac{1}{4!} g_0 \phi_0^4(x)\right]. \tag{11.30}$$

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:

$$\tilde{\Gamma}^{(2)}(p) = p^2 + \sum_{n=2} C_n(\varepsilon) g_0^n p^{2-n\varepsilon}$$

(remember that ε is infinitesimal).

To define a renormalized theory it is necessary to introduce a mass scale μ which will take care of the dimension of the ϕ^4 coupling constant:

$$\mathcal{S}_{\mathrm{r}}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2} Z(g)\left(\partial_\mu \phi\right)^2 + \frac{1}{4!}\mu^{4-d} g\, Z_g \phi^4\right]. \tag{11.31}$$

To the action, corresponds a set of relations between bare and renormalized correlation functions:

$$Z^{n/2}(g)\Gamma_{\mathrm{r}}^{(n)}(p_i;\mu,g) = \Gamma^{(n)}(p_i;g_0) \tag{11.32}$$

with

$$g_0 = \mu^{4-d} g Z_g \,/ Z^2\,. \tag{11.33}$$

Differentiation of equation (11.32) with respect to μ at g_0 fixed, yields RG equations which express that correlation functions depend on μ and g only through the combination $g_0(\mu,g)$:

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right]\tilde{\Gamma}_{\mathrm{r}}^{(n)}(p_i;\mu,g) = 0\,, \tag{11.34}$$

where β and η are given by expressions similar to equations (11.6,11.7):

$$\beta(g) = -(4-d)\left(\frac{\mathrm{d}\ln G(g)}{\mathrm{d}g}\right)^{-1}, \qquad \eta(g) = \beta(g)\frac{\mathrm{d}}{\mathrm{d}g}\ln Z(g). \tag{11.35}$$

As we have discussed in Section 10.10, it is then possible to define a massive theory by adding to the renormalized action a mass term of the form

$$\mathcal{S}_{\mathrm{r}}(m) = \mathcal{S}_{\mathrm{r}} + \frac{1}{2}m^2 \int \mathrm{d}^d x\, Z_2(g)\phi^2(x).$$

Note that since in the minimal subtraction scheme the renormalization constants are uniquely defined and independent of the ratio m/μ, one concludes by setting $m=\mu$ that the renormalization constants in the massless theory and the massive theory of Section 11.1 are the same.

11.5 RG Functions at Two-Loop Order in the ϕ^4 Field Theory

As an exercise, we now explicitly calculate the two-loop renormalization constants in the ϕ^4 field theory, and in the modified minimal subtraction scheme (see Appendix A11.1). We work with the massive theory though the calculations are somewhat easier in the massless theory. Note that for higher order calculations even more sophisticated methods can be used, where the masses in diagrams are chosen in the following way: only diagrams to which all divergent subdiagrams have been subtracted are considered. The remaining global divergence is then independent of momenta and internal masses. One sets to zero as many masses and momenta as possible, as long as one does not encounter IR divergences, consistently in the diagram and the subtracted subdiagrams.

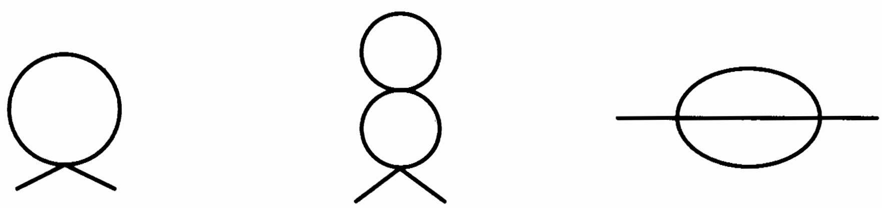
11.5.1 The diagrams

In Section 7.2.3 we have already given the diagrams contributing to the two-point function at this order. A short calculation gives also $\Gamma^{(4)}$:

$$\Gamma^{(2)}(p) = p^2 + m^2 + \frac{1}{2} g\,(\mathrm{a}) - \frac{g^2}{4}\,(\mathrm{b}) - \frac{g^2}{6}\,(\mathrm{c}) + O\left(g^3\right), \tag{11.36}$$

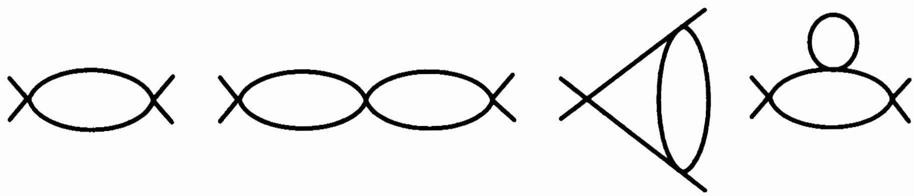
$$\Gamma^{(4)}(0,0,0,0) = g - \frac{3}{2} g^2\,(\mathrm{d}) + \frac{3}{4} g^3\,(\mathrm{e}) + 3g^3\,(\mathrm{f}) - \frac{3}{2} g^3\,(\mathrm{g}) + O\left(g^4\right). \tag{11.37}$$

To determine all renormalization constants, we thus need the diagrams displayed in figures 11.1 and 11.2.



(a) (b) (c)

Fig. 11.1 Diagrams contributing to $\Gamma^{(2)}$.



(d) (e) (f) (g)

Fig. 11.2 Diagrams contributing to $\Gamma^{(4)}$.

The diagram (g) contributing to $\Gamma^{(4)}$ corresponds to the insertion of the one-loop $\Gamma^{(2)}$ diagram into the one-loop $\Gamma^{(4)}$ diagram. When the one-loop diagrams are renormalized, it is automatically finite. At this order, only the diagram (c) has a divergence which is momentum dependent:

$$(\mathrm{c}) \equiv \Sigma^{(2)}(p,m) = \frac{1}{(2\pi)^{2d}} \int \frac{\mathrm{d}^d q_1\, \mathrm{d}^d q_2}{(q_1^2+m^2)(q_2^2+m^2)\left[(p+q_1+q_2)^2+m^2\right]}\,. \tag{11.38}$$

To renormalize the two-point function, we need (a), (b), $\Sigma(0,m)$ and $\partial\Sigma(p,m)/\partial p^2$ at $\mathbf{p}=0$. We first set

$$(\mathrm{a}) \equiv \frac{1}{(2\pi)^d}\int \frac{\mathrm{d}^d q}{q^2+m^2} = m^{d-2} I_1\,, \tag{11.39}$$

$$\Sigma^{(2)}(0,m) = m^{2d-6} I_2\,, \tag{11.40}$$

$$\left.\frac{\partial}{\partial p^2}\Sigma^{(2)}(p,m)\right|_{\mathbf{p}=0} = m^{2d-8} I_3\,. \tag{11.41}$$

The diagram (b) as well as the diagrams (d)–(f) at zero momentum can be expressed in terms of the integrals I_1 and I_2:

$$\begin{aligned}
&(\mathrm{d}) \equiv \frac{1}{(2\pi)^d}\int \frac{\mathrm{d}^d q}{(q^2+m^2)^2} = -\frac{\partial}{\partial m^2}(\mathrm{a}) = (1-d/2)m^{d-4} I_1\,,\\
&(\mathrm{b}) \equiv (\mathrm{a})(\mathrm{d}) = (1-d/2)m^{2d-6} I_1^2\,,\\
&(\mathrm{e}) \equiv [(\mathrm{d})]^2 = (1-d/2)^2 m^{2d-8} I_1^2\,,
\end{aligned} \tag{11.42}$$

and, finally,

$$\begin{aligned}
(\mathrm{f}) &= \frac{1}{(2\pi)^{2d}}\int \frac{\mathrm{d}^d q_1\,\mathrm{d}^d q_2}{(q_1^2+m^2)(q_2^2+m^2)\left[(q_1+q_2)^2+m^2\right]^2} = -\frac{1}{3}\frac{\partial}{\partial m^2}\Sigma^{(2)}(0,m)\\
&= -\tfrac{1}{3}(d-3)m^{2d-8} I_2\,.
\end{aligned} \tag{11.43}$$

The integral I_1. The integral is a special example of (A11.5). As explained in Section A11.1, we extract a factor N_d (equation (11.29)) for each loop. Setting then

$$d = 4-\varepsilon\,,$$

we find

$$I_1 = N_d \frac{\pi}{2\sin \pi d/2} = N_d\left(-\frac{1}{\varepsilon} + O(\varepsilon)\right). \tag{11.44}$$

Calculation of I_2. To derive an expression for both I_2 and I_3 we consider $\Sigma^{(2)}(p,m)$:

$$\Sigma^{(2)}(p,m=1) = \frac{1}{(2\pi)^{2d}}\int \frac{\mathrm{d}^d q_1\,\mathrm{d}^d q_2}{(q_1^2+1)(q_2^2+1)\left[(p+q_1+q_2)^2+1\right]}\,. \tag{11.45}$$

Using the method explained in Section 9.6 we write

$$\Sigma^{(2)}(p,1) = \frac{1}{(2\pi)^{2d}}\int_0^\infty \mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3 \int \mathrm{d}^d q_1 \mathrm{d}^d q_2 \exp\left[-A(q_1,q_2,p)\right] \tag{11.46}$$

with

$$A(q_1;q_2,p) = (t_1+t_3)q_1^2 + (t_2+t_3)q_2^2 + 2t_3 q_1 q_2 + 2t_3 p(q_1+q_2) + t_3 p^2 + (t_1+t_2+t_3)\,. \tag{11.47}$$

Integrating over q_1 and q_2 we find (equation (9.65))

$$\Sigma^{(2)}(p,1) = \frac{1}{t(4\pi)^d}\int_0^\infty \frac{\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3}{\Delta^{d/2}} \exp\left[-(t_1+t_2+t_3) - t_1 t_2 t_3 p^2/\Delta\right], \tag{11.48}$$

in which Δ is the determinant associated with the quadratic form

$$\Delta = t_1 t_2 + t_2 t_3 + t_3 t_1 \,. \tag{11.49}$$

We only need the first two terms of the expansion in powers of p^2. Let us first calculate I_2 which is more complicated:

$$I_2 = \Sigma^{(2)}(0,1) = \frac{1}{(4\pi)^d} \int_0^\infty \frac{\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}t_3}{\Delta^{d/2}}\, \mathrm{e}^{-(t_1+t_2+t_3)} \,. \tag{11.50}$$

As explained in Section A11.1 we now change variables, setting:

$$\begin{aligned} t_1 &= stu\,, \\ t_2 &= st(1-u), \\ t_3 &= s(1-t). \end{aligned} \tag{11.51}$$

The integral becomes

$$I_2 = \frac{1}{(4\pi)^d} \int_0^\infty \mathrm{d}s\, \mathrm{e}^{-s}\, s^{2-d} \int_0^1 \mathrm{d}u\,\mathrm{d}t \frac{t^{1-d/2}}{[1-t+tu(1-u)]^{d/2}}. \tag{11.52}$$

The integral over s can be performed. The remaining integral J, from which we want to extract the divergent and the finite parts, has divergences coming from $t=0$ and $t=1$, $u=0$ and $u=1$.

$$J = \int_0^1 \mathrm{d}u\,\mathrm{d}t\, t^{1-d/2}\, [1-t+tu(1-u)]^{-d/2} \,. \tag{11.53}$$

To separate divergences, we subtract and add 1 to each factor in the product:

$$\begin{aligned} J = \int_0^1 \mathrm{d}u\,\mathrm{d}t\, &\Big\{1 + \left(t^{1-d/2} - 1\right) + \left[\left(1-t+tu(1-u)\right)^{-d/2} - 1\right] \\ &+ \left(t^{1-d/2} - 1\right)\left[\left(1-t+tu(1-u)\right)^{-d/2} - 1\right]\Big\}. \end{aligned} \tag{11.54}$$

The last term is finite and we can set $d=4$. The others can be integrated immediately over t:

$$\begin{aligned} J = \frac{2}{\varepsilon} - 1 + \frac{1}{(1-\varepsilon/2)} &\int_0^1 \mathrm{d}u \left\{[u(1-u)]^{1-d/2} - 1\right\} [1-u(1-u)]^{-1} \\ + &\int_0^1 \mathrm{d}u\,\mathrm{d}t \frac{(1-t)}{t} \left[\left(1-t+tu(1-u)\right)^{-2} - 1\right]. \end{aligned} \tag{11.55}$$

The last integral in the expression yields

$$\int_0^1 \mathrm{d}u\,\mathrm{d}t \frac{(1-t)}{t} \left[\left(1-t+tu(1-u)\right)^{-2} - 1\right] = -\int_0^1 \mathrm{d}u \ln[u(1-u)] = 2\,.$$

A last piece has to be worked out:

$$\begin{aligned} &\int_0^1 \frac{\mathrm{d}u}{1-u(1-u)} \left\{[u(1-u)]^{1-d/2} - 1\right\} \\ &\quad = \int_0^1 \mathrm{d}u \left\{[u(1-u)]^{1-d/2} - 1\right\} + \int_0^1 \frac{\mathrm{d}u\, u(1-u)}{1-u(1-u)} \left[\frac{1}{u(1-u)} - 1\right] + O(\varepsilon)\,. \end{aligned}$$

Therefore,

$$\int_0^1 \frac{du}{1-u(1-u)}\left\{[u(1-u)]^{1-d/2}-1\right\}=\int_0^1 du\,[u(1-u)]^{1-d/2}+O(\varepsilon)\,. \tag{11.56}$$

We now use the identity (A11.6) and collect all contributions:

$$J=\frac{2}{\varepsilon}-1+\frac{1}{(1-\varepsilon/2)}\frac{\Gamma^2(\varepsilon/2)}{\Gamma(\varepsilon)}+2+O(\varepsilon)\,. \tag{11.57}$$

Since

$$\frac{\Gamma^2(\varepsilon/2)}{\Gamma(\varepsilon)}=\frac{4}{\varepsilon}+O(\varepsilon)\,, \tag{11.58}$$

we obtain the final result:

$$J=\frac{6}{\varepsilon}+3+O(\varepsilon)\,. \tag{11.59}$$

This leads, for I_2, to

$$I_2=N_d^2\,\frac{1}{4}\Gamma^2(d/2)\Gamma(3-d)\frac{6}{\varepsilon}\left(1+\frac{\varepsilon}{2}\right)+O(1)\,. \tag{11.60}$$

Using

$$\Gamma^2(d/2)\Gamma(3-d)=-\frac{(1-\varepsilon/2)^2}{\varepsilon(1-\varepsilon)}+O(\varepsilon)\,, \tag{11.61}$$

we deduce the singular part of integral I_2:

$$I_2=-N_d^2\,\frac{3}{2\varepsilon^2}\left(1+\frac{\varepsilon}{2}\right)+O(1)\,. \tag{11.62}$$

The integral I_3. The calculation of I_3 relies on the same method but is simpler. I_3 is the coefficient of p^2 in expression (11.48):

$$I_3=\frac{\partial}{\partial p^2}\Sigma^{(2)}(p,1)\Big|_{\mathbf{p}=0}=-\frac{1}{(4\pi)^d}\int_0^\infty \frac{dt_1\,dt_2\,dt_3\,t_1t_2t_3}{\Delta^{1+d/2}}\,e^{-(t_1+t_2+t_3)}\,.$$

After the change of variables (11.51) and the integration over s we find

$$I_3\sim-\frac{1}{(4\pi)^d}\Gamma(\varepsilon)\int_0^1 dt\,du\frac{(1-t)u(1-u)}{[1-t+ut(1-u)]^3}\,. \tag{11.63}$$

A short calculation then leads to

$$I_3\sim-N_d^2\frac{1}{8\varepsilon}+O(1)\,. \tag{11.64}$$

The four-point function at one-loop order. As a final exercise, let us also calculate explicitly the finite part of the one-loop four-point function which is given by the integral $I_4(p)$:

$$I_4(p)=\frac{m^{4-d}}{(2\pi)^d}\int\frac{d^dq}{(q^2+m^2)\left[(p+q)^2+m^2\right]}\,. \tag{11.65}$$

The calculation has already been performed in Section 9.6 for the case $m = 0$. The same method leads to

$$I_4(p) = N_d \frac{1}{2}(1 - d/2)\frac{\pi}{\sin \pi d/2} \int_0^1 \mathrm{d}t \left[\frac{p^2}{m^2}t(1-t) + 1\right]^{d/2-2} . \tag{11.66}$$

Expanding now for ε small, we obtain

$$I_4(p) = N_d \left[\frac{1}{\varepsilon} - \frac{1}{2} - \frac{1}{2}\int_0^1 \mathrm{d}t \ln\left(\frac{p^2}{m^2}t(1-t)+1\right) + O(\varepsilon)\right]. \tag{11.67}$$

Taking the finite part and performing explicitly the last integral we find

$$\left(I_4(p) - \frac{N_d}{\varepsilon}\right)_{\varepsilon=0} = -\frac{1}{8\pi^2}\left[\frac{1}{2}\sqrt{\frac{p^2+4m^2}{p^2}} \ln \frac{\sqrt{p^2+4m^2}+\sqrt{p^2}}{\sqrt{p^2+4m^2}-\sqrt{p^2}} - \frac{1}{2}\right]. \tag{11.68}$$

11.5.2 Renormalization constants and RG functions

We now introduce the renormalization constants, that is, replace in expressions (11.36) and (11.37) the propagator $1/(p^2+m^2)$ by

$$\left(p^2+m^2\right)^{-1} \mapsto \left(Zp^2 + m^2 Z_2\right)^{-1},$$

and the coupling constant g by

$$g \mapsto gZ_g \, .$$

We then expand the renormalization constants in powers of the coupling constant g and adjust the coefficients to cancel all divergences.

Renormalization constants. Integral I_3 immediately yields Z:

$$Z = 1 - N_d^2 \frac{g^2}{48\varepsilon} + O\left(g^3\right). \tag{11.69}$$

Expressing that $\Gamma^{(4)}$ is finite, we obtain Z_g:

$$Z_g = 1 + N_d \frac{3g}{2\varepsilon} + N_d^2\left(\frac{9}{4\varepsilon^2} - \frac{3}{4\varepsilon}\right) g^2 + O\left(g^3\right). \tag{11.70}$$

Finally, the propagator at zero momentum determines Z_2:

$$Z_2 = 1 + N_d \frac{g}{2\varepsilon} + N_d^2\left(\frac{1}{2\varepsilon^2} - \frac{1}{8\varepsilon}\right) g^2 + O\left(g^3\right). \tag{11.71}$$

RG functions. Equations (11.6–11.8) then yield the RG functions:

$$\tilde\beta(\tilde g) = N_d \beta(\tilde g) = -\varepsilon \tilde g + \frac{3}{2}\tilde g^2 - \frac{17}{12}\tilde g^3 + O\left(\tilde g^4\right), \tag{11.72}$$

$$\eta(\tilde g) = \frac{\tilde g^2}{24} + O\left(\tilde g^3\right), \tag{11.73}$$

$$\eta_2(\tilde g) = -\frac{\tilde g}{2} + \frac{5\tilde g^2}{24} + O\left(\tilde g^3\right) \tag{11.74}$$

with the notation

$$\tilde g = N_d g \, . \tag{11.75}$$

Field amplitude renormalization: positivity of the RG function. Note that for g small, that is, in the perturbative domain, the field renormalization (11.69) satisfies $Z < 1$. This is a general property in unitary theories implied by the spectral representation of the two-point function (see Section 6.9). It implies $\eta(g) > 0$ for g small.

11.6 Generalization to Several Component Fields

We now generalize previous calculations to a renormalized action of the form

$$\mathcal{S}(\phi)=\int \mathrm{d}^d x \left[\frac{1}{2} Z_{ij}\partial_\mu \phi_i \partial_\mu \phi_j + \frac{1}{4!} G_{ijkl}\phi_i\phi_j\phi_k\phi_l\right]. \tag{11.76}$$

Here, we consider only the massless theory and call μ the renormalization scale. We again use dimensional regularization and minimal subtraction. This explains the absence of mass counter-terms. We call g_{ijkl} the renormalized coupling constant, which is a symmetric tensor in its four indices. The quantity G_{ijkl}, which includes coupling constant renormalization, has the form

$$G_{ijkl} = \mu^{\varepsilon} g_{ijkl} + O\left(g^2\right). \tag{11.77}$$

To calculate the generalization of the renormalization constant Z_2 we use the renormalization of the insertion of the operator $\frac{1}{2}\phi_i(x)\phi_j(x)$.

The calculation of the renormalization constants reduces to the calculation of weight factors, the values of the divergent parts of the diagrams being the same as in Section 11.5.

11.6.1 Diagrams and renormalization constants

The two-point function has at two-loop order the expansion

$$\Gamma^{(2)}_{ij} = Z_{ij}p^2 - \tfrac{1}{6}G_{iklm}G_{jklm}\ (\mathrm{c}') + O\left(G^3\right). \tag{11.78}$$

The four-point function is at the same order:

$$\begin{aligned}\Gamma^{(4)}_{ijkl} &= G_{ijkl} - \tfrac{1}{2}\left(G_{ijmn}G_{mnkl}\,(\mathrm{d}') + 2 \text{ terms}\right) + \tfrac{1}{4}\left(G_{ijmn}G_{mnpq}G_{pqkl}\,(\mathrm{e}') + 2 \text{ terms}\right)\\ &\quad + \tfrac{1}{2}\left(G_{ijmn}G_{mpqk}G_{npql}\,(\mathrm{f}') + 5 \text{ terms}\right) + O\left(G^4\right).\end{aligned} \tag{11.79}$$

The additional terms restore the permutation symmetry of the four-point function in its four indices. The quantities (c′)–(f′) are, in the massless theory, the analogues of the diagrams of figures 11.1 and 11.2 (this accounts for the primes).

Let us, finally, calculate the $\frac{1}{2}\langle\phi_i(x)\phi_j(x)\phi_k(y_1)\phi_l(y_2)\rangle$ correlation function $\Gamma^{(1,2)}_{ij,kl}$:

$$\begin{aligned}\Gamma^{(1,2)}_{ij,kl} &= \tfrac{1}{2}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) - \tfrac{1}{2}G_{ijkl}\,(\mathrm{d}') + \tfrac{1}{4}G_{ijmn}G_{mnkl}\,(\mathrm{e}')\\ &\quad + \tfrac{1}{4}\left(G_{ikmn}G_{jlmn} + G_{jkmn}G_{ilmn}\right)(\mathrm{f}') + O\left(G^3\right).\end{aligned} \tag{11.80}$$

As we have argued in Section 11.4, the divergent parts of the massive diagrams of Section 11.5 and the massless theory are the same. Expanding the renormalization constants in powers of g_{ijkl}, and using the values of the divergent parts of diagrams (c), (d) and (f) of figures 11.1 and 11.2, we obtain

$$\begin{aligned}\mu^{-\varepsilon}G_{ijkl} &= g_{ijkl} + \frac{N_d}{2\varepsilon}\left(g_{ijmn}g_{mnkl} + 2 \text{ terms}\right) + \frac{N_d^2}{4\varepsilon^2}\left(g_{ijmn}g_{mnpq}g_{pqkl} + 2 \text{ terms}\right)\\ &\quad + \frac{N_d^2}{4\varepsilon^2}\left(1-\frac{\varepsilon}{2}\right)\left(g_{ijmn}g_{mpqk}g_{npql} + 5 \text{ terms}\right) + O\left(g^4\right),\end{aligned} \tag{11.81}$$

$$Z_{ij} = \delta_{ij} - \frac{N_d^2}{48\varepsilon}g_{iklm}g_{jklm} + O\left(g^3\right). \tag{11.82}$$

From equations (11.81,11.82) we deduce the expansion of the bare coupling constant, which we parametrize as $\mu^\varepsilon \gamma_{ijkl}$:

$$\gamma_{ijkl} = \mu^{-\varepsilon} G_{mnpq} \left(Z^{-1/2}\right)_{mi} \left(Z^{-1/2}\right)_{jn} \left(Z^{-1/2}\right)_{kp} \left(Z^{-1/2}\right)_{lq} . \tag{11.83}$$

We obtain

$$\begin{aligned}\gamma_{ijkl} = g_{ijkl} &+ \frac{N_d}{2\varepsilon}\left(g_{ijmn}g_{mnkl} + 2 \text{ terms}\right) + \frac{N_d^2}{4\varepsilon^2}\left(g_{ijmn}g_{mnpq}g_{pqkl} + 2 \text{ terms}\right) \\ &+ \frac{N_d^2}{4\varepsilon^2}\left(1-\frac{\varepsilon}{2}\right)\left(g_{ijmn}g_{mpqk}g_{npql} + 5 \text{ terms}\right) \\ &+ \frac{N_d^2}{96\varepsilon}\left(g_{ijkm}g_{mnpq}g_{npql} + 3 \text{ terms}\right) + O\left(g^4\right). \end{aligned} \tag{11.84}$$

Calling $Z^{(2)}_{ij,kl}$ the renormalization constant of the operator $\frac{1}{2}\phi_i(x)\phi_j(x)$, we derive from equation (11.80)

$$\begin{aligned} Z^{(2)}_{ij,kl} = \frac{1}{2}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) &+ \frac{N_d}{2\varepsilon} g_{ijkl} + \frac{N_d^2}{4\varepsilon^2} g_{ijmn}g_{mnkl} \\ &+ \frac{N_d^2}{8\varepsilon^2}\left(1-\frac{\varepsilon}{2}\right)\left(g_{ikmn}g_{jlmn} + g_{jkmn}g_{ilmn}\right) + O\left(g^3\right). \end{aligned} \tag{11.85}$$

Actually we need the matrix $\zeta^{(2)}_{ij,kl}$ which expresses the renormalized operator in terms of the bare fields:

$$\zeta^{(2)}_{ij,kl} = Z^{(2)}_{ij,mn} \left(Z^{-1/2}\right)_{mk} \left(Z^{-1/2}\right)_{nl} . \tag{11.86}$$

At two-loop order we find

$$\zeta^{(2)}_{ij,kl} = Z^{(2)}_{ij,kl} + \frac{N_d^2}{192\varepsilon}\left(g_{imnp}g_{kmnp}\delta_{jl} + 3 \text{ terms}\right) + O\left(g^3\right). \tag{11.87}$$

The additional terms restore the symmetry of exchange $(i \leftrightarrow j)$ and $(k \leftrightarrow l)$ and of the two pairs $(ij) \leftrightarrow (kl)$.

11.6.2 Renormalization group equations

Let us first derive the RG equations for a multicomponent massless field theory. The relation between bare and renormalized correlation functions now takes the form:

$$\Gamma^{(n)}_{i_1 i_2 \ldots i_n}(p_k, g, \mu) = \left(Z^{1/2}\right)_{i_1 j_1} \left(Z^{1/2}\right)_{i_2 j_2} \cdots \left(Z^{1/2}\right)_{i_n j_n} \Gamma^{(n)}_{\mathrm{B}, j_1 j_2 \ldots j_n}(p_k, g_0, \Lambda), \tag{11.88}$$

in which B stands for bare, g for g_{ijkl} and g_0 for $\mu^\varepsilon \gamma_{ijkl}$.

If we differentiate equation (11.88) with respect to μ at g_0 and Λ fixed, we obtain the RG equation:

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_{i'j'k'l'}\frac{\partial}{\partial g_{i'j'k'l'}}\right)\Gamma^{(n)}_{i_1 i_2 \ldots i_n} - \frac{1}{2}\sum_{m=1}^{n} \eta_{i_m j_m} \Gamma^{(n)}_{i_1 i_2 \ldots j_m \ldots i_n} = 0 \tag{11.89}$$

with the definitions

$$\beta_{i'j'k'l'}\frac{\partial \gamma_{ijkl}}{\partial g_{i'j'k'l'}} = -\varepsilon \gamma_{ijkl} , \tag{11.90}$$

$$\eta_{ij} = 2\beta_{i'j'k'l'} \left(\frac{\partial Z^{1/2}}{\partial g_{i'j'k'l'}} Z^{-1/2} \right)_{ij} . \tag{11.91}$$

Similarly, to renormalize correlation functions with $\frac{1}{2}\phi_i(x)\phi_j(x)$ insertions we have to multiply each insertion by the matrix $\zeta^{(2)}_{ij,kl}$. This leads to the RG equation:

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_{i'j'k'l'} \frac{\partial}{\partial g_{i'j'k'l'}}\right) \Gamma^{(l,n)}_{j_1k_1\ldots j_lk_l,i_1\ldots i_n} - \frac{1}{2}\sum_{m=1}^{n} \eta_{i_m a_m} \Gamma^{(l,n)}_{j_1k_1\ldots j_lk_l,i_1\ldots a_m\ldots i_n}$$
$$- \sum_{m=1}^{l} \eta^{(2)}_{j_mk_m,b_mc_m} \Gamma^{(l,n)}_{j_1k_1\ldots b_mc_m\ldots j_lk_l,i_1\ldots i_n} = 0 \tag{11.92}$$

with the definition

$$\eta^{(2)}_{ij,kl} = \beta_{i'j'k'l'} \left[\frac{\partial \zeta^{(2)}}{\partial g_{i'j'k'l'}} \left(\zeta^{(2)}\right)^{-1} \right]_{ij,kl} . \tag{11.93}$$

Renormalization group functions. From the expressions (11.82–11.87) we derive the expansion of the RG functions at two-loop order:

$$\begin{aligned}
\beta_{ijkl} &= -\varepsilon g_{ijkl} + \frac{N_d}{2}\left(g_{ijmn}g_{mnkl} + 2 \text{ terms}\right) + \frac{N_d^2}{4}\left(g_{ijmn}g_{mpqk}g_{npql} + 5 \text{ terms}\right) \\
&\quad + \frac{N_d^2}{48}\left(g_{ijkm}g_{mnpq}g_{npql} + 3 \text{ terms}\right) + O\left(g^4\right), & (11.94) \\
\eta_{ij} &= \frac{N_d^2}{24} g_{iklm}g_{jklm} + O\left(g^3\right), & (11.95) \\
\eta^{(2)}_{ij,kl} &= -\frac{N_d}{2} g_{ijkl} + \frac{N_d^2}{8}\left(g_{ikmn}g_{jlmn} + g_{jkmn}g_{ilmn}\right) \\
&\quad - \frac{N_d^2}{96}\left(g_{imnp}g_{kmnp}\delta_{jl} + 3 \text{ terms}\right) + O\left(g^3\right). & (11.96)
\end{aligned}$$

These expressions will be used in Chapter 26 in their general form. Let us here specialize to the $(\phi^2)^2$ field theory with $O(N)$ symmetry. We then have to substitute

$$g_{ijkl} = \frac{g}{3}\left(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right). \tag{11.97}$$

A short calculation leads to

$$\tilde\beta(\tilde g) = N_d\beta(\tilde g) = -\varepsilon\tilde g + \frac{1}{6}(N+8)\tilde g^2 - \frac{(3N+14)}{12}\tilde g^3 + O\left(\tilde g^4\right), \tag{11.98}$$

$$\eta(\tilde g) = \frac{(N+2)}{72}\tilde g^2 + O\left(\tilde g^3\right), \tag{11.99}$$

in the notation of equation (11.75).

The matrix $\eta^{(2)}_{ij,kl}$ becomes a linear combination of the identity matrix $\mathbf{I}$ and the projector $\mathbf{P}$:

$$I_{ij,kl} = \frac{1}{2}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right), \qquad P_{ij,kl} = \frac{1}{N}\delta_{ij}\delta_{kl} . \tag{11.100}$$

One finds

$$\eta^{(2)} = -\left(N\mathbf{P} + 2\mathbf{I}\right)\frac{\tilde g}{6} + \left[(N+10)\mathbf{I} + 4N\mathbf{P}\right]\frac{\tilde g^2}{72} + O\left(\tilde g^3\right). \tag{11.101}$$

The eigenvalue corresponding to a mass insertion ϕ^2 is given by the eigenvalue $+1$ of $\mathbf{P}$:

$$\eta_1^{(2)} = -\frac{1}{6}(N+2)\tilde{g}\left(1-\frac{5}{12}\tilde{g}\right)+O\left(\tilde{g}^3\right). \tag{11.102}$$

The other eigenvalue related to the eigenvalue 0 of $\mathbf{P}$ corresponds to the insertion of a mass term breaking the $O(N)$ symmetry:

$$\eta_2^{(2)} = -\frac{\tilde{g}}{3}+\frac{(N+10)}{72}\tilde{g}^2+O\left(\tilde{g}^3\right). \tag{11.103}$$

Remark. It is sometimes convenient to write the the flow equation for the coupling constants in terms of the ϕ^4 potential:

$$V(\phi)=\frac{1}{4!}g_{ijkl}\phi_i\phi_j\phi_k\phi_l\,. \tag{11.104}$$

At one-loop, for example, one finds

$$\lambda\frac{\mathrm{d}V}{\mathrm{d}\lambda}=-\varepsilon V+\frac{1}{32\pi^2}\left(\frac{\partial^2 V}{\partial\phi_p\partial\phi_q}\right)^2. \tag{11.105}$$

11.7 One-Loop RG Functions in a Theory with Scalar Bosons and Fermions

We now consider a simple field theory involving one scalar field $\phi(x)$ and N massless Dirac fermion fields $\psi^i(x)$ and $\bar{\psi}^i(x)$, coupled through a Yukawa type $\phi\bar{\psi}\psi$ interaction. The action is

$$\mathcal{S}(\bar{\psi},\psi,\phi)=\int\mathrm{d}^d x\left[-\bar{\psi}\cdot(\not{\partial}+g\phi)\,\psi+\frac{1}{2}\left(\partial_\mu\phi\right)^2+\frac{1}{2}m^2\phi^2+\frac{\lambda}{4!}\phi^4\right]. \tag{11.106}$$

It has an obvious $U(N)$ fermion symmetry and in addition a discrete symmetry that in even dimensions can be identified with a chiral symmetry:

$$\psi\mapsto\gamma_S\psi,\qquad\bar{\psi}\mapsto-\bar{\psi}\gamma_S\,,\qquad\phi(x)\mapsto-\phi(x), \tag{11.107}$$

which forbids a fermion mass term and odd powers of ϕ as ϕ self-interaction. The physics of this model, spontaneous chiral symmetry breaking and fermion mass generation, will be discussed in Section 31.7.

We have already shown that such a field theory is renormalizable in four dimensions by power counting. To calculate renormalization constants we use dimensional regularization and minimal subtraction. We render the renormalized coupling constants dimensionless in d dimensions by setting

$$\lambda_{0i}=m^{\varepsilon}f_i(\lambda), \tag{11.108}$$

in which λ_{0i} represents symbolically the bare coupling constants (g_0^2,λ_0) and λ_i the renormalized coupling constants (g^2,λ).

The eigenvalue corresponding to a mass insertion ϕ^2 is given by the eigenvalue $+1$ of $\mathbf{P}$:

$$\eta_1^{(2)} = -\frac{1}{6}(N+2)\tilde{g}\left(1-\frac{5}{12}\tilde{g}\right)+O\left(\tilde{g}^3\right). \tag{11.102}$$

The other eigenvalue related to the eigenvalue 0 of $\mathbf{P}$ corresponds to the insertion of a mass term breaking the $O(N)$ symmetry:

$$\eta_2^{(2)} = -\frac{\tilde{g}}{3}+\frac{(N+10)}{72}\tilde{g}^2+O\left(\tilde{g}^3\right). \tag{11.103}$$

Remark. It is sometimes convenient to write the the flow equation for the coupling constants in terms of the ϕ^4 potential:

$$V(\phi) = \frac{1}{4!}g_{ijkl}\phi_i\phi_j\phi_k\phi_l\,. \tag{11.104}$$

At one-loop, for example, one finds

$$\lambda\frac{\mathrm{d}V}{\mathrm{d}\lambda} = -\varepsilon V+\frac{1}{32\pi^2}\left(\frac{\partial^2 V}{\partial\phi_p\partial\phi_q}\right)^2. \tag{11.105}$$

11.7 One-Loop RG Functions in a Theory with Scalar Bosons and Fermions

We now consider a simple field theory involving one scalar field $\phi(x)$ and N massless Dirac fermion fields $\psi^i(x)$ and $\bar{\psi}^i(x)$, coupled through a Yukawa type $\phi\bar{\psi}\psi$ interaction. The action is

$$\mathcal{S}(\bar{\psi},\psi,\phi) = \int \mathrm{d}^d x\left[-\bar{\psi}\cdot\left(\not{\partial}+g\phi\right)\psi+\frac{1}{2}\left(\partial_\mu\phi\right)^2+\frac{1}{2}m^2\phi^2+\frac{\lambda}{4!}\phi^4\right]. \tag{11.106}$$

It has an obvious $U(N)$ fermion symmetry and in addition a discrete symmetry that in even dimensions can be identified with a chiral symmetry:

$$\psi\mapsto\gamma_S\psi,\qquad \bar{\psi}\mapsto-\bar{\psi}\gamma_S\,,\qquad \phi(x)\mapsto-\phi(x), \tag{11.107}$$

which forbids a fermion mass term and odd powers of ϕ as ϕ self-interaction. The physics of this model, spontaneous chiral symmetry breaking and fermion mass generation, will be discussed in Section 31.7.

We have already shown that such a field theory is renormalizable in four dimensions by power counting. To calculate renormalization constants we use dimensional regularization and minimal subtraction. We render the renormalized coupling constants dimensionless in d dimensions by setting

$$\lambda_{0i} = m^\varepsilon f_i(\lambda), \tag{11.108}$$

in which λ_{0i} represents symbolically the bare coupling constants (g_0^2,λ_0) and λ_i the renormalized coupling constants (g^2,λ).

We have given in Section 7.7 (equation (8.49)) the 1PI functional $\Gamma(\varphi, \psi, \bar{\psi})$ at one-loop order for an action which contains the action (11.106), $N = 1$, as a special example. Here, we obtain

$$\Gamma_{1\,\text{loop}} = -N \operatorname{tr}\ln\left[1 + g\not{\partial}^{-1}\varphi(x)\right] + \tfrac{1}{2}\operatorname{tr}\ln\Big[1 + \tfrac{1}{2}\lambda\left(-\Delta + m^2\right)^{-1}\varphi^2(x) \\ +2g^2\left(-\Delta + m^2\right)^{-1}\bar{\psi}(x)\cdot\left(\not{\partial} + g\varphi(x)\right)^{-1}\psi(x)\Big]. \tag{11.109}$$

Expanding in powers of φ, ψ and $\bar{\psi}$, we get all 1PI correlation functions at one-loop order. We consider only the divergent functions and omit all contributions involving only the ϕ^4 vertex which have already been evaluated in Section 11.3:

$$\langle\bar{\psi}^1(-p)\psi^1(p)\rangle^{-1} = i\not{p} + \frac{g^2}{(2\pi)^d}\int\frac{\mathrm{d}^d q}{(p-q)^2 + m^2}\frac{i\not{q}}{q^2}, \tag{11.110}$$

$$\langle\bar{\psi}^1(p_1)\psi^1(p_2)\phi(k)\rangle_{1\text{PI}} = g + g^3\int\frac{\mathrm{d}^d q}{(q+p_2)^2 + m^2}\frac{i\not{q}\,i\,(\not{q} - \not{k})}{q^2(q-k)^2}, \tag{11.111}$$

$$\langle\phi(-p)\phi(p)\rangle^{-1} = \langle\phi(-p)\phi(p)\rangle^{-1}\Big|_{g=0} + \frac{Ng^2}{(2\pi)^d}\int\mathrm{d}^d q\frac{\operatorname{tr} i\not{q}\,i\,(\not{q} + \not{p})}{q^2(p+q)^2}, \tag{11.112}$$

$$\langle\phi(p_1)\phi(p_2)\phi(p_3)\phi(p_4)\rangle_{1\text{PI}} = \langle\phi_1\phi_2\phi_3\phi_4\rangle_{1\text{PI}}\big|_{g=0} \\ + N\frac{g^4}{(2\pi)^d}\int\mathrm{d}^d q\frac{\operatorname{tr}\left[i\not{q}\,i\,(\not{q} + \not{p}_1)\,i\,(\not{q} + \not{p}_1 + \not{p}_2)\,i\,(\not{q} - \not{p}_4)\right]}{q^2(q+p_1)^2(q+p_1+p_2)^2(q-p_4)^2} \\ + 5 \text{ diagrams corresponding to permutations of } \{p_2, p_3, p_4\}. \tag{11.113}$$

Figure 11.3 displays the corresponding Feynman diagrams.

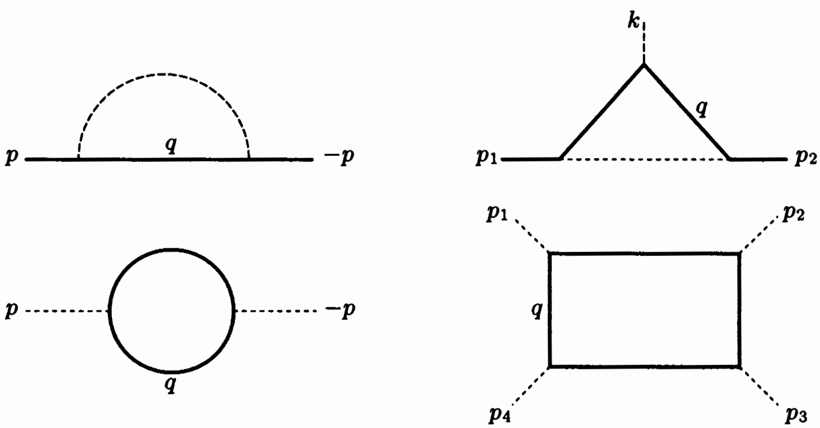


Fig. 11.3 Boson fermion diagrams (the fermions and bosons correspond to continuous and dotted lines, respectively).

11.7.1 Explicit calculations of the two-point functions

As an exercise let us first calculate completely the contributions to the two-point functions.

Fermions. We have to compute the integral $I_1(p)$:

$$I_1(p) = \frac{1}{(2\pi)^d} \int \frac{\mathrm{d}^d q}{(p-q)^2 + m^2} \frac{i\not{q}}{q^2}. \tag{11.114}$$

Rotational invariance implies that $I_1(p)$ is a linear combination of the form $A(p)+iB(p)\not{p}$. To obtain A and B we can take the trace of the integrand multiplied successively by 1 and $\not{p}$. Applying the rules explained in Appendix A8.3 we immediately see that $A = 0$, and thus no fermion mass is generated.

Using the method of Feynman parameters as explained in Section A11.1 we transform (11.114) into

$$I_1(p) = \frac{1}{(2\pi)^d} \int_0^1 \mathrm{d}x \int \frac{\mathrm{d}^d q \, i\not{q}}{\left[x\left((p-q)^2 + m^2\right) + (1-x)q^2\right]^2}. \tag{11.115}$$

We then shift q, $q - xp \mapsto q$, and integrate over q. We find

$$I_1(p) = i\not{p} \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} \int_0^1 \mathrm{d}x \, x^{d/2-1} \left[m^2 + (1-x)p^2\right]^{d/2-2}. \tag{11.116}$$

Expanding at first order in $\varepsilon = 4 - d$ and integrating over x we finally obtain

$$I_1(p) = \frac{i\not{p}}{2} m^{-\varepsilon} \frac{N_d}{\varepsilon} + i\not{p} \frac{1}{32\pi^2} \frac{p^2 + m^2}{p^2} \left[1 - \frac{p^2 + m^2}{p^2} \ln\left(\frac{p^2 + m^2}{m^2}\right)\right]. \tag{11.117}$$

Bosons. We have to calculate the integral $I_2(p)$:

$$I_2(p) = \frac{N}{(2\pi)^d} \int \mathrm{d}^d q \; \mathrm{tr} \, \frac{i\not{q}\,(i\not{q} + i\not{p})}{q^2 (p+q)^2}. \tag{11.118}$$

Evaluating the trace, we get

$$I_2(p) = -\frac{4N}{(2\pi)^d} \int \mathrm{d}^d q \frac{q^2 + p \cdot q}{q^2 (p+q)^2}. \tag{11.119}$$

Then, using

$$q^2 + p \cdot q = \tfrac{1}{2}\left[q^2 + (p+q)^2\right] - \tfrac{1}{2} p^2,$$

we can decompose $I_2(p)$ into a sum of simple integrals:

$$I_2(p) = -\frac{4N}{(2\pi)^d} \int \frac{\mathrm{d}^d q}{q^2} + \frac{4N}{(2\pi)^d} \frac{p^2}{2} \int \frac{\mathrm{d}^d q}{q^2 (p+q)^2}. \tag{11.120}$$

The first integral vanishes in dimensional regularization and the second is proportional to the massless limit of the integral I_4 calculated in Section 11.5. One finds

$$m^{\varepsilon} I_2(p) = 2Np^2 \frac{N_d}{\varepsilon} + \frac{N}{4\pi^2} p^2 \left[\tfrac{1}{2} - \ln(p/m)\right]. \tag{11.121}$$

11.7.2 Other divergences. Renormalization group functions

All one-loop diagrams can be calculated by similar methods. We evaluate here only the divergent part. Taking into account the results of Section 11.3 we find

$$\langle \bar{\psi}\psi\phi \rangle_{\text{1PI,1 loop, div.}} = -g^3 \frac{N_d}{\varepsilon}, \tag{11.122}$$

$$\langle \phi\phi\phi\phi \rangle_{\text{1PI,1 loop, div.}} = \left(-\frac{3}{2}\lambda^2 + 24Ng^4\right)\frac{N_d}{\varepsilon}. \tag{11.123}$$

Renormalization constants. We first replace action (11.106) by the renormalized action:

$$\mathcal{S}_{\text{r}}(\phi,\psi,\bar{\psi}) = \int \mathrm{d}^d x \left\{ -Z_\psi \left[\bar{\psi}(x)\cdot\left(\not{\partial} + g_0 Z_\phi^{1/2}\phi(x)\right)\psi(x)\right] + \frac{1}{2}Z_\phi\left[\left(\partial_\mu\phi(x)\right)^2 + m_0^2\phi^2(x)\right] + \frac{1}{4!}\lambda_0 Z_\phi^2\phi^4(x)\right\}. \tag{11.124}$$

We denote by m_0, g_0, λ_0, the bare parameters, by m, g, λ, the renormalized parameters, and use equation (11.108) to expand the bare coupling constants in terms of g and λ.

Substituting into the Feynman diagrams and identifying the divergent parts we obtain equations for the renormalization constants:

$$\begin{aligned}
i\not{p}\left(Z_\psi + \frac{1}{2}g^2\frac{N_d}{\varepsilon}\right) &= \text{finite} + O(2\text{ loops}),\\
m^{-\varepsilon/2}g_0 Z_\psi Z_\phi^{1/2} - g^3\frac{N_d}{\varepsilon} &= \text{finite} + O(2\text{ loops}),\\
Z_\phi\left(p^2 + m_0^2\right) + \left(-\frac{\lambda}{2}m^2 + 2Np^2g^2\right)\frac{N_d}{\varepsilon} &= \text{finite} + O(2\text{ loops}),\\
m^{-\varepsilon}\lambda_0 Z_\phi^2 + \left(-\frac{3}{2}\lambda^2 + 24Ng^4\right)\frac{N_d}{\varepsilon} &= \text{finite} + O(2\text{ loops}).
\end{aligned}$$

It follows that

$$Z_\psi = 1 - \frac{1}{2}g^2\frac{N_d}{\varepsilon} + O(2\text{ loops}), \tag{11.125}$$

$$Z_\phi = 1 - 2Ng^2\frac{N_d}{\varepsilon} + O(2\text{ loops}), \tag{11.126}$$

$$g_0 = m^{\varepsilon/2}g\left(1 + \frac{(2N+3)}{2}g^2\frac{N_d}{\varepsilon}\right) + O(2\text{ loops}), \tag{11.127}$$

$$\lambda_0 = m^\varepsilon\left[\lambda + \left(\frac{3}{2}\lambda^2 + 4N\lambda g^2 - 24Ng^4\right)\frac{N_d}{\varepsilon}\right] + O(2\text{ loops}), \tag{11.128}$$

$$m_0^2 = m^2\left[1 + \left(\frac{\lambda}{2} + 2Ng^2\right)\frac{N_d}{\varepsilon}\right] + O(2\text{ loops}). \tag{11.129}$$

RG functions. When bare and renormalized coupling constants are related by an equation of the form (11.108), the coupling constant RG functions $\beta_i(\lambda)$ are given by

$$\beta_i(\lambda) = m\frac{\partial}{\partial m}\bigg|_{\lambda_0}\lambda_i\,, \tag{11.130}$$

and thus using equation (11.108),

$$0 = m\,\frac{\partial}{\partial m}\lambda_{0i}\bigg|_{\lambda_0} = m^{\varepsilon}\left(\varepsilon f_i(\lambda) + \beta_j(\lambda)\frac{\partial f_i}{\partial \lambda_j}\right). \tag{11.131}$$

A short calculation then yields

$$\beta_{g^2} = -\varepsilon g^2 + N_d(2N+3)g^4 + \cdots, \tag{11.132a}$$

$$\beta_\lambda = -\varepsilon\lambda + N_d\left(\frac{3}{2}\lambda^2 + 4N\lambda g^2 - 24Ng^4\right) + \cdots. \tag{11.132b}$$

In the minimal subtraction scheme

$$\eta(\lambda) = -\lambda_i\frac{\partial}{\partial\lambda_i}\eta^{(1)}(\lambda),$$

where $\eta^{(1)}(\lambda)$ is the coefficient of $1/\varepsilon$ in the corresponding renormalization constant. Then,

$$\eta_\phi = 2NN_d g^2, \tag{11.133}$$

$$\eta_\psi = \tfrac{1}{2}N_d g^2, \tag{11.134}$$

$$\eta_2 = -\tfrac{1}{2}N_d\lambda - 2NN_d g^2. \tag{11.135}$$

Again, note that the functions η_ψ, η_ϕ are positive for small couplings, a result consistent with the general property $Z_\psi, Z_\phi < 1$ (see Section 6.9).

11.7.3 Pseudoscalar Yukawa interaction

A similar model corresponds to the action

$$S(\phi,\psi,\bar\psi) = \int \mathrm{d}^d x\Big\{-\bar\psi(x)\left[\not\partial + M + ig\gamma_S\phi(x)\right]\psi(x) + \frac{1}{2}\left(\partial_\mu\phi(x)\right)^2 + \frac{1}{2}m^2\phi^2(x) + \frac{1}{4!}\lambda\phi^4(x)\Big\}, \tag{11.136}$$

in which ϕ is a pseudoscalar boson field and $\psi, \bar\psi$ Dirac fermions.

The factor i in front of γ_S is imposed by invariance under reflection hermiticity, that is, euclidean hermitian conjugation (see Chapter 8 and specially Appendix A8 for details) followed by space reflection which implies

$$\gamma_1\gamma_S^\dagger\left(ig\phi(x)\right)^*\gamma_1 = \gamma_S ig\phi(x).$$

In addition, the action is reflection symmetric:

$$\psi(x)\mapsto\gamma_1\psi(\tilde x),\quad \bar\psi(x)\mapsto\bar\psi(\tilde x)\gamma_1,\quad \phi(x)\mapsto-\phi(\tilde x)$$

with $\tilde x = (-x_1, x_2, \ldots, x_d)$.

The transformation of $\phi(x)$ is imposed by the factor γ_S in front of the $\phi\bar\psi\psi$ interaction and justifies the denomination pseudoscalar for the field $\phi(x)$.

The renormalized action takes the form

$$S_{\mathrm{r}}(\phi,\psi,\bar\psi)=\int \mathrm{d}^d x\Big\{-Z_\psi\left[\bar\psi(x)\left(\not\partial+M_0+i\gamma_S g_0 Z_\phi^{1/2}\phi(x)\right)\psi(x)\right] \\ +\frac{1}{2}Z_\phi\left[\left(\partial_\mu\phi(x)\right)^2+m_0^2\phi^2(x)\right]+\frac{1}{4!}\lambda_0 Z_\phi^2\phi^4(x)\Big\}. \qquad (11.137)$$

Calculations of RG functions can be done using dimensional regularization and minimal subtraction because it is never necessary to evaluate any trace of products of γ_S by an even number of γ_μ matrices.

For what concerns the coupling and field renormalizations no new calculation is required because the models (11.106) and (11.136) are related by performing in action (11.136), restricted to $M=0$, a chiral rotation on the fields $\psi,\bar\psi$. For $M=0$, one sets

$$\psi=\mathrm{e}^{-i\pi\gamma_S/4}\,\psi'\,,\qquad \bar\psi=\bar\psi'\,\mathrm{e}^{-i\pi\gamma_S/4}\,.$$

Then the action becomes identical to the action of the model (11.106). This explains that the renormalization constants are the same.

The mass renormalization is determined by the $\bar\psi\psi$ two-point function. At one-loop order one can expand the function on the basis $\{1,\not p\}$ and use identities of the form

$$p\cdot q=\tfrac{1}{2}\left[q^2+M^2-(p-q)^2-m^2\right]+\tfrac{1}{2}\left(p^2+m^2-M^2\right).$$

One finds in the notation (11.117),

$$I_1(p)=\left[\frac{N_d}{\varepsilon}-\frac{1}{16\pi^2}\ln\left(\frac{Mm}{\mu^2}\right)\right]\left(M+\frac{i\not p}{2}\right)-\frac{1}{32\pi^2}\frac{i\not p}{p^2}\left(M^2+m^2\right)\ln\left(\frac{M}{m}\right) \\ -\frac{1}{16\pi^2}\left[M+\frac{i\not p}{2p^2}\left(p^2+m^2-M^2\right)\right]\left[\frac{1}{p^2}\sigma_+\sigma_-\ln\left(\frac{\sigma_++\sigma_-}{\sigma_+-\sigma_-}\right)\right. \\ \left.-1-\ln(M/m)\left(\frac{M^2-m^2}{p^2}\right)\right]+O(\varepsilon)\,, \qquad (11.138)$$

where we have set

$$\sigma_\pm(p^2)=\sqrt{p^2+(M\pm m)^2}. \qquad (11.139)$$

The divergent part of the ϕ two-point function at one-loop order is

$$I_2(p)=4\left(M^2+\frac{p^2}{2}\right)\frac{N_d}{\varepsilon}+O(1)\,. \qquad (11.140)$$

Bibliographical Notes

Topological properties and parametric representations of Feynman diagrams are studied in

N. Nakanishi, *Graph Theory and Feynman Integrals* (Gordon and Breach, New York 1971).

Feynman diagrams and dimensional regularization are discussed in

G. 't Hooft and M. Veltman, in *Particle Interactions at Very High Energies*, D. Speiser, F. Halzen and J. Weyers eds. (Plenum Press, New York 1973) vol. 4B.

The properties of RG functions in the minimal subtraction scheme is discussed in

G. 't Hooft, *Nucl. Phys.* B61 (1973) 455.

A good presentation can be found in

D.J. Gross in *Methods in Field Theory*, Les Houches 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

For more systematic techniques to calculate RG functions at higher orders see

A.A. Vladimirov, *Teor. Mat. Fiz.* 43 (1980) 210.

APPENDIX A11

A11.1 Feynman Parameters

Feynman parameters. In explicit calculations of Feynman diagrams a simple identity is often useful. One starts from

$$\frac{1}{a^\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty \mathrm{d}t\, t^{\alpha-1}\,\mathrm{e}^{-at}\,. \qquad (A11.1)$$

Therefore,

$$\prod_{i=1}^n (a_i)^{-\alpha_i} = \prod_{i=1}^n (\Gamma(\alpha_i))^{-1} \int_0^\infty \left(\prod_{i=1}^n \mathrm{d}t\, t_i^{\alpha_i-1}\right) \exp\left(-\sum_{i=1}^n a_i t_i\right). \qquad (A11.2)$$

Then, setting

$$t_i = s u_i \qquad (A11.3)$$

with

$$u_i \ge 0\,, \qquad \sum_{i=1}^n u_i = 1\,,$$

we can integrate over s to obtain

$$\prod_{i=1}^n (a_i)^{-\alpha_i} = \Gamma(\alpha_1+\alpha_2+\cdots+\alpha_n) \prod_{i=1}^n (\Gamma(\alpha_i))^{-1}$$
$$\times \int_0^\infty \delta\left(\sum_i u_i - 1\right) \prod_{i=1}^n \mathrm{d}u_i u_i^{\alpha_i-1} \left(\sum_{i=1}^n a_i u_i\right)^{-\sum_i \alpha_i}. \qquad (A11.4)$$

If the quantities $a_1,\ldots,a_n$ 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:

$$\frac{1}{(2\pi)^d} \int \frac{\mathrm{d}^d p}{(p^2+1)^{n+1}} = \frac{1}{(2\pi)^d \Gamma(n+1)} \int_0^\infty t^n \mathrm{d}t \int \mathrm{d}^d p\, \mathrm{e}^{-tp^2-t} = \frac{\Gamma(n+1-d/2)}{(4\pi)^{d/2}\Gamma(n+1)}\,. \qquad (A11.5)$$

Note that representation (9.65) leads to an expression similar to (A11.4). 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 (A11.3), the integral over the homogeneous variable s can be performed.

In the calculations we have also used the integral

$$B(\alpha,\beta) \equiv \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \qquad (A11.6)$$

12 RENORMALIZATION OF COMPOSITE OPERATORS. SHORT DISTANCE EXPANSION

In this chapter, we discuss two related problems: the renormalization of composite operators, that is, local polynomials of the field, and the short distance expansion (SDE) of the product of operators. The relation is easy to understand: we 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 $x \to y$ 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 composite operators in general from the point of view of power counting. We use the relations between bare and renormalized operators to establish a set of CS equations for the insertion of operators of dimension 4 in the ϕ^4_4 field theory. We then show that in a given field theory there exist linear relations between operators due to the equations of motion and relations derived in Section 7.3.

In the second part of this chapter, devoted to the SDE, we first establish the existence of a SDE for the product of two basic fields, and discuss the SDE at leading order in the ϕ^4_4 field theory.

We then pass from short distance behaviour to large momentum behaviour and derive a CS equation for the coefficient of the expansion at leading order. We, finally, 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 which appears when one studies the large momentum behaviour of real time correlation functions (in contrast to euclidean imaginary time).

12.1 Renormalization of Operator Insertions

We have shown in Section 10.6 how to renormalize insertions of the operator ϕ^2. We could have considered other vertices like ϕ^4, $(\partial_\mu\phi)^2$... They all generate new divergences which have to be eliminated by additional renormalizations.

In Section 9.4 we explained how to calculate the superficial degree of divergence of the insertion of a local operator $\mathcal{O}(\phi,x)$ by adding a source $g(x)$ for this operator in the action:

$$\mathcal{S}(\phi) \mapsto \mathcal{S}(\phi) + \int \mathrm{d}x\, \mathcal{O}(\phi,x)g(x).$$

With this choice of sign, differentiation of $\mathcal{W}(J,g)$ with respect to $g(x)$ generates insertions of $-\mathcal{O}(\phi,x)$ in connected correlation functions. However, since $\delta\Gamma/\delta g(x) = -\delta\mathcal{W}/\delta g(x)$, $\delta\Gamma/\delta g(x)$ corresponds to the insertion of $\mathcal{O}(\phi,x)$ in proper vertices.

As a convention, we assign a canonical dimension $[g(x)]$ to the source $g(x)$, 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

$$[g] = d - [\mathcal{O}(\phi)].$$

With this convention, the consequences of power counting and renormalization theory have a simple formulation: the sum of counter-terms needed to render $\Gamma(\phi,g)$ finite is the most general local functional of $\phi(x)$ and $g(x)$ allowed by power counting, that is, ***the most general linear combination of all vertices in $\phi(x)$ and $g(x)$ of non-positive dimensions.*** Note that this in particular implies that an operator of a given dimension inserted one time in general mixes under renormalization with all operators of equal or lower dimension. It is thus natural to study the renormalization of operators of increasing dimension.

Let us, therefore, first verify this result in the case of the insertion of $\phi^2(x)$ in a ϕ^4 field theory in four dimensions which we have already discussed in Chapter 10.

Regularization. For simplicity we assume in this chapter a momentum cut-off regularization and denote the cut-off by Λ.

12.1.1 The $\phi^2(x)$ insertion

We use the conventions of Chapter 10 for the bare and renormalized actions:

$$\mathcal{S}(\phi) = \int \mathrm{d}^4 x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} m^2 \phi^2 + \frac{1}{4!} g \phi^4\right], \tag{12.1}$$

$$\mathcal{S}_\mathrm{r}(\phi_\mathrm{r}) = \int \mathrm{d}^4 x \left[\frac{1}{2} Z \left(\partial_\mu \phi_\mathrm{r}\right)^2 + \frac{1}{2}\left(m_\mathrm{r}^2 + \delta m^2\right) \phi_\mathrm{r}^2 + \frac{1}{4!} g_\mathrm{r} Z_g \phi_\mathrm{r}^4\right]. \tag{12.2}$$

In the ϕ^4 field theory in four dimensions the operator $\phi^2(x)$ has canonical dimension 2:

$$[\phi^2] = 2\,.$$

We denote by $t(x)$ the source for $\phi^2(x)$. Its dimension $[t]$ is

$$[t(x)] = 2\,.$$

Then, in addition to the vertices involving only the field ϕ, the following vertices arise as counter-terms:

$$\int t(x)\phi^2(x)\mathrm{d}^4 x \quad \text{which has dimension zero,}$$

$$\int t^2(x)\mathrm{d}^4 x \quad \text{which also has dimension zero,}$$

$$\int t(x)\mathrm{d}^4 x \quad \text{which has dimension } -2\,.$$

The renormalized action then has the form

$$\mathcal{S}_\mathrm{r}(\phi_\mathrm{r}, t) = \mathcal{S}_\mathrm{r}(\phi_\mathrm{r}) + \tfrac{1}{2} Z_2 \int \mathrm{d}^4 x\, t(x) \phi_\mathrm{r}^2(x) + \int \mathrm{d}^4 x \left[\tfrac{1}{2} a t^2(x) + b t(x)\right]. \tag{12.3}$$

The last two terms contribute only to the vacuum amplitude. Expression (12.3) implies a set of relations between bare and renormalized generating functionals:

$$\mathcal{Z}_\mathrm{r}(J, t) = \mathcal{Z}(J/\sqrt{Z}, t Z_2/Z) \exp\left[-\int \mathrm{d}^4 x \left(\tfrac{1}{2} a t(x)^2 + b t(x)\right)\right]. \tag{12.4}$$

For the connected functions this gives

$$\mathcal{W}_\mathrm{r}(J, t) = \mathcal{W}(J/\sqrt{Z}, t Z_2/Z) - \int \mathrm{d}^4 x \left(\tfrac{1}{2} a t(x)^2 + b t(x)\right). \tag{12.5}$$

After Legendre transformation with respect to J, one obtains

$$\Gamma_\mathrm{r}(\varphi, t) = \Gamma(\varphi\sqrt{Z}, t Z_2/Z) + \int \mathrm{d}^4 x \left(\tfrac{1}{2} a t(x)^2 + b t(x)\right). \tag{12.6}$$

Expanding in powers of t and φ, one recovers the relations between bare and renormalized proper vertices proved in Chapter 10.

12.1.2 Operators of dimensions 3 and 4

The ϕ^3 insertion. Let us first consider the insertion of the ϕ^3 operator. The operator $\phi^3(x)$ has dimension 3, the corresponding source $t(x)$ has dimension 1. The renormalized action $\mathcal{S}_\mathrm{r}(\phi, t)$ thus has the form

$$\mathcal{S}_\mathrm{r}(\phi,t) = \mathcal{S}_\mathrm{r}(\phi) + \int \mathrm{d}^4x \left[t(x) \left(\frac{1}{3!} Z_3 \phi^3(x) + a\nabla^2\phi(x) + b\phi(x) \right) \right.$$
$$\left. + c\left(\partial_\mu t(x)\right)^2 + t^2(x)\left(d\phi^2(x) + e\right) + f t^3(x)\phi(x) + g t^4(x) \right]. \qquad (12.7)$$

This expression, in particular, shows that the operator $\phi^3(x)$ mixes under renormalization with $\phi(x)$, that the double insertion of ϕ^3 generates a counter-term proportional to ϕ^2 .... To be able to write the equivalent of relations (12.4–12.6) we have, therefore, to explicitly introduce a source for $\phi^2(x)$. We leave it to the reader to write the renormalized action with sources for ϕ^3 and ϕ^2 and the relations between renormalized and bare proper vertices. We postpone the discussion of the CS equations of correlation functions with ϕ^3 insertion until Section 12.2 in order to be able to incorporate information coming from the field equation of motion.

Operators of dimension 4. Quite generally, if an operator has a dimension strictly smaller than the space dimension d, the source has a strictly positive dimension and the renormalized action is a polynomial in the source. If the source is coupled to operators of dimension d, corresponding to vertices of dimension zero (here $\phi^4(x)$, $(\partial_\mu\phi(x))^2$), an infinite series in the source is generated by the renormalization procedure, together with all operators of lower or equal dimensions.

In the ϕ^4_4 field theory, if one inserts once an operator of dimension 4, 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

$$\begin{aligned} \mathcal{O}_1(\phi) &= \tfrac{1}{2} m_\mathrm{r}^2 \phi^2(x), & \mathcal{O}_2(\phi) &= -\tfrac{1}{2}\nabla^2\left(\phi^2(x)\right), \\ \mathcal{O}_3(\phi) &= \tfrac{1}{2}\left[\partial_\mu\phi(x)\right]^2, & \mathcal{O}_4(\phi) &= \tfrac{1}{4!}\phi^4(x), \end{aligned} \qquad (12.8)$$

form a basis of linearly independent operators mixing 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 $\mathcal{O}_3$:

$$\tfrac{1}{2}\nabla^2\left(\phi^2(x)\right) = \phi(x)\nabla^2\phi(x) + \left[\partial_\mu\phi(x)\right]^2.$$

Operator $\mathcal{O}_1(\phi)$, and, therefore, also operator $\mathcal{O}_2(\phi)$, are multiplicatively renormalizable.

We can thus write the relation between bare and renormalized correlation functions $\Gamma^{(n)}_{\mathcal{O}_i}$ with $\mathcal{O}_i$ insertion in the form

$$\left(\Gamma^{(n)}_{\mathcal{O}_i}\right)_\mathrm{r} = Z^{n/2} \sum_j Z_{ij} \Gamma^{(n)}_{\mathcal{O}_j}. \qquad (12.9)$$

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 10.2.

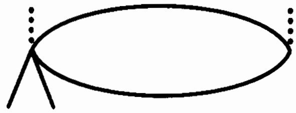


Fig. 12.1 Divergent contribution to $\phi^2\phi^4$ insertion.

CS equations. From equation (12.9) we can derive a CS equation for $\left(\Gamma^{(n)}_{\mathcal{O}_i}\right)_{\mathrm{r}}$. However, here some care is required. The CS operation involves ϕ^2 insertions and the product, for example, $(\phi^4 x)_{\mathrm{r}}(\phi^2(y))_{\mathrm{r}}$ inserted in a correlation function is not finite: since the source for ϕ^4 has dimension zero and the source for ϕ^2 dimension 2 the product of ϕ^2 by both sources has dimension 4. Figure 12.1 gives the first divergent diagram.

This implies:

$$\tfrac{1}{4!}\left[\left(\phi^4(x)\right)_{\mathrm{r}}\left(\phi^2(y)\right)_{\mathrm{r}}\right]_{\mathrm{r}}=\tfrac{1}{4!}\left(\phi^4(x)\right)_{\mathrm{r}}\left(\phi^2(y)\right)_{\mathrm{r}}+C_4\delta(x-y)\left(\phi^2(x)\right)_{\mathrm{r}}, \tag{12.10}$$

in which C_4 is a new renormalization constant. Identity (12.10) is only true as an insertion in an n-point correlation function, $n\neq 0$.

After Fourier transformation and for an insertion of ϕ^2 at zero momentum, equation (12.10) becomes

$$\left[\left(\tilde{\mathcal{O}}_4(p)\right)_{\mathrm{r}}\left(\tilde{\mathcal{O}}_1(0)\right)_{\mathrm{r}}\right]_{\mathrm{r}}=\left(\tilde{\mathcal{O}}_4(p)\right)_{\mathrm{r}}\left(\tilde{\mathcal{O}}_1(0)\right)_{\mathrm{r}}+C_4\left(\tilde{\mathcal{O}}_1(p)\right)_{\mathrm{r}}. \tag{12.11}$$

A similar equation holds for the operator $\mathcal{O}_3$.

We now apply the CS operator, $m_{\mathrm{r}}\partial/\partial m_{\mathrm{r}}$, at g and Λ fixed, on equation (12.9).

A new set of RG functions is generated involving the matrix $\tilde{\eta}_{ij}$:

$$\tilde{\eta}_{ij}(g_{\mathrm{r}},\Lambda/m_{\mathrm{r}})=\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}}Z_{ik}\right)Z^{-1}_{kj}. \tag{12.12}$$

As a consequence of relation (12.11), two elements of the matrix $\tilde{\eta}_{ij}$ are not finite when the cut-off becomes infinite: $\tilde{\eta}_{31}$ and $\tilde{\eta}_{41}$. Their divergent part cancels the divergences coming from the insertion of ϕ^2 as represented by equation (12.11). Defining, then,

$$\eta_{i1}(g_{\mathrm{r}})=\tilde{\eta}_{i1}-m_{\mathrm{r}}^2\sigma(g_{\mathrm{r}})C_i\,, \tag{12.13}$$

we now obtain two finite RG functions.

For the other matrix elements we just set

$$\eta_{ij}=\tilde{\eta}_{ij}\,. \tag{12.14}$$

The CS equations then read

$$\left\{\left[m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}}+\beta(g_{\mathrm{r}})\frac{\partial}{\partial g_{\mathrm{r}}}-\frac{n}{2}\eta(g_{\mathrm{r}})\right]\delta_{ij}-\eta_{ij}(g_{\mathrm{r}})\right\}\left(\Gamma^{(n)}_{\mathcal{O}_j}\right)_{\mathrm{r}}(p_k)=m_{\mathrm{r}}^2\,\sigma(g_{\mathrm{r}})\left(\Gamma^{(1,n)}_{\mathcal{O}_i}\right)_{\mathrm{r}}(0;p_k). \tag{12.15}$$

The matrix η_{ij} has the form

$$\eta_{ij}=\begin{bmatrix}\mathbf{c} & 0\\ \mathbf{b} & \mathbf{a}\end{bmatrix},$$

in which **a**, **b** and **c** are 2×2 matrices, and **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 $(\partial_\mu \phi(x))^2$ insertion. It is similar to the $\phi^4\phi^2$ insertion. The relation between product of renormalized operators and renormalized product now is

$$\begin{aligned}\left[\left(\phi^4(x)\right)_{\mathrm{r}}\left(\phi^4(y)\right)_{\mathrm{r}}\right]_{\mathrm{r}} &= \left(\phi^4(x)\right)_{\mathrm{r}}\left(\phi^4(y)\right)_{\mathrm{r}} + \delta(x-y)\sum_{i=1}^{4} D_{4i}\left[\mathcal{O}_i\left(\phi(x)\right)\right]_{\mathrm{r}} \\ &\quad + \partial_\mu\delta(x-y)E_4\partial_\mu\mathcal{O}_1\left(\phi(x)\right)_{\mathrm{r}} + \partial^2\delta\,(x-y)F_4\mathcal{O}_1\left(\phi(x)\right)_{\mathrm{r}}\,, \quad (12.16)\end{aligned}$$

in which D_{4i}, E_4 and F_4 are new renormalization constants. A similar equation is valid for $[\partial_\mu\phi(x)]^2$. Again equation (12.16) is valid only as an insertion.

12.1.3 Operator insertion: general case

Power counting arguments, based on the dimension of operators and sources, tell us quite generally that if $\mathcal{O}(\phi, x)$ is an operator of canonical dimension D, $[\mathcal{O}(\phi)] = D$, then it renormalizes as

$$\left[\mathcal{O}(\phi, x)\right]_{\mathrm{r}} = \sum_{\alpha:\ [\mathcal{O}_\alpha]\le D} Z_\alpha \mathcal{O}_\alpha(\phi, x). \qquad (12.17)$$

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,

$$\begin{aligned}\left\{\left[\mathcal{O}(\phi,x)\right]_{\mathrm{r}}\left[\mathcal{O}'(\phi,y)\right]_{\mathrm{r}}\right\}_{\mathrm{r}} &= \left[\mathcal{O}(\phi,x)\right]_{\mathrm{r}}\left[\mathcal{O}'(\phi,y)\right]_{\mathrm{r}} \\ &+ \sum_{\alpha:[\mathcal{O}_\alpha]+[P_\alpha]\le D+D'-d} C_\alpha \mathcal{O}_\alpha(\phi,x)P_\alpha(\partial_\mu)\delta(x-y), \qquad (12.18)\end{aligned}$$

in which $P_\alpha(\partial_\mu)$ is a polynomial in ∂_μ. For example in the ϕ^4_4 field theory, the product $(\phi^6(x))_{\mathrm{r}} t(\phi^8(y))_{\mathrm{r}}$ involves all operators of dimension lower than or equal to 10.

12.2 Quantum Field Equations

We have discussed the renormalization of composite operators. However, not all renormalizations are independent in a given field theory, because quantum field equations and other identities discussed in Section 7.3 imply relations between operators.

12.2.1 Insertion of the ϕ^3 operator

Let us again discuss the example of the $\phi^3(x)$ operator in the framework of the ϕ^4 field theory.

We, therefore, consider the action $\mathcal{S}(\phi, t, u)$:

$$\mathcal{S}(\phi,t,u) = \int \mathrm{d}^4x \left[\tfrac{1}{2}\left(\partial_\mu\phi\right)^2 + \tfrac{1}{2}m^2\phi^2 + \frac{1}{4!}g\phi^4 + \tfrac{1}{2}t(x)\phi^2 + \frac{1}{3!}u(x)\phi^3\right]. \qquad (12.19)$$

(A regularization is assumed.) Differentiation with respect to $u(x)$ and $t(x)$ generates $-\frac{1}{3!}\phi^3$ and $-\frac{1}{2}\phi^2$ insertions.

The simplest relations are derived from the quantum field equation

$$\int [\mathrm{d}\phi]\left[\frac{\delta \mathcal{S}}{\delta\phi(x)} - J(x)\right]\exp\left(-\mathcal{S}+J\phi\right)=0\,.$$

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 $u(x)$. We then obtain

$$\left[\left(-\nabla^2+m^2\right)_\Lambda + t(x)\right]\frac{\delta \mathcal{Z}}{\delta J(x)} - g\frac{\delta \mathcal{Z}}{\delta u(x)} - u(x)\frac{\delta \mathcal{Z}}{\delta t(x)} = J(x)\mathcal{Z}(J,u,t). \tag{12.20}$$

This yields an equation for $\mathcal{W}=\ln\mathcal{Z}$:

$$\left[\left(-\nabla^2+m^2\right)_\Lambda + t(x)\right]\frac{\delta \mathcal{W}}{\delta J(x)} - g\frac{\delta \mathcal{W}}{\delta u(x)} - u(x)\frac{\delta \mathcal{W}}{\delta t(x)} = J(x). \tag{12.21}$$

The Legendre transformation is straightforward. We remember that

$$\frac{\delta\mathcal{W}}{\delta u(x)} = -\frac{\delta\Gamma}{\delta u(x)},\qquad \frac{\delta\mathcal{W}}{\delta t(x)} = -\frac{\delta\Gamma}{\delta t(x)},$$

and find

$$\left[\left(-\nabla^2+m^2\right)_\Lambda + t(x)\right]\varphi(x) + g\frac{\delta\Gamma}{\delta u(x)} + u(x)\frac{\delta\Gamma}{\delta t(x)} - \frac{\delta\Gamma}{\delta\varphi(x)} = 0\,. \tag{12.22}$$

To explore a first consequence of this relation, let us set $t(x)=u(x)=0$:

$$\left.\frac{\delta\Gamma}{\delta u(x)}\right|_{u=t=0} = \Gamma_{\phi^3}\,, \tag{12.23}$$

Γ_{ϕ^3} is the generating functional of 1PI ϕ-field correlation functions with one $\frac{1}{3!}\phi^3$ insertion.

$$g\Gamma_{\phi^3(x)} = \frac{\delta\Gamma}{\delta\varphi(x)} - \left(-\nabla^2+m^2\right)_\Lambda\varphi(x). \tag{12.24}$$

This relation shows that, up to explicit subtractions affecting only the $\langle\phi^3\phi\rangle$ correlation function, the insertion of ϕ^3 is equivalent to the insertion of ϕ itself.

The diagrammatic interpretation of equation (12.24) 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 (see figure 12.2).

Fig. 12.2 ϕ^3 insertion.

However, diagrams without a ϕ^4 vertex cannot be generated and this explains the subtractions.

We now introduce the generating functional of renormalized proper vertices

$$\Gamma_{\mathrm{r}}(\varphi) = \Gamma(\varphi\sqrt{Z}),$$

and, thus,

$$\frac{\delta\Gamma_{\mathrm{r}}(\varphi)}{\delta\varphi(x)} = \sqrt{Z}\,\frac{\delta\Gamma(\varphi\sqrt{Z})}{\delta\varphi(x)}.$$

We use this relation in equation (12.24):

$$\left(-\nabla^2 + m^2\right)_{\Lambda} Z\varphi(x) + g\sqrt{Z}\,\Gamma_{\phi^3(x)}(\varphi\sqrt{Z}) = \frac{\delta\Gamma_{\mathrm{r}}(\varphi)}{\delta\varphi(x)}. \tag{12.25}$$

The r.h.s. is finite in the infinite cut-off limit. We now introduce the ϕ^4 field theory renormalization constants as defined by equations (12.1,12.2):

$$\left[-\nabla^2 + Z^{-2}\left(m_{\mathrm{r}}^2 + \delta m^2\right)\right] Z\varphi(x) + g_{\mathrm{r}} Z_g Z^{-3/2}\Gamma_{\phi^3}(\varphi\sqrt{Z}) = \frac{\delta\Gamma_{\mathrm{r}}(\varphi)}{\delta\varphi(x)}. \tag{12.26}$$

This relation shows that all ϕ-field correlation functions with one insertion of the operator $g_{\mathrm{r}} Z_g \phi_{\mathrm{r}}^3(x)$ are finite except the $\langle\phi^3\phi\rangle$ correlation function which needs two additional subtractions. We determine the corresponding renormalization constants by imposing

$$\begin{aligned} \left(\Gamma^{(1)}_{\phi^3}\right)_{\mathrm{r}}(p,-p)\Big|_{p=0} &= 0\,, \\ \frac{\partial}{\partial p^2}\left(\Gamma^{(1)}_{\phi^3}\right)_{\mathrm{r}}(p,-p)\Big|_{p=0} &= 0\,. \end{aligned} \tag{12.27}$$

Equation (12.26) expanded in powers of φ then yields explicitly,

$$\Gamma_{\mathrm{r}}^{(n+1)}(q;p_1,\ldots,p_n) = g_{\mathrm{r}}\left(\Gamma^{(n)}_{\phi^3}\right)_{\mathrm{r}}(q;p_1,\ldots,p_n) + \delta_{n1}\left(p^2+m^2\right). \tag{12.28}$$

Note that with this definition $(\Gamma^{(3)}_{\phi^3})_{\mathrm{r}}$ satisfies the renormalization condition

$$\left(\Gamma^{(3)}_{\phi^3}\right)_{\mathrm{r}}(0;0,0,0) = 1\,. \tag{12.29}$$

Equation (12.22) also contains information about multiple insertions of ϕ^3. For example, after some algebraic manipulations one finds

$$g^2\Gamma_{\phi^3(x_1)\phi^3(x_2)} + g\delta(x_1-x_2)\Gamma_{\phi^2(x_1)} + \left(-\nabla^2+m^2\right)\delta(x_1-x_2) = \frac{\delta^2\Gamma}{\delta\varphi(x_1)\delta\varphi(x_2)}. \tag{12.30}$$

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.

12.2.2 Other relations: renormalization of operators of dimension 4

We have shown in Section 7.3 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 variables $\phi \mapsto \phi'$,

$$\phi'(x) = \phi(x) + \varepsilon(x)\phi(x),$$

the variation of the action (12.1) in the presence of a source is

$$\delta\left[\mathcal{S}(\phi) - J\phi\right] = \varepsilon(x)\left[\phi(x)\left(-\nabla^2 + m^2\right)_\Lambda \phi(x) + \frac{g}{3!}\phi^4(x) - J(x)\phi(x)\right].$$

From the invariance of the functional integral it follows that

$$\mathcal{W}_{\phi(x)(-\nabla^2+m^2)_\Lambda\phi(x)} + \frac{g}{3!}\mathcal{W}_{\phi^4(x)} = J(x)\frac{\delta \mathcal{W}}{\delta J(x)}. \tag{12.31}$$

The discussion of the large cut-off limit of this equation, or the corresponding one obtained after Legendre transformation,

$$\Gamma_{\phi(x)(-\nabla^2+m^2)_\Lambda\phi(x)} + \frac{g}{3!}\Gamma_{\phi^4(x)} = \varphi(x)\frac{\delta\Gamma}{\delta\varphi(x)}, \tag{12.32}$$

is more delicate than in the case of equation (12.24): the operator $\phi(x)\left(-\nabla^2 + m^2\right)_\Lambda \phi(x)$ which in Pauli–Villars regularization is

$$\phi(x)\left(-\nabla^2 + m^2\right)_\Lambda \phi(x) \equiv \phi(x)\left[-\nabla^2 + m^2 + \alpha_1\frac{\nabla^4}{\Lambda^2} - \alpha_2\frac{\nabla^6}{\Lambda^3} + \cdots\right]\phi(x),$$

contains operators of canonical dimensions larger than 4 divided by powers of the cut-off. We shall discuss in Chapter 27 the problem of irrelevant operators which is directly related with the large cut-off limit of such operators. Let us here only state the result: in the large cut-off limit the operator $\phi\left(-\nabla^2 + m^2\right)_\Lambda \phi$ is equivalent to a linear combination of all operators of dimensions 4 and 2.

Equation (12.32) implies after renormalization an identity satisfied by the operators $\left(\mathcal{O}_i(\phi)\right)_\mathrm{r}$ as defined in equations (12.8,12.9):

$$\sum_{i=1}^{4} C_i(g_\mathrm{r})\left(\Gamma^{(n)}_{\mathcal{O}_i}\right)_\mathrm{r}(q;p_1,\ldots,p_n) = \sum_{m=1}^{n}\Gamma^{(n)}(p_1,\ldots,p_m+q,\ldots,p_n). \tag{12.33}$$

If we impose some renormalization conditions to define explicitly the insertions of operators $\mathcal{O}_i$, in the spirit of Chapter 10,

$$\begin{aligned}
\left(\Gamma^{(2)}_{\mathcal{O}_1}\right)_\mathrm{r}(0;0,0) &= m_\mathrm{r}^2\,,\\
\left(\Gamma^{(2)}_{\mathcal{O}_3}\right)_\mathrm{r}(q;p_1,p_2) &= -p_1\cdot p_2 + O\left(p^4\right),\\
\left(\Gamma^{(4)}_{\mathcal{O}_3}\right)_\mathrm{r}(0;0,0,0,0) &= 0\,,\\
\left(\Gamma^{(2)}_{\mathcal{O}_4}\right)_\mathrm{r}(q;p_1,p_2) &= O\left(p^4\right),\\
\left(\Gamma^{(4)}_{\mathcal{O}_4}\right)_\mathrm{r}(0;0,0,0,0) &= 1\,,
\end{aligned} \tag{12.34}$$

then we can calculate the coefficients $C_i(g_\mathrm{r})$ only from $\left(\Gamma^{(n)}_{\mathcal{O}_1}\right)_\mathrm{r}$ and its derivatives at zero momentum.

12.3 Short Distance Expansion (SDE) of Operator Products

Several chapters will be devoted to the discussion of the large or small momentum behaviour of correlation functions. Our essential tool will be the CS or RG equations. However, these equations are directly useful only for non-exceptional momenta. For large momenta we shall be able to characterize the behaviour of

$$\Gamma^{(n)}(\rho p_1 + r_1, \rho p_2 + r_2, \ldots, \rho p_n + r_n), \qquad \rho \to \infty,$$

provided no subset of momenta p_i has a vanishing sum:

$$\sum_{i \in I \neq \emptyset} p_i \neq 0 \quad \forall I \not\equiv (1, 2, \ldots n).$$

For exceptional momenta a new tool is needed: the short distance expansion (SDE) of product of operators. In this section, we shall mainly discuss the SDE of the product of two fields at leading order, but we shall indicate how the method we use can be generalized to more complicated cases. It is entirely based on the theory of renormalization of composite operators we have just presented.

Definition. We consider the 1PI correlation function:

$$\Gamma^{(n+2)}\left(x + \frac{v}{2}, x - \frac{v}{2}, y_1, \ldots, y_n\right) = \left\langle \phi\left(x + \frac{v}{2}\right)\phi\left(x - \frac{v}{2}\right)\phi(y_1)\cdots\phi(y_n)\right\rangle_{1\,\mathrm{PI}}, \tag{12.35}$$

in which all arguments are fixed, except the vector $\mathbf{v}$ which goes to zero. We want to study the $\mathbf{v} \to \mathbf{0}$ limit. In a theory which is sufficiently regularized, an expansion in powers of v can be obtained by expanding the product of fields $\phi(x + v/2)\phi(x - v/2)$ in powers of v:

$$\begin{aligned}\phi\left(x + v/2\right)\phi\left(x - v/2\right) = \phi^2(x) + \frac{1}{4}v_{\mu_1}v_{\mu_2}\left[\phi(x)\partial_{\mu_1}\partial_{\mu_2}\phi(x) - \partial_{\mu_1}\phi(x)\partial_{\mu_2}\phi(x)\right] \\ + \sum_{n=2}^{\infty}\frac{1}{2^{2n}}\frac{1}{2n!}v_{\mu_1}\ldots v_{\mu_{2n}}\mathcal{O}_{\mu_1\ldots\mu_{2n}}\left(\phi(x)\right),\end{aligned} \tag{12.36}$$

in which $\mathcal{O}_{\mu_1\ldots\mu_{2n}}(\phi)$ is a local operator quadratic in ϕ with $2n$ derivatives.

If we now insert expansion (12.36) 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 12.1, the various composite operators appearing in expansion (12.36) have to be renormalized. Therefore, we have to expand each bare operator on a basis of renormalized operators. In terms of renormalized operators $\mathcal{O}^{\alpha}_{\mathrm{r}}(\phi)$ expansion (12.36) then takes the form

$$\phi_{\mathrm{r}}(x + v/2)\phi_{\mathrm{r}}(x - v/2) = \sum_{\alpha} C_{\alpha}(v, \Lambda)\mathcal{O}^{\alpha}_{\mathrm{r}}(x), \tag{12.37}$$

in which, at cut-off Λ fixed, the coefficients $C_{\alpha}(v, \Lambda)$ are regular, even functions of the vector $\mathbf{v}$. An $\mathcal{O}^{\alpha}_{\mathrm{r}}(x)$ receives contributions from bare operators of equal or higher dimensions. We conclude that for $|\mathbf{v}|$ small the coefficient functions $C_{\alpha}(v, \Lambda)$ behave like

$$C_{\alpha}(\lambda v, \Lambda) \sim \lambda^{[\mathcal{O}_{\alpha}] - 2[\phi]}, \quad \text{for } \lambda \to 0. \tag{12.38}$$

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 $\mathbf{v} = \mathbf{0}$, diverge in a way predicted by power counting. Since $\mathbf{v}$ is small but non-vanishing, the coefficients will grow with the cut-off until Λ is of order $1/|\mathbf{v}|$. 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 (12.38) will survive, the operators of lowest dimensions will dominate expansion (12.37) for $|\mathbf{v}|$ small and the behaviour of the coefficients $C_\alpha(v,\Lambda)$ is given by equation (12.38) up to powers of $\ln v$.

The expansion (12.37) is the short distance expansion (SDE) of the product of two operators ϕ.

12.3.1 SDE at leading order

Further insight is gained into the structure of the SDE by realizing that the product $\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)$ can be considered as a regularization by point splitting of the composite operator $(\phi^2(x))_\mathrm{r}$. Let us discuss this point in detail in the framework of the ϕ^4_4 field theory. We then expect

$$\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2) \underset{|\mathbf{v}|\to 0}{\sim} C_1(v)\left[\phi^2(x)\right]_\mathrm{r} . \tag{12.39}$$

The singularities of $C_1(v)$ for $|\mathbf{v}|$ small should be directly related to the divergences of the renormalization constant Z_2 which renders $\phi^2_\mathrm{r}(x)$ finite.

In what follows, it will, therefore, be convenient to treat the product

$$\tfrac{1}{2}\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)$$

as one composite operator depending on the point x and the parameter v, in particular from the point of view of connectivity and one-particle irreducibility. To make this explicit, we introduce a notation for this operator: $\frac{1}{2}[\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)]$.

We define

$$\begin{aligned}&\tfrac{1}{2}\left\langle [\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)]\,\phi_\mathrm{r}(y_1)\ldots\phi_\mathrm{r}(y_n)\right\rangle \\ &\quad\equiv \tfrac{1}{2}\left\langle \phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)\phi_\mathrm{r}(y_1)\ldots\phi_\mathrm{r}(y_n)\right\rangle .\end{aligned} \tag{12.40}$$

However, the relation between connected correlation functions is then different:

$$\begin{aligned}&\left\langle [\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)]\,\phi_\mathrm{r}(y_1)\ldots\phi_n(y_n)\right\rangle_\mathrm{c} \\ &\qquad = \tfrac{1}{2}\left\langle \phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)\phi_\mathrm{r}(y_1)\ldots\phi_\mathrm{r}(y_n)\right\rangle_\mathrm{c} \\ &\qquad\quad + \tfrac{1}{2}\sum_{I\cup J=(1,\ldots,n)} \left\langle \phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(y_{i_1})\ldots\phi_\mathrm{r}(y_{i_p})\right\rangle_\mathrm{c} \\ &\qquad\quad \times \left\langle \phi_\mathrm{r}(x-v/2)\phi_\mathrm{r}(y_{j_1})\ldots\phi_\mathrm{r}(y_{j_{n-p}})\right\rangle_\mathrm{c} ,\end{aligned} \tag{12.41}$$

in which I and J are all non-empty partitions of $(1,\ldots,n)$.

Rather than writing explicitly the corresponding relations between proper vertices, we give the relation in terms of generating functionals. Denoting by $\mathcal{Z}(x+v/2,x-v/2;J)$ the generating functional of correlation functions with the operator insertion:

$$\mathcal{Z}(x+v/2,x-v/2;J) = \frac{1}{2}\frac{\delta^2\mathcal{Z}(J)}{\delta J(x+v/2)\delta J(x-v/2)}, \tag{12.42}$$

we find for connected correlation functions, with obvious notation,

$$\begin{aligned}\mathcal{W}(x+v/2,x-v/2;J) &= \mathcal{Z}^{-1}(J)\mathcal{Z}(x+v/2,x-v/2;J)\\ &= \frac{1}{2}\frac{\delta^2\mathcal{W}(J)}{\delta J(x+v/2)\delta J(x-v/2)}\\ &\quad+\frac{1}{2}\frac{\delta\mathcal{W}}{\delta J(x+v/2)}\frac{\delta\mathcal{W}}{\delta J(x-v/2)},\end{aligned} \tag{12.43}$$

and, finally, for the generating functional of proper vertices

$$\begin{aligned}\Gamma(x+v/2,x-v/2;\varphi) &= \frac{1}{2}\varphi(x+v/2)\varphi(x-v/2)\\ &\quad+\frac{1}{2}\left[\frac{\delta^2\Gamma}{\delta\varphi(x+v/2)\delta\varphi(x-v/2)}\right]^{-1}.\end{aligned} \tag{12.44}$$

Note that this equation is similar to equation (7.109) which we have used to prove the irreducibility of $\Gamma(\varphi)$.

To now ensure the limit (12.39), we determine $C_1(v)$ by imposing that the insertion of the operator $\frac{1}{2}C_1^{-1}(v)\,[\phi(x+v/2)\phi(x-v/2)]$ in the two-point function satisfies for any v the renormalization condition (10.28). Defining

$$\begin{aligned}C_1^{-1}(v)\int \mathrm{d}x\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{e}^{ipx+iq_1y_1+iq_2y_2}\left.\frac{\delta^2\Gamma_\mathrm{r}(x+v/2,x-v/2;\varphi)}{\delta\varphi(y_1)\delta\varphi(y_2)}\right|_{\varphi=0}\\ =(2\pi)^4\delta(p+q_1+q_2)\tilde{\Gamma}_\mathrm{r}^{(1,2)}(v;p;q_1,q_2),\end{aligned} \tag{12.45}$$

we impose

$$\tilde{\Gamma}_\mathrm{r}^{(1,2)}(v;0;0,0)=1\,. \tag{12.46}$$

Setting $q_1=q_2=0$ in equation (12.45), we derive

$$C_1(v)=\left.\frac{\delta^2\Gamma_\mathrm{r}(v/2,-v/2;\varphi)}{\delta\tilde{\varphi}(0)\delta\tilde{\varphi}(0)}\right|_{\varphi=0}, \tag{12.47}$$

in which $\tilde{\varphi}(q)$ is the Fourier transform of $\varphi(x)$.

By differentiating equation (12.44) twice with respect to φ, we can relate the r.h.s. of equation (12.47) to the φ-field proper vertices:

$$C_1(v)=1-\frac{1}{2}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\mathrm{e}^{-ikv}\left[W_\mathrm{r}^{(2)}(k)\right]^2\Gamma_\mathrm{r}^{(4)}(k,-k,0,0). \tag{12.48}$$

The equation can also be rewritten as

$$C_1(v)=1+\frac{1}{2}m_\mathrm{r}^4\left\langle\phi_\mathrm{r}(v/2)\phi_\mathrm{r}(-v/2)\tilde{\phi}_\mathrm{r}(0)\tilde{\phi}_\mathrm{r}(0)\right\rangle_\mathrm{c}. \tag{12.49}$$

We have introduced a mixed connected correlation functions, $\tilde{\phi}(p)$ being the Fourier transform of the field $\phi(x)$, and used renormalization conditions (10.25).

The coefficient $C_1(v)$ is defined in such a way that the renormalized operator $\phi_\mathrm{r}(x+v/2)\phi_\mathrm{r}(x-v/2)C_1^{-1}(v)$ then converges towards the operator $(\phi_\mathrm{r}^2(x))_\mathrm{r}$, correctly normalized.

Let us, for example, express the consequence for the four-point function:

$$C_1(v)\Gamma_{\mathrm{r}}^{(1,2)}(p;q_1,q_2) \underset{|\mathbf{v}|\to 0}{\sim} 1-\frac{1}{2}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,\mathrm{e}^{-ikv}\,W_{\mathrm{r}}^{(2)}(p/2-k)W_{\mathrm{r}}^{(2)}(p/2+k) \\ \times\Gamma_{\mathrm{r}}^{(4)}(p/2+k,p/2-k,q_1,q_2). \tag{12.50}$$

The neglected contributions are of order $v^2\left(\ln v^2\right)^p$ at any finite order in perturbation theory.

Renormalization theory tells us that equation (12.39) is valid as long as the replacement of the operator product by $\phi^2(x)$ does not generate new renormalizations. We can, therefore, use equation (12.39) in all $\phi(x)$ and $\phi^2(x)$ correlation functions except

(i) the two-point function $\langle\phi(x+v/2)\phi(x-v/2)\rangle$ which leads to $\left\langle\phi^2(x)\right\rangle$;

(ii) the four-point function $\left\langle\phi(x+v/2)\phi(x-v/2)\phi^2(y)\right\rangle$ which leads to the ϕ^2 two-point function $\left\langle\phi^2(x)\phi^2(y)\right\rangle$.

Both require an additional additive renormalization. The strategy in such cases is to first apply the CS differential operator on the correlation to generate additional $\int\phi^2(x)\mathrm{d}x$ insertions until relation (12.39) can be used. As a consequence, the SDE is modified by contributions which are solutions of the homogeneous CS equations.

12.3.2 One-loop calculation of the leading coefficient of the SDE

Equation (12.48) 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:

$$C_1(v)=1-\frac{g}{2}\int\frac{\mathrm{d}^4k}{16\pi^2}\frac{\mathrm{e}^{-ikv}}{\left(k^2+m^2\right)^2}+O\left(g^2\right). \tag{12.51}$$

It is clear from this expression that, as expected, $C_1(v)$ has at one-loop order the form

$$C_1(v)\sim A\ln\left(|\mathbf{v}|\,m\right)+B+O\left(v^2\right). \tag{12.52}$$

An often useful idea to extract an asymptotic expansion of this form is to calculate the Mellin transform $\mu(\alpha)$ of the function:

$$\mu(\alpha)=\int_0\mathrm{d}v\,v^{\alpha-1}C_1(v), \tag{12.53}$$

in which v is the length $|\mathbf{v}|$ of the vector $\mathbf{v}$. The expansion (12.52) then implies

$$\mu(\alpha)=-\frac{A}{\alpha^2}+(A\,\ln m+B)\,\frac{1}{\alpha}+O\left(1\right)\quad\text{for}\quad\alpha\to 0. \tag{12.54}$$

Applying this technique, one has to evaluate

$$f(\alpha)=\int_0\mathrm{d}v\,v^{\alpha-1}\int\frac{\mathrm{d}^4k}{16\pi^2}\frac{\mathrm{e}^{-ikv}}{\left(k^2+m^2\right)^2}. \tag{12.55}$$

As usual we rewrite the integral as

$$f(\alpha)=\int_0\mathrm{d}v\,v^{\alpha-1}\int_0^\infty t\,\mathrm{d}t\frac{\mathrm{d}^4k}{16\pi^2}\,\mathrm{e}^{-t(k^2+m^2)-ikv}\,. \tag{12.56}$$

Integrating over k, v and t, in this order, we finally obtain

$$f(\alpha) = \frac{1}{8}\left(\frac{2}{m}\right)^{\alpha} \frac{\Gamma^2(1+\alpha/2)}{\alpha^2}. \tag{12.57}$$

An expansion for $\alpha \sim 0$ yields the coefficients A and B:

$$C_1(v) = 1 - \frac{g}{16}\left[-\ln\left(\frac{vm}{2}\right) + \psi(1)\right] + O\left(g^2\right), \tag{12.58}$$

in which $\psi(z)$ is the logarithmic derivative of the function $\Gamma(z)$.

At this order, the bare parameters can be replaced by renormalized parameters.

12.4 Large Momentum Expansion of the SDE Coefficients: CS Equations

To the behaviour of the product of fields $\phi(x+v/2)\phi(x-v/2)$ at short distance 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, the constant terms in the asymptotic expansion of $C_1(v)$, yield, after Fourier transformation, terms proportional to $\delta^4(k)$ which do not contribute to the large momentum behaviour. At the same time, the algebraic structure is, for the same reason, somewhat simplified.

Let us take the example of equation (12.50). We introduce the Fourier transform $C_1(k)$ of $C_1(v)$ (as stated before, we omit the tilde sign indicating Fourier transformation when, due to the change of arguments, x, y, z, v to k, p, q, there is no ambiguity):

$$C_1(k) = \int \mathrm{e}^{ikv}\, C_1(v)\mathrm{d}^4 v\,. \tag{12.59}$$

After Fourier transformation, in the large k limit, equation (12.50) yields

$$\Gamma_{\mathrm{r}}^{(4)}(p/2+k, p/2-k, q_1, q_2) \sim -2C_1(k)\left[\Gamma_{\mathrm{r}}^{(2)}(k)\right]^2 \Gamma_{\mathrm{r}}^{(1,2)}(p; q_1, q_2), \tag{12.60}$$

in which the neglected terms are of order $(\ln k)^p/k^2$ at any finite order in perturbation theory.

More generally, due to momentum conservation, the disconnected contributions in equation (12.43) 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 r.h.s. of equation (12.44) in powers of φ yields two types of contributions: one term which becomes 1PI after amputation of the lines corresponding to the fields $\phi(x+v/2)$ and $\phi(x-v/2)$, and the other terms which remain reducible.

Figure 12.3 gives the example of the six-point function.

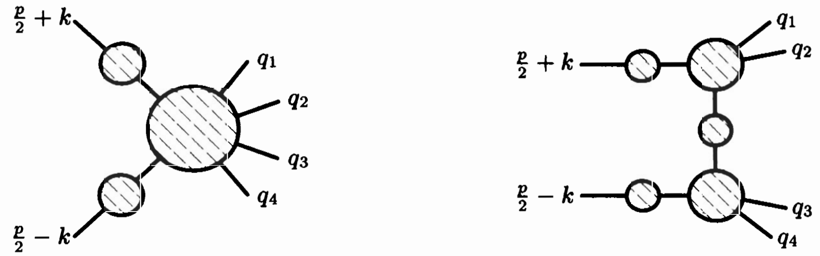


Fig. 12.3

Due to momentum conservation, in reducible terms the propagator which connects the proper vertices 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.

At leading order we can, therefore, write for $n > 0$

$$\Gamma_{\mathrm{r}}^{(n+2)}\left(\frac{p}{2}+k, \frac{p}{2}-k, q_1, \ldots, q_n\right) \underset{|k|\to\infty}{\sim} -2C_1(k)\left[\Gamma_{\mathrm{r}}^{(2)}(k)\right]^2 \Gamma_{\mathrm{r}}^{(1,n)}(p; q_1, \ldots, q_n). \quad (12.61)$$

As explained in Section 12.3, this equation generalizes to ϕ^2 insertions provided either n is positive or $l > 1$:

$$\Gamma_{\mathrm{r}}^{(l,n+2)}\left(p_1, \ldots, p_l; \frac{p}{2}+k, \frac{p}{2}-k, q_1, \ldots, q_n\right) \sim -2C_1(k)\left[\Gamma_{\mathrm{r}}^{(2)}(k)\right]^2 \times \Gamma_{\mathrm{r}}^{(l+1,n)}(p_1, \ldots, p_l, p; q_1, \ldots, q_n). \quad (12.62)$$

Callan–Symanzik equation for the first coefficient of the SDE. From now on, all quantities are assumed to be renormalized and we omit the subscript r.

We first note by comparing equations (12.48) and (12.61) that in the momentum representation the relevant function is $\Gamma^{(4)}(k, -k, 0, 0)$ which we shall call $B(k)$ in what follows:

$$B(k) \equiv \Gamma^{(4)}(k, -k, 0, 0) \underset{|\mathbf{k}|\to\infty}{\sim} -2C_1(k)\left[\Gamma^{(2)}(k)\right]^2. \quad (12.63)$$

Equation (12.61) becomes

$$\Gamma^{(n+2)}(p/2+k, p/2-k, q_1, \ldots, q_n) \sim B(k)\Gamma^{(1,n)}(p; q_1, \ldots, q_n). \quad (12.64)$$

Similarly,

$$\Gamma^{(1,n+2)}(0; p/2+k, p/2-k, q_1, \ldots, q_n) \sim B(k)\Gamma^{(2,n)}(0, p; q_1, \ldots, q_n). \quad (12.65)$$

Let us introduce a notation for the CS differential operator:

$$\mathrm{D} \equiv m\frac{\partial}{\partial m} + \beta(g)\frac{\partial}{\partial g}. \quad (12.66)$$

We write the CS equations for $\Gamma^{(n+2)}$:

$$\left[\mathrm{D} - \tfrac{1}{2}(n+2)\eta(g)\right]\Gamma^{(n+2)}(\ldots) = m^2\sigma(g)\Gamma^{(1,n+2)}(0; \ldots). \quad (12.67)$$

We now use equations (12.64) and (12.65) in the large $|k|$ limit:

$$\left[\mathrm{D} - \tfrac{1}{2}(n+2)\eta(g)\right]\left[B(k)\Gamma^{(1,n)}(p; q_1, \ldots, q_n)\right] = m^2\sigma(g)B(k)\Gamma^{(2,n)}(0, p; q_1, \ldots, q_n). \quad (12.68)$$

Using the CS equation for $\Gamma^{(1,n)}$,

$$\mathrm{D}\Gamma^{(1,n)} = \left[\frac{n}{2}\eta(g) + \eta_2(g)\right]\Gamma^{(1,n)} + m^2\sigma(g)\Gamma^{(2,n)}(0, \ldots), \quad (12.69)$$

we finally obtain an equation for $B(k)$:

$$\left[\mathrm{D} - \eta(g) + \eta_2(g)\right]B(k) \sim 0\,. \quad (12.70)$$

We can also write an equation for $C_1(k)$ using relation (12.63):

$$\left[\mathrm{D} + \eta(g) + \eta_2(g)\right]C_1(k/m, g) \sim 0\,. \quad (12.71)$$

We can compare this equation with equations (10.33,10.34) which imply

$$\left[\mathrm{D} - \eta(g) - \eta_2(g)\right]Z_2(\Lambda/m, g) = 0\,. \quad (12.72)$$

This indeed shows that $C_1(k/m, g)$ plays the same role as $Z_2^{-1}(\Lambda/m, g)$.

Due to momentum conservation, in reducible terms the propagator which connects the proper vertices 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.

At leading order we can, therefore, write for $n > 0$

$$\Gamma_{\mathrm{r}}^{(n+2)}\left(\frac{p}{2}+k, \frac{p}{2}-k, q_1, \ldots, q_n\right) \underset{|k|\to\infty}{\sim} -2C_1(k)\left[\Gamma_{\mathrm{r}}^{(2)}(k)\right]^2 \Gamma_{\mathrm{r}}^{(1,n)}(p; q_1, \ldots, q_n). \quad (12.61)$$

As explained in Section 12.3, this equation generalizes to ϕ^2 insertions provided either n is positive or $l > 1$:

$$\Gamma_{\mathrm{r}}^{(l,n+2)}\left(p_1, \ldots, p_l; \frac{p}{2}+k, \frac{p}{2}-k, q_1, \ldots, q_n\right) \sim -2C_1(k)\left[\Gamma_{\mathrm{r}}^{(2)}(k)\right]^2 \times \Gamma_{\mathrm{r}}^{(l+1,n)}(p_1, \ldots, p_l, p; q_1, \ldots, q_n). \quad (12.62)$$

Callan–Symanzik equation for the first coefficient of the SDE. From now on, all quantities are assumed to be renormalized and we omit the subscript r.

We first note by comparing equations (12.48) and (12.61) that in the momentum representation the relevant function is $\Gamma^{(4)}(k, -k, 0, 0)$ which we shall call $B(k)$ in what follows:

$$B(k) \equiv \Gamma^{(4)}(k, -k, 0, 0) \underset{|\mathbf{k}|\to\infty}{\sim} -2C_1(k)\left[\Gamma^{(2)}(k)\right]^2. \quad (12.63)$$

Equation (12.61) becomes

$$\Gamma^{(n+2)}(p/2+k, p/2-k, q_1, \ldots, q_n) \sim B(k)\Gamma^{(1,n)}(p; q_1, \ldots, q_n). \quad (12.64)$$

Similarly,

$$\Gamma^{(1,n+2)}(0; p/2+k, p/2-k, q_1, \ldots, q_n) \sim B(k)\Gamma^{(2,n)}(0, p; q_1, \ldots, q_n). \quad (12.65)$$

Let us introduce a notation for the CS differential operator:

$$\mathrm{D} \equiv m\frac{\partial}{\partial m} + \beta(g)\frac{\partial}{\partial g}. \quad (12.66)$$

We write the CS equations for $\Gamma^{(n+2)}$:

$$\left[\mathrm{D} - \tfrac{1}{2}(n+2)\eta(g)\right]\Gamma^{(n+2)}(\ldots) = m^2\sigma(g)\Gamma^{(1,n+2)}(0; \ldots). \quad (12.67)$$

We now use equations (12.64) and (12.65) in the large $|k|$ limit:

$$\left[\mathrm{D} - \tfrac{1}{2}(n+2)\eta(g)\right]\left[B(k)\Gamma^{(1,n)}(p; q_1, \ldots, q_n)\right] = m^2\sigma(g)B(k)\Gamma^{(2,n)}(0, p; q_1, \ldots, q_n). \quad (12.68)$$

Using the CS equation for $\Gamma^{(1,n)}$,

$$\mathrm{D}\Gamma^{(1,n)} = \left[\frac{n}{2}\eta(g) + \eta_2(g)\right]\Gamma^{(1,n)} + m^2\sigma(g)\Gamma^{(2,n)}(0, \ldots), \quad (12.69)$$

we finally obtain an equation for $B(k)$:

$$\left[\mathrm{D} - \eta(g) + \eta_2(g)\right]B(k) \sim 0. \quad (12.70)$$

We can also write an equation for $C_1(k)$ using relation (12.63):

$$\left[\mathrm{D} + \eta(g) + \eta_2(g)\right]C_1(k/m, g) \sim 0. \quad (12.71)$$

We can compare this equation with equations (10.33,10.34) which imply

$$\left[\mathrm{D} - \eta(g) - \eta_2(g)\right]Z_2(\Lambda/m, g) = 0. \quad (12.72)$$

This indeed shows that $C_1(k/m, g)$ plays the same role as $Z_2^{-1}(\Lambda/m, g)$.

12.5 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 into account in expansion (12.37) all operators of dimensions 2 and 4, 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, let us differentiate expansion (12.37) twice with respect to v:

$$\begin{split}\frac{1}{4}\Big[\phi\left(x+\frac{v}{2}\right)\partial_\mu\partial_\nu\phi\left(x-\frac{v}{2}\right)+\phi\left(x-\frac{v}{2}\right)\partial_\mu\partial_\nu\phi\left(x+\frac{v}{2}\right)-\partial_\mu\phi\left(x+\frac{v}{2}\right)\partial_\nu\phi\left(x-\frac{v}{2}\right)\\ -\partial_\mu\phi\left(x-\frac{v}{2}\right)\partial_\nu\phi\left(x+\frac{v}{2}\right)\Big]=\sum_\alpha\partial_\mu\partial_\nu C_\alpha(v)\mathcal{O}^\alpha_{\mathrm{r}}\left(\phi(x)\right).\end{split}\tag{12.73}$$

The product in the r.h.s. can be considered as a form regularized by point splitting of $\frac{1}{2}\left(\phi(x)\partial_\mu\partial_\nu\phi(x)-\partial_\mu\phi(x)\partial_\nu\phi(x)\right)$ which is a linear combination of operators of dimension 4, and spins 0 and 2. In Section 12.1 we have discussed the renormalization of operators of dimension 4 and spin zero. 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 $\partial_\mu\partial_\nu\left(\phi^2(x)\right)$. Rotation invariance in space tells us that operators of different spin do not mix under renormalization. Therefore, in addition to the relations (12.9) we have

$$\begin{split}\mathcal{O}_5\left(\phi(x)\right)&\equiv\partial_\mu\phi(x)\partial_\nu\phi(x)-\tfrac{1}{4}\delta_{\mu\nu}\left(\partial\phi(x)\right)^2\\ &=Z_5^{-1}\left(\mathcal{O}_5\left(\phi(x)\right)\right)_{\mathrm{r}}-B_5\left(\partial_\mu\partial_\nu-\tfrac{1}{4}\delta_{\mu\nu}\nabla^2\right)\left(\phi^2(x)\right)_{\mathrm{r}}.\end{split}\tag{12.74}$$

We can impose the renormalization conditions:

$$\begin{split}\Gamma^{(2)}_{\mathcal{O}_5}(q;p_1,p_2)&=-\left(p_1^\mu p_2^\nu+p_1^\nu p_2^\mu-\tfrac{1}{2}\delta_{\mu\nu}p_1p_2\right)+O\left(p^4\right),\\ \Gamma^{(4)}_{\mathcal{O}_5}(0;0,0,0,0)&=0.\end{split}\tag{12.75}$$

We can now use exactly the same strategy as for the leading term in the SDE. We insert expansion (12.73) 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 1PI correlation functions in the way explained in Section 12.3, that is, considering the product of fields in (12.73) as a composite operator. We finally use the renormalization conditions (12.34) and (12.75) 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 $|\mathbf{v}|$ goes to zero. In this limit the coefficients $\partial_\mu\partial_\nu C_\alpha(v)$ have singularities in $\mathbf{v}$ which 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 obtain 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 1PI correlation 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 (12.10), the difference being that the new renormalization constant, which appears in front of the contact term, is now a function of $\mathbf{v}$.

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 (12.71) and (12.72).

SDE of products of arbitrary local operators. For general local operators A and B we expect

$$A(x+v/2)B(x-v/2) = \sum_{\alpha} C^{\alpha}_{AB}(v)\mathcal{O}^{\alpha}_{\mathrm{r}}(x), \tag{12.76}$$

in which at any finite order in perturbation theory

$$C^{\alpha}_{AB}(\lambda v) \sim \lambda^{[\mathcal{O}^{\alpha}]-[A]-[B]} \quad \text{for} \quad \lambda \to 0\,, \tag{12.77}$$

up to powers of $\ln\lambda$ and the $\mathcal{O}^{\alpha}_{\mathrm{r}}(x)$ form a complete basis of local operators.

Let us take the example of

$$A \equiv B \equiv \left(\phi^2(x)\right)_{\mathrm{r}} .$$

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 tells us that among these operators $\phi^2(x)$ has the most divergent coefficient. Therefore, at leading order,

$$\left(\phi^2(x+v/2)\right)_{\mathrm{r}} \left(\phi^2(x-v/2)\right)_{\mathrm{r}} \underset{|\mathbf{v}|\to 0}{\sim} C^{1}_{\phi^2\phi^2}(v)\left(\phi^2(x)\right)_{\mathrm{r}}\,, \tag{12.78}$$

in which the coefficient can be determined by using the renormalization condition for $(\phi^2)_{\mathrm{r}}$:

$$\Gamma^{(1,2)}(q;p_1,p_2) = 1 \quad \text{for} \quad q = p_1 = p_2 = 0\,,$$

and expressing that equation (12.78) should be exact when inserted in the two-point function at the subtraction point where all momenta vanish.

It is then easy to derive RG equations for this new coefficient by inserting the relation in the CS equations for the proper vertices $\Gamma^{(l,n)}$.

We do not wish to go into further detail since the discussion is very 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 composite operators.

12.6 Light Cone Expansion (LCE) 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).

It is necessary to classify all operators not only according to their canonical dimensions, but also their spin s which characterizes their transformation properties under space rotations. The LCE takes the form

$$\phi(x+v/2)\phi(x-v/2) = \sum_{\alpha,s} C^{s}_{\alpha}\left(v^2\right) P^{s}_{\mu_1\ldots\mu_s}(v)\mathcal{O}^{s,\alpha}_{\mu_1\ldots\mu_s}(x). \tag{12.79}$$

The polynomial $P^s_{\mu_1\ldots\mu_s}(v)$ is a homogeneous, traceless for $s>0$, polynomial of the vector v_μ and the operators $\mathcal{O}^{s,\alpha}_{\mu_1\ldots\mu_s}$ form a complete basis of local operators.

For example,

$$P^2_{\mu_1,\mu_2}(v) = v_{\mu_1}v_{\mu_2} - \delta_{\mu_1\mu_2}v^2/d\,.$$

When v^2 goes to zero with v_μ finite, the polynomials P^s have a finite limit and, therefore, the coefficients $C^s_\alpha(v^2)$ 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 $C^s_\alpha(v^2)$ behaves like

$$C^s_\alpha(\lambda^2 v^2) \sim \lambda^{[\mathcal{O}^{s,\alpha}]-2[\phi]-s} \text{ up to powers of } \ln\lambda \text{ for } \lambda\to 0\,. \tag{12.80}$$

Therefore, the important parameter is no longer the dimension of the operator $\mathcal{O}^{s,\alpha}$ but a quantity called the twist τ:

$$\tau = [\mathcal{O}^{s,\alpha}] - s\,. \tag{12.81}$$

The operators of lowest twist dominate the LCE of product of operators.

In expansion (12.79) 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

$$\phi(x)\,(\partial_{\mu_1}\ldots\partial_{\mu_2}\ldots\partial_{\mu_n} - \text{ traces})\,\phi(x),$$

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 easy to write RG equations for the corresponding coefficients $C^s_\alpha(v^2)$.

Bibliographical Notes

The renormalization of composite operators is discussed by

W. Zimmermann in *Lectures in Elementary Particles and Quantum Field Theory* already quoted in Chapter 10.

The short distance expansion was postulated in

L.P. Kadanoff, *Phys. Rev. Lett.* 23 (1969) 1430; K.G. Wilson, *Phys. Rev.* 179 (1969) 1499, D3 (1971) 1818.

RG equations for the leading coefficients of the SDE are established in

K. Symanzik, *Commun. Math. Phys.* 23 (1971) 49; C.G. Callan, *Phys. Rev.* D5 (1972) 3202.

A thorough investigation is carried out in

W. Zimmermann, *Ann. Phys. (NY)* 77 (1973) 536 and 570.

See also

C.G. Callan in *Methods in Field Theory,* already quoted in Chapter 9.

The extension of Wilson's work to the light cone expansion is found in

R.A. Brandt and G. Preparata, *Nucl. Phys.* B27 (1971) 541.

The polynomial $P^s_{\mu_1\ldots\mu_s}(v)$ is a homogeneous, traceless for $s > 0$, polynomial of the vector v_μ and the operators $\mathcal{O}^{s,\alpha}_{\mu_1\ldots\mu_s}$ form a complete basis of local operators.

For example,

$$P^2_{\mu_1,\mu_2}(v) = v_{\mu_1}v_{\mu_2} - \delta_{\mu_1\mu_2}v^2/d\,.$$

When v^2 goes to zero with v_μ finite, the polynomials P^s have a finite limit and, therefore, the coefficients $C^s_\alpha(v^2)$ 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 $C^s_\alpha(v^2)$ behaves like

$$C^s_\alpha(\lambda^2 v^2) \sim \lambda^{[\mathcal{O}^{s,\alpha}]-2[\phi]-s} \text{ up to powers of } \ln\lambda \text{ for } \lambda \to 0\,. \tag{12.80}$$

Therefore, the important parameter is no longer the dimension of the operator $\mathcal{O}^{s,\alpha}$ but a quantity called the twist τ:

$$\tau = [\mathcal{O}^{s,\alpha}] - s\,. \tag{12.81}$$

The operators of lowest twist dominate the LCE of product of operators.

In expansion (12.79) 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

$$\phi(x)\,(\partial_{\mu_1}\ldots\partial_{\mu_2}\ldots\partial_{\mu_n} - \text{ traces})\,\phi(x),$$

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 easy to write RG equations for the corresponding coefficients $C^s_\alpha(v^2)$.

Bibliographical Notes

The renormalization of composite operators is discussed by

W. Zimmermann in *Lectures in Elementary Particles and Quantum Field Theory* already quoted in Chapter 10.

The short distance expansion was postulated in

L.P. Kadanoff, *Phys. Rev. Lett.* 23 (1969) 1430; K.G. Wilson, *Phys. Rev.* 179 (1969) 1499, D3 (1971) 1818.

RG equations for the leading coefficients of the SDE are established in

K. Symanzik, *Commun. Math. Phys.* 23 (1971) 49; C.G. Callan, *Phys. Rev.* D5 (1972) 3202.

A thorough investigation is carried out in

W. Zimmermann, *Ann. Phys. (NY)* 77 (1973) 536 and 570.

See also

C.G. Callan in *Methods in Field Theory,* already quoted in Chapter 9.

The extension of Wilson's work to the light cone expansion is found in

R.A. Brandt and G. Preparata, *Nucl. Phys.* B27 (1971) 541.

3 SYMMETRIES AND RENORMALIZATION

Up to now, we have discussed the problem of renormalization from the sole point of view of power counting. In Chapters 13–21 we discuss the consequences for renormalization of symmetries of the action. Indeed, when the action in the tree approximation has some symmetry properties, it can be expected that the renormalized action will not have the most general form allowed by pure power counting arguments but will instead keep some trace of the initial symmetry. Technically, this means that, as a consequence of the symmetry, the divergences generated in perturbation theory are not of generic form, and, therefore, the renormalization constants are not all independent. In this chapter, we deal only with *linear continuous symmetries* corresponding to compact Lie groups because they imply interesting formal properties; consequences of discrete symmetries can also be studied but with somewhat different methods (see the remark at the end of Section 13.5). Also, we deal below only with infinitesimal group transformations and, therefore, topological properties of groups will play no role.

Our general strategy will be as follows:

(i) We first introduce a regularization which preserves the symmetry.

(ii) We then prove identities, generally called Ward–Takahashi (WT) identities, consequences of the symmetry of the action and satisfied by the generating functional of 1PI correlation functions.

(iii) These identities imply relations between the divergences of correlation functions and thus between the counter-terms which render the theory finite. From these relations we derive the *generic* form of the counter-terms. Such an analysis is based on a loop expansion of perturbation theory.

(iv) We finally read off the properties of the renormalized action.

More generally non-trivial identities survive when terms in the action induce a *soft breaking* of symmetry. We specifically consider the examples of linear symmetry breaking and the important limiting case of spontaneous symmetry breaking, and quadratic symmetry breaking.

Finally, in Section 13.6 we apply the formalism to the physical example of chiral symmetry breaking in low-energy effective models of hadrons.

In the appendix, we outline the relation between WT identities and current conservation. We derive the energy–momentum tensor and dilatation current.

13.1 Preliminary Remarks

Before entering the discussion, we describe our notations and conventions for group and Lie algebras. A few additional algebraic remarks concerning the representation of the Lie algebra in terms of differential operators are also useful.

13.1.1 Conventions and notations

We consider continuous symmetries corresponding to various Lie groups and algebras. In this context, we adopt the following set of conventions except if explicitly stated otherwise: for continuous symmetries we only consider compact Lie groups. Fields ϕ_i

from the group point of view will be N-component vectors transforming linearly under a representation $\mathcal{R}(G)$ of a Lie group G:

$$\phi'_i = R_{ij}(\mathbf{g})\phi_j\,, \quad \mathbf{g} \in G\,. \tag{13.1}$$

To the group corresponds a Lie algebra $\mathcal{L}(G)$ whose generators can be represented by $N \times N$ real antisymmetric matrices t^α. The trace of two antisymmetric matrices defines a scalar product. We can use it to normalize the matrices by

$$\operatorname{tr} t^\alpha t^\beta = -N\delta_{\alpha\beta}\,. \tag{13.2}$$

With this convention the structure constants $f_{\alpha\beta\gamma}$ of the Lie algebra,

$$[t^\alpha, t^\beta] = f_{\alpha\beta\gamma} t^\gamma, \tag{13.3}$$

are completely antisymmetric in the three indices. The basis of the Lie algebra is fixed up to an orthogonal transformation. In the special case of unitary groups, we also sometimes represent the generators by hermitian or anti-hermitian matrices (this will be a matter of convenience) and then normalize them by

$$\operatorname{tr} t^\alpha t^\beta = N\delta_{\alpha\beta}\,,$$

(in the hermitian case).

As a consequence, as in the orthogonal case, the structure constants defined by

$$[t^\alpha, t^\beta] = i f_{\alpha\beta\gamma} t^\gamma,$$

and, thus,

$$f_{\alpha\beta\gamma} = \frac{2}{N} \operatorname{Im} \operatorname{tr}\left(t^\alpha t^\beta t^\gamma\right),$$

are completely antisymmetric.

To a group element close to the identity $\mathbf{g} = \mathbf{1} + \boldsymbol{\omega}$ corresponds a variation $\delta\phi_i = \phi'_i - \phi_i$ of the vector ϕ_i:

$$\delta\phi_i = t^\alpha_{ij}\phi_j\omega_\alpha\,, \tag{13.4}$$

in which the ω_α parametrize the infinitesimal transformation $\boldsymbol{\omega}$.

13.1.2 Lie algebra and differential operators

Let us first establish a simple group theoretical property which will become increasingly useful when we discuss the renormalization of symmetries in more complicated situations. The variation of a differentiable function $\mathcal{S}(\phi)$ under an infinitesimal transformation (13.4) is

$$\delta\mathcal{S}(\phi) = t^\alpha_{ij}\phi_j \frac{\partial\mathcal{S}}{\partial\phi_i}\omega_\alpha\,.$$

In particular, an invariant function $\mathcal{S}(\phi)$ satisfies

$$t^\alpha_{ij}\phi_j \frac{\partial\mathcal{S}}{\partial\phi_i} = 0\,. \tag{13.5}$$

The differential operators

$$\mathcal{D}_\alpha = t^\alpha_{ij}\phi_j \frac{\partial}{\partial \phi_i} \tag{13.6}$$

are thus the generators of the Lie algebra of the group G realized as differential operators acting on functions of ϕ_i. The expected commutation relations

$$[\mathcal{D}_\alpha, \mathcal{D}_\beta] = f_{\alpha\beta\gamma}\mathcal{D}_\gamma\,, \tag{13.7}$$

can be verified by direct calculation, using the commutation relations (13.3) of the generators t^α_{ij}.

Note, finally, that the equation (13.5) can also be considered as a system of differential equations for $\mathcal{S}(\phi)$.

$$\mathcal{D}_\alpha \mathcal{S}(\phi) = 0\,. \tag{13.8}$$

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 $\mathcal{D}_\alpha$, it is also a solution of all equations corresponding to operators belonging to the Lie algebra generated by $\mathcal{D}_\alpha$. The system (13.8) 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 $\mathcal{D}_\alpha$, that is, if the $\mathcal{D}_\alpha$ 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 (13.8).

We shall be concerned with situations in which ϕ is a field depending on a space variable x, and $\mathcal{S}$ is the action, functional of ϕ. The operator $\mathcal{D}_\alpha$ then has the typical form

$$\mathcal{D}_\alpha = \int \mathrm{d}x\, t^\alpha_{ij}\phi_j(x)\frac{\delta}{\delta\phi_i(x)},$$

but the analysis is the same.

13.2 Linear Global Symmetries

Definition. We call *global symmetry* a symmetry which corresponds to a transformation of the fields whose parameters are space-independent. More precisely let $\phi_i(x)$ be a set of fields transforming linearly under a representation $\mathcal{R}(G)$ of a compact Lie group G:

$$\phi'_i(x) = R_{ij}(\mathbf{g})\phi_j(x). \tag{13.9}$$

The transformation (13.9) is global if the group element $\mathbf{g}$ does not depend on the space variable 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. In what follows we explore the consequences of invariance only under infinitesimal group transformations. In the notations of Section 13.1.1, we can write the variation $\delta\phi$ of ϕ under transformation (13.9) as

$$\delta\phi_i(x) = t^\alpha_{ij}\phi_j(x)\omega_\alpha\,, \tag{13.10}$$

in which ω_α are the space-independent parameters of the transformation.

A classical action $\mathcal{S}(\phi)$ invariant under such a transformation then satisfies

$$\mathcal{D}_\alpha \mathcal{S}(\phi) \equiv \int \mathrm{d}x\, t^\alpha_{ij} \phi_j(x) \frac{\delta \mathcal{S}}{\delta \phi_i(x)} = 0 \,. \tag{13.11}$$

Regularization. In the case of linearly realized global symmetries, it is always possible to find a regularization which preserves the symmetry of the action. For purely boson field theories, we can use dimensional, lattice or momentum cut-off regularizations. In the latter case, we modify the propagator by adding to the tree action $\mathcal{S}(\phi)$ quadratic invariant terms involving higher order derivatives (like in expression (9.32)):

$$\phi_i(x)(-\nabla^2 + m^2)\phi_i(x) \mapsto \phi_i(x)(-\nabla^2 + m^2) \prod_{r=1}^{r_{\max}} (1 - \nabla^2/M_r^2)\phi_i(x), \tag{13.12}$$

in which the masses M_r are proportional to the cut-off Λ. By choosing $r_{\max}$ large enough, it is always possible to render the theory finite. The regularization terms are obviously symmetric since they are invariant under arbitrary orthogonal transformations.

In the case of massless chiral fermions, if the transformation law involves the matrix γ_S,

$$\delta\psi_i(x) = \gamma_S t^\alpha_{ij} \psi_j(x),$$

the substitution

$$\bar\psi(x) \not\partial \psi(x) \mapsto \bar\psi(x) \not\partial \prod_{r=1}^{r_{\max}} (1 - \nabla^2/M_r^2)\psi(x),$$

preserves the chiral symmetry.

Derivation of WT identities. We now consider the generating functional of correlation functions $\mathcal{Z}(J)$ corresponding to the symmetric action $\mathcal{S}(\phi)$:

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, J_i(x)\phi_i(x)\right]. \tag{13.13}$$

We have shown in Section 7.3 that general identities satisfied by the generating functional $\mathcal{Z}(J)$, like the equation of motion, can be obtained by expressing that the functional integral is invariant in an infinitesimal change of variables. Here, we use this observation to derive the consequences of equation (13.11) for $\mathcal{Z}(J)$. We perform a change of variables of the form of a transformation (13.10), setting

$$\phi_i(x) = \phi'_i(x) + t^\alpha_{ij} \phi'_j(x) \omega_\alpha \,, \tag{13.14}$$

in the functional integral (13.13). As a consequence of the symmetry as expressed by equation (13.11), the action $\mathcal{S}(\phi)$ and, therefore, the regularized action $\mathcal{S}_\Lambda(\phi)$ are left invariant by the transformation (13.14). The measure of integration $[\mathrm{d}\phi_i]$ in the functional integral (13.13) is the flat euclidean measure which is invariant under orthogonal transformation (more generally in any unimodular transformation, that is, corresponding to matrices of determinant 1). The only variation comes from the source term. This implies

$$0 = \delta\mathcal{Z}(J) = \int [\mathrm{d}\phi']\, \delta\, [\text{source term}] \exp\left[-\mathcal{S}(\phi') + \int \mathrm{d}x\, J_i(x)\phi'_i(x)\right].$$

The variation of the source term is

$$\delta\,[\text{source term}] = \int \mathrm{d}x\, J_i(x) t^\alpha_{ij} \phi'_j(x) \omega_\alpha \,.$$

This leads to the equation

$$0 = \omega_\alpha \int [\mathrm{d}\phi] \int \mathrm{d}x\, J_i(x) t^\alpha_{ij} \phi_j(x) \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\, J_k(x)\phi_k(x)\right]. \tag{13.15}$$

We have now renamed ϕ'_i, ϕ_i since ϕ' is a dummy integration variable.

Equation (13.15), being valid for any set of parameters ω_α, can be rewritten for each component α. Finally, the identity

$$\begin{aligned}\int [\mathrm{d}\phi]\, \phi_i(x) \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}y\, J_k(y)\phi_k(y)\right] \\ = \frac{\delta}{\delta J_i(x)} \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}y\, J_k(y)\phi_k(y)\right],\end{aligned} \tag{13.16}$$

allows us to rewrite equation (13.15) as an equation for the functional $\mathcal{Z}(J)$:

$$\int \mathrm{d}x\, t^\alpha_{ij} J_i(x) \frac{\delta \mathcal{Z}(J)}{\delta J_j(x)} = 0\,. \tag{13.17}$$

Equation (13.17) immediately implies an identical equation for the generating functional $\mathcal{W}(J) = \ln \mathcal{Z}(J)$ of connected correlation functions:

$$\int \mathrm{d}x\, t^\alpha_{ij} J_i(x) \frac{\delta \mathcal{W}(J)}{\delta J_j(x)} = 0\,. \tag{13.18}$$

Expanding equation (13.18) in a power series of the source $J(x)$, we obtain identities between the connected correlation functions which describe the physical implications of the symmetry of the action.

However, for renormalization purposes it is more useful to derive an equation for the 1PI functional $\Gamma(\varphi)$. We, therefore, perform a Legendre transformation:

$$\begin{aligned}\Gamma(\varphi) + \mathcal{W}(J) &= \int \mathrm{d}x\, J_i(x)\varphi_i(x), \\ \varphi_i(x) &= \frac{\delta \mathcal{W}}{\delta J_i(x)}.\end{aligned} \tag{13.19}$$

Expressing equation (13.18) in terms of φ and Γ we obtain

$$\int \mathrm{d}x\, t^\alpha_{ij} \varphi_i(x) \frac{\delta \Gamma}{\delta \varphi_j(x)} = 0\,, \tag{13.20}$$

which, expanded in powers of φ, yields WT identities for proper vertices. The equation implies that the regularized functional $\Gamma(\varphi)$ is invariant under the transformation (13.10).

Renormalization. We now perform a loop expansion of $\Gamma(\varphi)$:

$$\Gamma(\varphi) = \sum_{l=0}^{\infty} \Gamma_l(\varphi) g^l. \tag{13.21}$$

The parameter g is any coupling constant playing the role of $\hbar$ and introduced to order the loop expansion. Since equation (13.20) is linear in $\Gamma(\varphi)$ and independent of g, all functionals $\Gamma_l(\varphi)$ also satisfy equation (13.20).

The functional $\Gamma_0(\varphi)$ is just the action $\mathcal{S}(\varphi)$ and satisfies, by assumption, equation (13.20). The regularized one-loop functional $\Gamma_1(\varphi)$ satisfies (13.20):

$$\int \mathrm{d}x\, t^{\alpha}_{ij}\varphi_i(x)\frac{\delta\Gamma_1(\varphi)}{\delta\varphi_j(x)} = 0\,. \tag{13.22}$$

We now perform an 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 (13.22) 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 $\Gamma_1^{\mathrm{div}}(\varphi)$:

$$\int \mathrm{d}x\, t^{\alpha}_{ij}\varphi_i(x)\frac{\delta\Gamma_1^{\mathrm{div}}(\varphi)}{\delta\varphi_j(x)} = 0\,. \tag{13.23}$$

General renormalization theory tells us that $\Gamma_1^{\mathrm{div}}(\varphi)$ is a general local functional of the fields restricted only by power counting; equation (13.23) tells us in addition that it is symmetric. Adding $-\Gamma_1^{\mathrm{div}}(\varphi)$ 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 (13.20). After one-loop renormalization $\Gamma_2(\varphi)$ has only local divergences which also satisfy equation (13.23) and all arguments can be repeated. It is clear that the arguments extend to all orders.

The conclusion is that the renormalized action $\mathcal{S}_{\mathrm{r}}$ is the most general local functional of the field $\phi_i(x)$ compatible with power counting and invariant under the transformation (13.10).

Each reader familiar with perturbative calculations will realize that this is a pedantic derivation of a rather obvious result. However, since the same strategy, suitably adapted, allows us to discuss much more general situations, we believe that it has been useful to explain it first in a case in which it can be easily understood.

Note, finally, that we have renormalized using a minimal subtraction scheme. Additional finite renormalizations which are consistent with the symmetry can still be performed.

13.3 Linear Symmetry Breaking

For some applications (see for example Sections 13.6, 26.5) it is useful to consider the following situation: the action $\mathcal{S}(\phi)$ is the sum of a symmetric part $\mathcal{S}_{\mathrm{sym}}(\phi)$, that is, invariant under the transformation (13.9), and a term breaking the symmetry linear in the fields $\phi_i(x)$:

$$\mathcal{S}(\phi) = \mathcal{S}_{\mathrm{sym}}(\phi) - \int c_i\phi_i(x)\mathrm{d}x\,, \tag{13.24}$$

in which $\mathbf{c}$ is a constant vector.

An example of such a situation is provided by the action:

$$\mathcal{S}(\phi) = \int \mathrm{d}x \left[\frac{1}{2}\left(\partial_\mu\phi\right)^2 + \frac{1}{2}m^2\phi^2 + \frac{1}{4!}g\left(\phi^2\right)^2 - \mathbf{c}\cdot\phi\right], \tag{13.25}$$

in which $\boldsymbol{\phi}(x)$ is an N-component vector. The action $\mathcal{S}(\boldsymbol{\phi})$ is the sum of an $O(N)$ invariant part and a linear symmetry breaking term.

The perturbative expansion corresponding to action (13.25) is obtained by the following method: one first looks for a classical minimum of the action which corresponds to a constant field $\mathbf{v}_0$ satisfying

$$\frac{\delta \mathcal{S}_{\mathrm{sym}}(\mathbf{v}_0)}{\delta \phi_i} - c_i = 0\,, \tag{13.26}$$

with the condition

$$\frac{\delta^2 \mathcal{S}_{\mathrm{sym}}(\mathbf{v}_0)}{\delta \phi_i \delta \phi_j} \geq 0\,,$$

in the matrix sense.

In the example (13.25) v_{0i} satisfies the equation

$$\left(m^2 + \frac{g}{6}\mathbf{v}_0^2\right) v_{0i} = c_i\,. \tag{13.27}$$

If the action has several minima, one is in general instructed to choose the absolute minimum of the potential but this is irrelevant from the point of view of formal perturbation theory. The quantity $\mathbf{v}_0$ is, in the tree approximation, the expectation value (vacuum expectation value in the particle physics language) of the field $\boldsymbol{\phi}$.

One then translates the field $\boldsymbol{\phi}$, setting

$$\boldsymbol{\phi}(x) = \mathbf{v}_0 + \boldsymbol{\chi}(x). \tag{13.28}$$

After translation, the action no longer contains a linear term and the perturbative calculation proceeds in the standard manner. However, the example (13.25) shows that after translation the mass term is no longer symmetric and a non-symmetric χ^3 interaction has been generated. Correlation functions will no longer be symmetric and the form of the UV divergences from the point of view of the symmetry is *a priori* unknown. It is thus important to understand whether the structure of the renormalized action reflects in some way the structure of the action (13.24).

The answer here follows from a simple argument. With obvious notation, we have

$$\mathcal{Z}(\mathbf{J}) = \mathcal{Z}_{\mathrm{sym}}(\mathbf{J} + \mathbf{c}), \tag{13.29}$$

and, thus,

$$\mathcal{W}(\mathbf{J}) = \mathcal{W}_{\mathrm{sym}}(\mathbf{J} + \mathbf{c}). \tag{13.30}$$

Equation (13.18) then, in particular, implies

$$\int \mathrm{d}x\, t^{\alpha}_{ij} \left[J_i(x) + c_i\right] \frac{\delta \mathcal{W}(\mathbf{J})}{\delta J_j(x)} = 0\,. \tag{13.31}$$

Expanding in powers of $J_i(x)$, we obtain a set of relations (WT identities) between connected correlation functions which can be most conveniently expressed in the momentum representation:

$$c_i t^{\alpha}_{ij} \tilde{W}^{(n+1)}_{jk_1,\ldots,k_n}(0, p_1, \ldots, p_n) + \sum_{r=1}^{n} t^{\alpha}_{k_r j} \tilde{W}^{(n)}_{k_1,\ldots,k_{r-1},j,k_{r+1},\ldots,k_n}(p_1, \ldots, p_n) = 0\,. \tag{13.32}$$

in which $\phi(x)$ is an N-component vector. The action $\mathcal{S}(\phi)$ is the sum of an $O(N)$ invariant part and a linear symmetry breaking term.

The perturbative expansion corresponding to action (13.25) is obtained by the following method: one first looks for a classical minimum of the action which corresponds to a constant field $\mathbf{v}_0$ satisfying

$$\frac{\delta \mathcal{S}_{\text{sym}}(\mathbf{v}_0)}{\delta \phi_i} - c_i = 0\,, \tag{13.26}$$

with the condition

$$\frac{\delta^2 \mathcal{S}_{\text{sym}}(\mathbf{v}_0)}{\delta \phi_i \delta \phi_j} \geq 0\,,$$

in the matrix sense.

In the example (13.25) v_{0i} satisfies the equation

$$\left(m^2 + \frac{g}{6}\mathbf{v}_0^2\right) v_{0i} = c_i\,. \tag{13.27}$$

If the action has several minima, one is in general instructed to choose the absolute minimum of the potential but this is irrelevant from the point of view of formal perturbation theory. The quantity $\mathbf{v}_0$ is, in the tree approximation, the expectation value (vacuum expectation value in the particle physics language) of the field ϕ.

One then translates the field ϕ, setting

$$\phi(x) = \mathbf{v}_0 + \boldsymbol{\chi}(x). \tag{13.28}$$

After translation, the action no longer contains a linear term and the perturbative calculation proceeds in the standard manner. However, the example (13.25) shows that after translation the mass term is no longer symmetric and a non-symmetric χ^3 interaction has been generated. Correlation functions will no longer be symmetric and the form of the UV divergences from the point of view of the symmetry is *a priori* unknown. It is thus important to understand whether the structure of the renormalized action reflects in some way the structure of the action (13.24).

The answer here follows from a simple argument. With obvious notation, we have

$$\mathcal{Z}(\mathbf{J}) = \mathcal{Z}_{\text{sym}}(\mathbf{J} + \mathbf{c}), \tag{13.29}$$

and, thus,

$$\mathcal{W}(\mathbf{J}) = \mathcal{W}_{\text{sym}}(\mathbf{J} + \mathbf{c}). \tag{13.30}$$

Equation (13.18) then, in particular, implies

$$\int \mathrm{d}x\, t^{\alpha}_{ij}\, [J_i(x) + c_i]\, \frac{\delta \mathcal{W}(\mathbf{J})}{\delta J_j(x)} = 0\,. \tag{13.31}$$

Expanding in powers of $J_i(x)$, we obtain a set of relations (WT identities) between connected correlation functions which can be most conveniently expressed in the momentum representation:

$$c_i t^{\alpha}_{ij} \tilde{W}^{(n+1)}_{jk_1,\ldots,k_n}(0, p_1, \ldots, p_n) + \sum_{r=1}^{n} t^{\alpha}_{k_r j} \tilde{W}^{(n)}_{k_1,\ldots,k_{r-1},j,k_{r+1},\ldots,k_n}(p_1, \ldots, p_n) = 0\,. \tag{13.32}$$

The 1PI functional Γ is given by the Legendre transformation:

$$\begin{aligned}\Gamma(\varphi) + \mathcal{W}(\mathbf{J}) &= \int \mathrm{d}x\, J_i(x)\varphi_i(x), \\ \varphi_i(x) &= \frac{\delta \mathcal{W}}{\delta J_i(x)} = \frac{\delta \mathcal{W}_{\mathrm{sym}}(\mathbf{J}+\mathbf{c})}{\delta J_i(x)}.\end{aligned} \tag{13.33}$$

In the symmetric situation these relations read instead:

$$\begin{aligned}\Gamma_{\mathrm{sym}}(\xi) + \mathcal{W}_{\mathrm{sym}}(\mathbf{J}) &= \int \mathrm{d}x\, J_i(x)\xi_i(x), \\ \xi_i(x) &= \frac{\delta \mathcal{W}_{\mathrm{sym}}(\mathbf{J})}{\delta J_i(x)}.\end{aligned} \tag{13.34}$$

Replacing $J_i(x)$ by $J_i(x) + c_i$ in the relations (13.34) we obtain

$$\begin{aligned}\Gamma_{\mathrm{sym}}(\varphi) + \mathcal{W}_{\mathrm{sym}}(\mathbf{J}+\mathbf{c}) &= \int \mathrm{d}x\, (J_i(x) + c_i)\varphi_i(x), \\ \varphi_i(x) &= \frac{\delta \mathcal{W}_{\mathrm{sym}}(\mathbf{J}+\mathbf{c})}{\delta J_i(x)},\end{aligned} \tag{13.35}$$

and, therefore, comparing (13.33) with (13.35),

$$\Gamma(\varphi) = \Gamma_{\mathrm{sym}}(\varphi) - \int \mathrm{d}x\, c_i\varphi_i(x). \tag{13.36}$$

This identity shows that the divergences of the functionals $\Gamma(\varphi)$ and $\Gamma_{\mathrm{sym}}(\varphi)$ are identical. If we, therefore, replace the regularized symmetric action by the renormalized symmetric action, the theory is finite for any value of c_i. This is informally expressed by saying that the linear breaking term is not renormalized.

To obtain the 1PI correlation functions of ϕ_i, we then have to translate φ by the ϕ field expectation value setting (see Section 7.5)

$$\varphi_i(x) = v_i + \chi_i(x) \tag{13.37}$$

with

$$\left.\frac{\delta\Gamma}{\delta\varphi_i(x)}\right|_{\varphi_i(x)=v_i} = 0 \implies \frac{\delta\Gamma_{\mathrm{sym}}}{\delta\varphi_i(x)}(v_i) = c_i\,, \tag{13.38}$$

and $\delta^2\Gamma(\mathbf{v})/\delta\phi_i\delta\phi_j \geq 0$.

The 1PI correlation functions are then the coefficients of the expansion of $\Gamma(\varphi)$ in powers of χ. In the tree approximation, one recovers $v_i = v_{0i}$.

The WT identities for $\Gamma(\varphi)$ can be inferred from the identity (13.20) for Γ_{sym}:

$$\int \mathrm{d}x\, t^{\alpha}_{ij}\left[\frac{\delta\Gamma}{\delta\varphi_i(x)} + c_i\right]\varphi_j(x) = 0\,, \tag{13.39}$$

which after the translation (13.37) becomes

$$\int \mathrm{d}x\, t^{\alpha}_{ij}\left[\frac{\delta\Gamma}{\delta\chi_i}(\chi + \mathbf{v}) + c_i\right](\chi_j + v_j) = 0\,. \tag{13.40}$$

Application. Let us show that this identity leads to some non-trivial relations between the 1PI correlation functions. Setting $\chi = 0$ we obtain

$$t^{\alpha}_{ij} c_i v_j = 0 \,, \tag{13.41}$$

which shows the breaking vector **c** and the expectation value **v** are left invariant by the same subgroup of G. In the example of the $O(N)$ symmetry, equation (13.41) implies that the vector **v** is proportional to the vector **c**.

Differentiating once with respect to $\chi_k(y)$ and then setting χ equal to zero, we relate the one- and two-point functions:

$$\int \mathrm{d}x \left[v_j t^{\alpha}_{ij} \Gamma^{(2)}_{ik}(x,y) + t^{\alpha}_{ik} c_i \delta(x-y) \right] = 0 \tag{13.42}$$

with

$$\Gamma^{(2)}_{ij}(x,y) = \left. \frac{\delta^2 \Gamma(\chi + v)}{\delta \chi_i(x) \delta \chi_j(y)} \right|_{\chi=0} .$$

In terms of the Fourier transform $\tilde{\Gamma}^{(2)}_{ij}(p)$ of the two-point function,

$$\Gamma^{(2)}_{ij}(x,y) = \int \frac{\mathrm{d}^d p}{(2\pi)^d} \, \mathrm{e}^{-ip(x-y)} \, \tilde{\Gamma}^{(2)}_{ij}(p), \tag{13.43}$$

equation (13.42) becomes

$$v_j t^{\alpha}_{ji} \tilde{\Gamma}^{(2)}_{ik}(0) + t^{\alpha}_{ki} c_i = 0 \,. \tag{13.44}$$

This equation determines the geometrical structure of the zero momentum propagator in the presence of the linear symmetry breaking term.

In the example (13.25), the identity (13.44) yields the value of the propagator of the components of the field orthogonal to the vector **c**, at zero momentum:

$$\tilde{\Gamma}^{(2)}_{T}(0) = c/v \,.$$

Equation (13.44) is the last equation which involves c_i explicitly. The terms of higher degree in χ are functions only of the expectation value v_i. 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 zero and the n-point function $\tilde{\Gamma}^{(n)}$:

$$v_j t^{\alpha}_{jk} \tilde{\Gamma}^{(n+1)}_{ki_1 \ldots i_n}(0, p_1, \ldots, p_n) + \sum_{r=1}^{n} t^{\alpha}_{i_r k} \tilde{\Gamma}^{(n)}_{i_1, \ldots, i_{r-1}, k, i_{r+1}, \ldots, i_n}(p_1, \ldots, p_n) = 0 \,. \tag{13.45}$$

For example, this equation for $n = 2$ reads

$$v_j t^{\alpha}_{ji} \tilde{\Gamma}^{(3)}_{ikl}(0, p, -p) + t^{\alpha}_{li} \tilde{\Gamma}^{(2)}_{ik}(p) + t^{\alpha}_{ki} \tilde{\Gamma}^{(2)}_{il}(p) = 0 \,.$$

If we choose to renormalize by fixing the value of the primitively divergent correlation functions at some given point in momentum space, then the set of WT identities implies relations between the different parameters. Apart from the vector **v**, the non-symmetric theory depends on the same number of independent parameters as the symmetric theory. In the example of the $O(N)$ symmetric $(\phi^2)^2$ field theory in four dimensions, it is possible to impose one arbitrary renormalization condition on $\Gamma^{(4)}_{1111}(p_i)$ and two conditions on $\Gamma^{(2)}_{11}(p)$. All others are given by the WT identities (13.45) used for $n = 1$ to 4.

13.4 Spontaneous Symmetry Breaking

Spontaneous symmetry breaking (SSB) is a possible limit of linear symmetry breaking when the breaking parameter goes to zero. As we discuss below, in this limit, the action becomes symmetric, but depending on the values of other parameters (in our examples the coefficient of the ϕ^2 operator), the physics may or may not be symmetric. Many physical models in particle physics are based on the concept of SSB. The reason is that the mechanism of SSB allows one to construct models with broken symmetries that depend on no more parameters than the symmetric models. The appearance of massless particles (Goldstone modes) is, in general, the most characteristic feature of such models (in the absence of gauge symmetries).

The physics of SSB will be extensively discussed, in the statistical context, in Chapters 23–31 devoted to critical phenomena. One important issue from the point of view of perturbation theory is the following. SSB, in the perturbative framework, is associated with degenerate classical minima. Each minimum is the starting point of a perturbative expansion. Naively one would expect that it is necessary to sum over the contributions of all minima. If we calculate only expectation values of group invariant correlation functions all minima give the same contribution. Summing over all minima yields a factor which disappears in the normalization of the functional integral. However, for non-invariant correlation functions a summation over all minima is equivalent to a group average and projects onto invariant functions: as a consequence, all non-vanishing correlation functions are invariant and the field has no expectation value.

In fact the correct procedure in the case of degenerate classical minima depends on the real physical situation beyond perturbation theory. In the absence of phase transitions, one must sum over all minima. Quantum fluctuations restore the symmetry broken in the classical approximation and the true ground state is unique.

In the case of phase transitions, there is a breaking of ergodicity in the ordered phase, and one must, instead, choose one specific minimum. The ground state is degenerate.

Below, we assume that the situation of SSB is realized.

Classical analysis: the $O(N)$ example. We first consider the example of the $O(N)$ symmetric model and discuss the expectation value of the field in the tree or classical approximation. The action density for a constant field $\boldsymbol{\phi}$ is

$$\mathcal{E}(\boldsymbol{\phi}) \equiv \mathcal{S}(\boldsymbol{\phi})\,/\text{volume} = \frac{1}{2}m^2\boldsymbol{\phi}^2 + \frac{1}{4!}g(\boldsymbol{\phi}^2)^2 - \mathbf{c}\cdot\boldsymbol{\phi}\,. \tag{13.46}$$

As long as $\mathbf{c}$ does not vanish, it is possible to pass continuously from a situation in which the parameter m^2 is positive to a situation in which m^2 is negative (remember that m is no longer the physical mass) without encountering any singularity. For instance, the expectation value $\mathbf{v}$ is, at $\mathbf{c}$ fixed, a regular function of m^2 at $m^2 = 0$. If instead $\mathbf{c}$ vanishes, the expectation value $\mathbf{v}$ vanishes identically for $m^2 > 0$ and takes a non-trivial value for $m^2 < 0$ such that

$$|\mathbf{v}| = \sqrt{-6m^2/g}, \tag{13.47}$$

as can be easily understood by displaying the action density in both cases (see figure 13.1).

In the latter case, 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 (13.47).

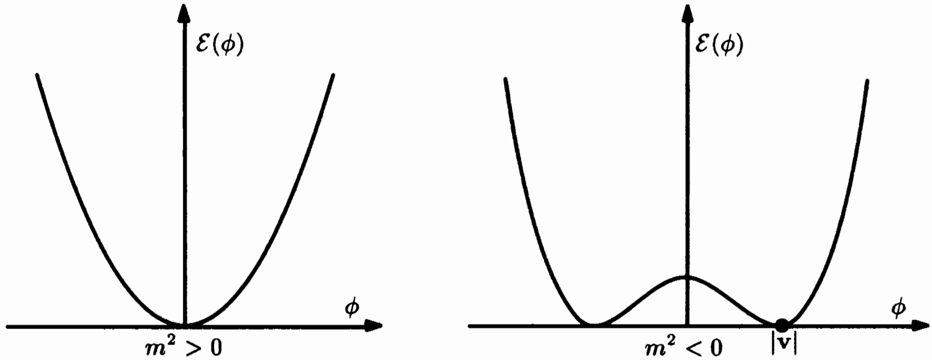


Fig. 13.1 Section of the ϕ-action density.

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:

$$\boldsymbol{\phi}(x) = \mathbf{v} + \boldsymbol{\chi}(x)\,.$$

The $\boldsymbol{\chi}$-field mass matrix is obtained by calculating the second derivatives of the action density at the minimum. Using equation (13.47), we find

$$\left.\frac{\partial \mathcal{E}}{\partial \phi_i \partial \phi_j}\right|_{\boldsymbol{\phi}=\mathbf{v}} = \frac{g}{3} v_i v_j\,.$$

This matrix has $N-1$ zero eigenvalues corresponding to eigenvectors orthogonal to $\mathbf{v}$. This is not surprising since the potential is flat along a group orbit. The physical consequence is that spontaneous breaking of a continuous symmetry implies the appearance of massless (Goldstone) modes: from the point of view of particle physics massless scalar particles called Goldstone bosons.

General symmetry group. We now examine a more 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 which leaves the vector $\mathbf{v}$ invariant. By definition the p generators of the Lie algebra $\mathcal{L}(H)$ of H satisfy

$$\mathcal{L}(H): \qquad 1 \le \alpha \le p \;\Rightarrow\; t^{\alpha}_{ij} v_j = 0\,,$$

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

$$\mathcal{L}(G/H): \qquad \sum_{\alpha>p} t^{\alpha}_{ij} v_j \omega_\alpha = 0 \;\Rightarrow\; \omega_\alpha = 0 \quad \text{for all } \alpha\,.$$

For $\alpha > p$, the vectors $(v^\alpha)_i = t^{\alpha}_{ij} v_j$ are thus linearly independent. We then parametrize the field $\boldsymbol{\phi}$ in the form of a group element acting on a vector:

$$\boldsymbol{\phi}(x) = \exp\left(\sum_{\alpha>p} t^\alpha \xi^\alpha(x)\right)\left(\mathbf{v} + \boldsymbol{\rho}(x)\right) = \mathbf{v} + \xi^\alpha(x) t^\alpha \mathbf{v} + \boldsymbol{\rho}(x) + \cdots\,,$$

in which $\boldsymbol{\rho}(x)$ has components only in the subspace orthogonal to all vectors $t^\alpha \mathbf{v}$. In the $O(N)$ example, $\boldsymbol{\rho}$ has only one component along $\mathbf{v}$. This parametrization is such that the mapping of fields $\{\boldsymbol{\rho}(x), \xi^\alpha(x)\} \mapsto \boldsymbol{\phi}(x) - \mathbf{v}$ can be inverted for small fields. This property ensures that if the fluctuations of the field $\boldsymbol{\phi}$ around its expectation value are in some sense small, perturbation theory is at least qualitatively sensible.

Inserting this parametrization into the action we note the following: the contributions to the action which 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 spontaneous breaking of symmetry of a group G to a subgroup H, the group which leaves the field expectation value invariant, yields a number of massless Goldstone modes (bosons) equal to the number of generators of G which do not belong to H. This result is valid in the classical approximation. We now generalize it to the full quantum theory.

WT identities and spontaneous symmetry breaking. To connect continuously the two phases, symmetric and with SSB, without encountering any singularity, we start from the situation $m^2 > 0$, $\mathbf{c} = 0$; we give to $\mathbf{c}$ a non-trivial value, perform the continuation from $m^2 > 0$ to $m^2 < 0$, and again take the vanishing $\mathbf{c}$ limit. We then assume the existence of non-trivial solutions to the equation

$$\left.\frac{\delta \Gamma}{\delta \varphi_i(x)}\right|_{\varphi_i(x)=v_i} = 0 . \tag{13.48}$$

In Section 7.10 we have explained how the existence of solutions to this equation is consistent with the convexity of the function $\Gamma(\mathbf{v})$ (equation (7.81)).

Since the WT identities (13.40) hold for any value of the parameters and we have made an analytic continuation, we still have in the $m^2 < 0$, $\mathbf{c} = 0$ limit

$$\int \mathrm{d}x \, t^\alpha_{ij} \frac{\delta \Gamma(\chi + v)}{\delta \chi_i(x)} (\chi_j + v_j) = 0 , \tag{13.49}$$

the direction of v_i being fixed by equation (13.41).

Goldstone modes. One important consequence of WT identities is obtained by taking the $\mathbf{c} = \mathbf{0}$ limit in equation (13.44):

$$v_j t^\alpha_{ji} \tilde{\Gamma}^{(2)}_{ik}(0) = 0 . \tag{13.50}$$

To explain the meaning of this equation, 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 $(v^\alpha)_i = t^\alpha_{ij} v_j$ are linearly independent, equation (13.50) implies that the real symmetric matrix $\tilde{\Gamma}_{ij}(0)$ has as many eigenvectors with eigenvalue zero as there are generators in $\mathcal{L}(G/H)$, confirming the classical analysis. The corresponding components of the field are Nambu–Goldstone modes associated with the spontaneous breaking of the G-symmetry, associated with massless particles.

13.5 Quadratic Symmetry Breaking

A symmetry may be broken by terms of higher canonical dimensions. We give here a detailed discussion only of the case of the quadratic symmetry breaking. We then briefly indicate how the results generalize in the case of breaking terms of even higher dimensions. We consider the action

$$\mathcal{S}(\phi) = \mathcal{S}_{\text{sym}}(\phi) + \frac{1}{2}\mu_{ij} \int \mathrm{d}x \, \phi_i(x)\phi_j(x), \tag{13.51}$$

in which μ_{ij} is a symmetric traceless constant matrix. We assume μ_{ij} traceless without loss of generality since a term proportional to the unit matrix can always be absorbed into the symmetric action $\mathcal{S}_{\text{sym}}(\phi)$. In the action (13.51) the interactions are symmetric but the mass terms break the symmetry.

We can try again to derive WT identities by performing an infinitesimal change of variables in the functional integral:

$$\phi_i = \phi'_i + t^{\alpha}_{ij}\omega_\alpha \phi'_j \,.$$

The variation of the integrand now comes both from the source term and the breaking term:

$$\delta \left[\mathcal{S}(\phi) - \int \mathrm{d}x \, J_i(x)\phi_i(x)\right] = \omega_\alpha \int \mathrm{d}x \left[\mu_{ij}\phi_i(x) t^{\alpha}_{jk}\phi_k(x) - J_i(x) t^{\alpha}_{ij}\phi_j(x)\right].$$

Expressing as usual that the result of the functional integral is not modified by a change of variables we obtain the equation:

$$\int [\mathrm{d}\phi]\, \delta \left[\mathcal{S}(\phi) - \int \mathrm{d}x \, J_i(x)\phi_i(x)\right] \exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x \, J_i(x)\phi_i(x)\right] = 0 \,.$$

Using equation (13.51) and replacing factors of the form $\phi_i(x)$ by $\delta / \delta J_i(x)$ we derive an equation for $\mathcal{Z}(J)$:

$$\int \mathrm{d}x \left[\mu_{ij} t^{\alpha}_{jk} \frac{\delta^2}{\delta J_i(x) \delta J_k(x)} - J_i(x) t^{\alpha}_{ij} \frac{\delta}{\delta J_j(x)}\right] \mathcal{Z}(J) = 0 \,. \tag{13.52}$$

Two features distinguish this equation from the equation of the symmetric case ($\mu_{ij} = 0$):

(i) It involves functional second derivatives with respect to the sources. The corresponding equation for $\mathcal{W}(J)$ then also involves a term of the form $\delta^2 \mathcal{W} / [\delta J(x)]^2$. If we now try to perform a Legendre transformation, we have to introduce the quantity

$$\frac{\delta^2 \mathcal{W}}{[\delta J(x)]^2} = \left[\frac{\delta^2 \Gamma}{\delta\varphi(x)\delta\varphi(y)}\right]^{(-1)} \Bigg|_{x=y}, \tag{13.53}$$

the inverse being understood in the sense of kernels. The WT identities take a very complicated form.

(ii) It is no longer a relation really between ϕ-field correlation functions because the two functional derivatives are taken at the same point; it, instead, also involves insertions of the composite operator $\phi_i(x)\phi_j(x)$.

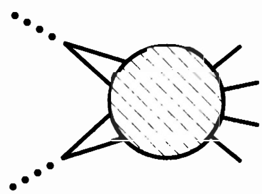


Fig. 13.2

The difficulties we encounter 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 $\frac{1}{2}\mu_{ij}\int \mathrm{d}x\,\phi_i(x)\phi_j(x)$ (see figure 13.2).

These insertions, as we have already extensively discussed, 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 such an 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 (13.51). If we want to write WT identities for renormalized quantities, we have to also renormalize the operator $\int \mathrm{d}x\,\mu_{ij}t^{\alpha}_{jk}\phi_i(x)\phi_k(x)$ 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.

The general method. The method we now introduce, partially inspired by the analysis of the equation of motion of Section 12.2, is fairly general and will allow us to discuss many cases of renormalization with symmetries, as the subsequent chapters devoted to different types of symmetries will show.

The basic idea is to calculate the variation of the breaking term under an infinitesimal group transformation, to collect all new linearly independent composite operators generated in this way and to add source terms for them in the action. Let us call $C_\alpha(\phi)$ such operators. In our example

$$C_\alpha(\phi) \equiv \int \mathrm{d}x\,\mu_{ij}t^{\alpha}_{jk}\phi_i(x)\phi_j(x).$$

Since we have now added a source for C_α in the action, we have to worry about the effect of an infinitesimal transformation on $C_\alpha(\phi)$:

$$\delta C_\alpha(\phi) = \int \mathrm{d}x\,\frac{\delta C_\alpha}{\delta\phi_i(x)}t^{\beta}_{ij}\phi_j(x)\omega_\beta\,.$$

Two cases may arise: either $\delta C_\alpha(\phi)$ is a linear combination of $\phi_i(x)$ and $C_\alpha(\phi)$ and we proceed deriving WT identities, or new independent operators are generated and we again add sources for them in the action.

We repeat the procedure as long as necessary. It is easy to verify that under these conditions the WT identities will always be first order differential equations for the generating functional $\mathcal{Z}$ considered as a functional of all sources. Several examples will illustrate this point.

Application. 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. We, therefore, immediately consider the general action $\mathcal{S}(\phi, K)$:

$$\mathcal{S}(\phi, K) = \mathcal{S}_{\text{sym}}(\phi) + \frac{1}{2}\int \mathrm{d}x\, K_{ij}(x)\phi_i(x)\phi_j(x) \tag{13.54}$$

with

$$K_{ij}(x) = K_{ji}(x), \qquad K_{ii}(x) = 0\,.$$

Actually we need only insertions at zero momentum and we could, thus, restrict ourselves to constant sources K_{ij}. However, it is not more difficult to use space-dependent sources. Furthermore, when the symmetric theory is massless, zero momentum insertions could lead to IR divergences which in this way are avoided.

We consider the corresponding generating functional $\mathcal{Z}(J, K)$:

$$\mathcal{Z}(J, K) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi, K) + \int \mathrm{d}x\, J_i(x)\phi_i(x)\right]. \tag{13.55}$$

An infinitesimal change of variable (13.10) leads to

$$0 = \int [\mathrm{d}\phi]\, \delta_\alpha \left[-\mathcal{S}(\phi, K) + \int \mathrm{d}x\, J_i\phi_i\right] \exp\left[-\mathcal{S}(\phi, K) + \int \mathrm{d}x\, J_i\phi_i\right] \tag{13.56}$$

with

$$\delta_\alpha \left[\mathcal{S}(\phi, K) - \int \mathrm{d}x\, J_i\phi_i\right] = \int \mathrm{d}x\, t^\alpha_{ij}\phi_j(x)\left[K_{ik}(x)\phi_k(x) - J_i(x)\right]. \tag{13.57}$$

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 $\delta^2/\delta J_j(x)\delta/\delta J_k(x)$.

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 $\delta\mathbf{K}$,

$$F(\mathbf{K} + \delta\mathbf{K}) - F(\mathbf{K}) = \int \frac{\delta F}{\delta K_{ij}(x)}\,\delta K_{ij}(x)\mathrm{d}x + O(\|\delta\mathbf{K}\|^2). \tag{13.58}$$

The derivative is then defined in the sense of differential geometry: it is the linear operator acting on $\delta K_{ij}(x)$ in the r.h.s. of equation (13.58).

In the present example, differentiation with respect to K_{ij} generates the traceless part of the product $\phi_i\phi_j$. Since the equation (13.57) involves only the traceless part of the same product, the definition (13.58) allows us to rewrite equation (13.56):

$$\int \mathrm{d}x \left\{ \left[t^\alpha_{ij}K_{ik}(x) + t^\alpha_{ik}K_{ij}(x)\right] \frac{\delta}{\delta K_{kj}(x)} - J_i(x)t^\alpha_{ij}\frac{\delta}{\delta J_j(x)} \right\} \mathcal{Z}(J, K) = 0\,. \tag{13.59}$$

Because equation (13.59) is a first order 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:

$$\Gamma(\varphi, K) + \mathcal{W}(J, K) = \int \mathrm{d}x\, J_i(x)\varphi_i(x), \qquad \varphi_i(x) = \frac{\delta\mathcal{W}(J, K)}{\delta J_i(x)}\,.$$

The sources $K_{ij}(x)$ do not participate in the Legendre transformation and have to be considered as external parameters. This, then, immediately implies (equation (7.73)):

$$\left.\frac{\delta \mathcal{W}}{\delta K_{ij}(x)}\right|_J = -\left.\frac{\delta \Gamma}{\delta K_{ij}(x)}\right|_\varphi . \tag{13.60}$$

This 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 (13.59) is simple and leads, for $\Gamma(\varphi, K)$, to

$$\int \mathrm{d}x \left\{ \left[t^\alpha_{ij} K_{ik}(x) + t^\alpha_{ik} K_{ij}(x)\right] \frac{\delta \Gamma(\varphi, K)}{\delta K_{kj}(x)} + \varphi_i(x) t^\alpha_{ij} \frac{\delta \Gamma(\varphi, K)}{\delta \varphi_j(x)} \right\} = 0 . \tag{13.61}$$

The equation has a straightforward interpretation: $\Gamma(\varphi, K)$ is invariant under the double transformation

$$\begin{aligned} \delta \varphi_i(x) &= t^\alpha_{ij} \omega_\alpha \varphi_j(x), \\ \delta K_{kj}(x) &= \omega_\alpha \left[t^\alpha_{ji} K_{ik}(x) + t^\alpha_{ki} K_{ij}(x)\right] . \end{aligned} \tag{13.62}$$

By performing a transformation both on $\boldsymbol{\phi}$ and $\mathbf{K}$, we have rendered the breaking term $\frac{1}{2} \int \mathrm{d}x\, K_{ij}(x) \phi_i(x) \phi_j(x)$ group invariant. It is then almost obvious that the symmetry of $\mathcal{S}(\phi, K)$ under the equivalent of the transformations (13.62) implies the WT identities (13.61).

Using the arguments given for the case $\mathbf{K} = 0$, we can show that if the regularized functional $\Gamma(\varphi, K)$ satisfies (13.61), the renormalized functional $\Gamma(\varphi, K)$ and the renormalized action will satisfy the same identity: the renormalized action $\mathcal{S}_\mathrm{r}(\phi, K)$ is the most general local functional of ϕ and K compatible with power counting and is invariant under the group transformation (13.62).

For a ϕ^4-like field theory in four dimensions the action is the integral of a local function of dimension four, the field ϕ has dimension $[\phi] = 1$ and the dimension $[K]$ of the source $\mathbf{K}$ is two because it is coupled to an operator of dimension 2 (see Section 12.1):

$$[\phi] = 1 , \qquad [K] = 2 , \qquad [\mathcal{S}(\phi, K)] = 4 .$$

The renormalized action $\mathcal{S}_\mathrm{r}(\phi, K)$ thus has the general form

$$\begin{aligned} \mathcal{S}_\mathrm{r}(\phi, K) = [\mathcal{S}_\mathrm{sym}]_\mathrm{r}(\phi) &+ \frac{1}{2} \int \mathrm{d}x\, K_{ij}(x) A_{ij}(\phi(x)) \\ &+ \frac{1}{2} b_{ij,kl} \int \mathrm{d}x\, K_{ij}(x) K_{kl}(x), \end{aligned} \tag{13.63}$$

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 (13.62). In the simplest case, $A_{ij}(\phi)$ has the form

$$A_{ij}(\phi(x)) = Z_2(\phi_i(x)\phi_j(x) - \frac{1}{N} \delta_{ij} \phi^2(x)), \tag{13.64}$$

in which N is the number of components of $\boldsymbol{\phi}$.

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.

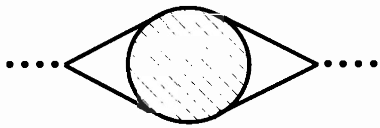


Fig. 13.3

The term quadratic in K_{ij} does not depend on ϕ and can be factorized in front of the functional 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 figure 13.3.

The renormalized action for the case of quadratic symmetry breaking is then obtained by setting $K_{ij}(x)$ to be a constant:

$$K_{ij}(x) = \mu_{ij}\,.$$

We see from the analysis of this more complex situation 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. The renormalized theory may, therefore, depend on more parameters than one would have naively anticipated when adding the breaking term.

Finally, let us note that, as in the case of the linear symmetry breaking, one can use the WT identities (13.61) 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 (13.61) by expanding in $\varphi(x)$ and $K(x)$, in which they appear.

Breaking terms of higher canonical dimensions. The preceding analysis can be easily 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 zero. 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.

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 theory finite renormalizes the theory 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. This question has been examined in Section 12.1. 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.

13.6 Chiral Symmetry Breaking in Strong Interactions

One of the striking feature of Strong Interactions in low energy Particle Physics is the observation of approximate spontaneously broken $SU(N) \times SU(N)$ chiral symmetries, which manifest themselves in particular in the small masses of the pseudoscalar mesons. In particular the π-meson is specially light, an indication that the breaking of the $SU(2) \times SU(2)$ symmetry is quite small. With our present understanding, this property is a consequence of the small masses of the **u** and **d** quarks (see for example Section 8.2.3) and the vector-like coupling of quarks to gluons. The mass of the **s** quark and thus the explicit breaking of the $SU(3) \times SU(3)$ symmetry are larger as can be seen from the masses of the K and η pseudoscalar mesons.

Note that, according to our preceding analysis, since a fermion mass operator in a renormalizable field theory 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 34). Progress is slow though quite encouraging. The most serious technical difficulties are related to the dynamics of quarks.

Here, we explain instead how one can construct effective low energy theories based on observed hadrons like protons, neutrons, π-mesons... 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. We, therefore, face the situation we have discussed at some length in Section 13.3.

13.6.1 The chiral symmetry: general structure

We first consider the action for N free massless Dirac fermions in even dimensions:

$$\mathcal{S}(\psi,\bar{\psi}) = -\int \mathrm{d}^d x\, \bar{\psi}_i(x) \not{\partial} \psi_i(x). \tag{13.65}$$

It has a chiral $U(N) \times U(N)$ symmetry corresponding to the transformations

$$\boldsymbol{\psi}' = \left[\tfrac{1}{2}(\mathbf{1}+\gamma_S)\mathbf{U}_+ + \tfrac{1}{2}(\mathbf{1}-\gamma_S)\mathbf{U}_-\right]\boldsymbol{\psi}\,, \tag{13.66}$$

$$\bar{\boldsymbol{\psi}}' = \bar{\boldsymbol{\psi}}\left[\tfrac{1}{2}(\mathbf{1}+\gamma_S)\mathbf{U}_-^\dagger + \tfrac{1}{2}(\mathbf{1}-\gamma_S)\mathbf{U}_+^\dagger\right], \tag{13.67}$$

where $\mathbf{U}_\pm$ are two $N \times N$ unitary matrices corresponding to the two $U(N)$ groups.

We now couple the fermions to scalar bosons forming a complex $N \times N$ matrix $\mathbf{M}(x)$. One verifies that the interaction term,

$$-g\int \mathrm{d}^d x\, \bar{\psi}_i \left[\tfrac{1}{2}(\mathbf{1}+\gamma_S)M_{ij} + \tfrac{1}{2}(\mathbf{1}-\gamma_S)M_{ij}^\dagger\right]\psi_j\,, \tag{13.68}$$

is invariant under the transformations (13.66,13.67) provided the matrix $\mathbf{M}$ transforms like

$$\mathbf{M}' = \mathbf{U}_-\mathbf{M}\mathbf{U}_+^\dagger\,. \tag{13.69}$$

The total action also satisfies reflection hermiticity as defined in Section 8.2.1. It can be made invariant under a space reflection P (Section 8.2.2) if $\mathbf{M}$ transforms like

$$\mathbf{M}_P(x) = \mathbf{M}^\dagger(\tilde{x}), \tag{13.70}$$

in which $\tilde{x}$ is obtained from x by changing the sign of one component. Therefore, the matrix $\mathbf{\Sigma} = (\mathbf{M}+\mathbf{M}^\dagger)/\sqrt{2}$ represents a set of scalar fields and $\mathbf{\Pi} = (\mathbf{M}-\mathbf{M}^\dagger)/\sqrt{2}$ a set of pseudoscalar fields. Finally, under a charge conjugation C, $\mathbf{M}$ transforms like

$$\mathbf{M}_C = \mathbf{M}^*, \quad \text{if } d/2 \text{ is odd}, \text{ and } \quad \mathbf{M}_C = {}^T\mathbf{M}, \quad \text{if } d/2 \text{ is even}. \tag{13.71}$$

A possible action for the boson fields symmetric under $U(N)\times U(N)$ transformations is then:

$$\mathcal{S}(\mathbf{M}) = \int \mathrm{d}^d x \,\operatorname{tr}\left(\partial_\mu \mathbf{M}\partial_\mu \mathbf{M}^\dagger + V(\mathbf{M}\mathbf{M}^\dagger)\right), \tag{13.72}$$

where $V(\varphi)$ is a polynomial of the matrix φ. If in addition we add a term proportional to $\det \mathbf{M} + \det \mathbf{M}^\dagger$, we reduce 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 (13.70) and (13.71), is

$$\mathcal{S}_B(\mathbf{M}) = -\frac{1}{\sqrt{2}}\int \mathrm{d}^d x \,\operatorname{tr}\mathbf{C}\left(\mathbf{M}+\mathbf{M}^\dagger\right), \tag{13.73}$$

in which $\mathbf{C}$ is a hermitian matrix:

$$\mathbf{C} = \mathbf{C}^\dagger .$$

To the transformations (13.66,13.67) and (13.69) correspond two currents (for more details see Appendix 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 $U(N)\times U(N)$ ($\mathbf{U}_+ = \mathbf{U}_-$) which conserves parity:

$$-iV_\mu^\alpha(x) = -\bar{\psi}t^\alpha\gamma_\mu\psi + \operatorname{tr} t^\alpha \left\{[\partial_\mu\mathbf{M}^\dagger, \mathbf{M}] + [\partial_\mu\mathbf{M}, \mathbf{M}^\dagger]\right\}, \tag{13.74}$$

and the axial current $\mathbf{A}_\mu(x)$ associated with the complementary set of generators in the Lie algebra, that is, $\mathcal{L}$ $(U(N)\times U(N)/U(N))$:

$$-iA_\mu^\alpha(x) = -\bar{\psi}t^\alpha\gamma_S\gamma_\mu\psi + \operatorname{tr} t^\alpha \left\{[\partial_\mu\mathbf{M}^\dagger, \mathbf{M}]_+ + [\partial_\mu\mathbf{M}, \mathbf{M}^\dagger]_+\right\}. \tag{13.75}$$

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:

$$\partial_\mu V_\mu^\alpha(x) = -i\operatorname{tr}\{[t^\alpha, \mathbf{C}]\,\mathbf{\Sigma}\}, \tag{13.76}$$

$$\partial_\mu A_\mu^\alpha(x) = \operatorname{tr}\left\{[t^\alpha, \mathbf{C}]_+\,\mathbf{\Pi}\right\}. \tag{13.77}$$

in which $\tilde{x}$ is obtained from x by changing the sign of one component. Therefore, the matrix $\mathbf{\Sigma} = (\mathbf{M}+\mathbf{M}^\dagger)/\sqrt{2}$ represents a set of scalar fields and $\mathbf{\Pi} = (\mathbf{M}-\mathbf{M}^\dagger)/\sqrt{2}$ a set of pseudoscalar fields. Finally, under a charge conjugation C, $\mathbf{M}$ transforms like

$$\mathbf{M}_C = \mathbf{M}^*, \quad \text{if } d/2 \text{ is odd}, \text{ and } \quad \mathbf{M}_C = {}^T\mathbf{M}, \quad \text{if } d/2 \text{ is even}. \tag{13.71}$$

A possible action for the boson fields symmetric under $U(N) \times U(N)$ transformations is then:

$$\mathcal{S}(\mathbf{M}) = \int \mathrm{d}^d x \; \mathrm{tr}\left(\partial_\mu \mathbf{M} \partial_\mu \mathbf{M}^\dagger + V(\mathbf{M}\mathbf{M}^\dagger)\right), \tag{13.72}$$

where $V(\varphi)$ is a polynomial of the matrix φ. If in addition we add a term proportional to $\det \mathbf{M} + \det \mathbf{M}^\dagger$, we reduce 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 (13.70) and (13.71), is

$$\mathcal{S}_B(\mathbf{M}) = -\frac{1}{\sqrt{2}} \int \mathrm{d}^d x \; \mathrm{tr}\,\mathbf{C}\left(\mathbf{M}+\mathbf{M}^\dagger\right), \tag{13.73}$$

in which $\mathbf{C}$ is a hermitian matrix:

$$\mathbf{C} = \mathbf{C}^\dagger .$$

To the transformations (13.66,13.67) and (13.69) correspond two currents (for more details see Appendix 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 $U(N) \times U(N)$ ($\mathbf{U}_+ = \mathbf{U}_-$) which conserves parity:

$$-iV_\mu^\alpha(x) = -\bar{\psi} t^\alpha \gamma_\mu \psi + \mathrm{tr}\, t^\alpha \left\{ [\partial_\mu \mathbf{M}^\dagger, \mathbf{M}] + [\partial_\mu \mathbf{M}, \mathbf{M}^\dagger] \right\}, \tag{13.74}$$

and the axial current $\mathbf{A}_\mu(x)$ associated with the complementary set of generators in the Lie algebra, that is, $\mathcal{L}\ (U(N) \times U(N)/U(N))$:

$$-iA_\mu^\alpha(x) = -\bar{\psi} t^\alpha \gamma_S \gamma_\mu \psi + \mathrm{tr}\, t^\alpha \left\{ [\partial_\mu \mathbf{M}^\dagger, \mathbf{M}]_+ + [\partial_\mu \mathbf{M}, \mathbf{M}^\dagger]_+ \right\}. \tag{13.75}$$

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:

$$\partial_\mu V_\mu^\alpha(x) = -i\, \mathrm{tr}\left\{ [t^\alpha, \mathbf{C}]\, \mathbf{\Sigma} \right\}, \tag{13.76}$$

$$\partial_\mu A_\mu^\alpha(x) = \mathrm{tr}\left\{ [t^\alpha, \mathbf{C}]_+\, \mathbf{\Pi} \right\}. \tag{13.77}$$

13.6.2 A special case: the linear σ-model

The case $N = 2$ is of particular interest because the pion mass is specially small, and thus the explicit breaking of chiral symmetry small.

Previous analysis leads to a theory with eight real boson 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 Appendix A8.1.4, to eliminate possible confusion with the traditional notation for fields). Therefore, $\mathbf{M}$ and $\tau_2 \mathbf{M}^* \tau_2$ have the same transformation law. The representation can be reduced and the matrix $\mathbf{M}$ parametrized in terms of two fields $\sigma(x)$ and $\boldsymbol{\pi}(x)$ in the form

$$\mathbf{M} = \tau_2 \mathbf{M}^* \tau_2 \equiv \frac{1}{\sqrt{2}}(\sigma + i\boldsymbol{\tau}\cdot\boldsymbol{\pi}) = \frac{1}{\sqrt{2}}\begin{bmatrix} \sigma + i\pi_0 & \pi_2 + i\pi_1 \\ -\pi_2 + i\pi_1 & \sigma - i\pi_0 \end{bmatrix}. \tag{13.78}$$

The group $SU(2)\times SU(2)$ is the covering group of $O(4)$ which is also the symmetry group of the bosonic 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 4-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 $\sigma(x)$. The action can then be written as

$$\mathcal{S} = \int \mathrm{d}^d x \Big\{ -\bar{N}(x)\left[\not{\partial} + g(\sigma + i\gamma_5 \boldsymbol{\tau}\cdot\boldsymbol{\pi})\right] N(x) + \tfrac{1}{2}\left((\partial_\mu\sigma)^2 + (\partial_\mu\boldsymbol{\pi})^2\right) + V(\sigma^2+\boldsymbol{\pi}^2) - c\sigma \Big\} \tag{13.79}$$

with

$$V(\rho) = \frac{1}{2}m^2\rho + \frac{1}{4!}\lambda\rho^2. \tag{13.80}$$

The fermion doublet $N(x)$ is identified with the two nucleon fields, proton and neutron.

The action (13.79) has an exact $SU(2)\times U(1)$ symmetry, to which corresponds the conservation of the vector current, and implements the idea of *partially conserved axial current* (PCAC) for $SU(2)$. In the standard normalization, which differs by a factor 2 from definition (13.75) (see equations (13.88,13.100)),

$$\partial_\mu \mathbf{A}_\mu(x) = c\boldsymbol{\pi}(x). \tag{13.81}$$

Finally, it follows from equation (13.70) that $\sigma(x)$ is a scalar field and $\boldsymbol{\pi}(x)$ a pseudoscalar field (the pi-meson). 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.

13.6.3 Tree approximation

Bosonic sector. We discuss the pattern of symmetry breaking in the classical approximation. Furthermore, we consider only the case $N = 2$ because it is the simplest and physically the most important. In the absence of fermions we simply have the ϕ^4 field theory with $O(4)$ symmetry. Equation (13.27) gives the relation between the expectation value v of the field σ and the symmetry breaking parameter c in the classical approximation:

$$v(m^2 + \lambda v^2/6) = c\,. \tag{13.82}$$

Setting

$$\sigma(x) = v + s(x),$$

in action (13.79), we read off the masses of the π and σ particles at the tree order:

$$m_\pi^2 = m^2 + \lambda v^2/6\,, \qquad m_\sigma^2 = m^2 + \lambda v^2/2\,. \tag{13.83}$$

The hypothesis which accounts for the success of PCAC phenomenology is that the explicit symmetry breaking term is small and one is close to a situation of SSB. For the model (13.79) this means in particular that m_π is small compared to m_σ. With this hypothesis it is possible to predict some general features of the low energy π–π scattering. Introducing the standard invariant variables

$$s = -(p_1 + p_2)^2\ , \quad t = -(p_1 + p_3)^2\ , \quad u = -(p_1 + p_4)^2, \tag{13.84}$$

we can write the connected amputated π-field four-point function at this order:

$$\begin{aligned}\left[W^{(4)}_{ijkl}\right]_{\text{amp}} &= \frac{s - m_\pi^2}{v^2}\frac{m_\sigma^2 - m_\pi^2}{m_\sigma^2 - s}\delta_{ij}\delta_{kl} + \frac{t - m_\pi^2}{v^2}\frac{m_\sigma^2 - m_\pi^2}{m_\sigma^2 - t}\delta_{ik}\delta_{jl} \\ &\quad + \frac{u - m_\pi^2}{v^2}\frac{m_\sigma^2 - m_\pi^2}{m_\sigma^2 - u}\delta_{il}\delta_{jk}\,. \end{aligned}\tag{13.85}$$

We have used the relations (13.83) to eliminate m and λ. The physical scattering amplitude is obtained by setting all momenta on the mass shell: $p_i^2 = -m_\pi^2$ and then $s + t + u = 4m_\pi^2$.

The expectation value v is experimentally accessible from the weak π-meson decay as a consequence of relation (13.81) and is denoted traditionally by f_π. Since m_σ is supposed to be large compared to m_π, the expression (13.85) 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. Although the π–π scattering amplitude, of course, cannot be measured directly, indirect methods provide an experimental confirmation of the resulting pattern.

Fermion sector. In the unbroken phase, the mass of the fermion vanishes for $c = 0$. The largest contribution to the fermion mass m_N is thus generated by the Yukawa coupling and the σ expectation value

$$m_\psi = gv\,.$$

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 potential due to π exchange. We then have the relation between physical quantities:

$$g_{\pi NN} = m_N/f_\pi\,. \tag{13.86}$$

This relation is the Goldberger–Treiman relation in the tree approximation and agrees semi-quantitatively with experiment since

$$g_{\pi NN} = 13.6\,, \quad \frac{m_N}{f_\pi} \simeq \frac{939.}{93.3} = 10.\,. \tag{13.87}$$

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; it agrees well with experimental data.

Beyond the tree approximation. Since the field theory model is renormalizable, it is possible to calculate loop corrections. Then, several problems arise. First, there is a question of principle. As we shall argue later, the ϕ^4 field theory, as well as the theory (13.79) with fermions, is most likely inconsistent in four dimensions for non-vanishing coupling (see Chapters 25–30,35). More precisely, although the theory 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 const./$\ln(\Lambda/\mu)$—this is the ***triviality problem.*** 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. Still the loop corrections may be useful to improve the tree level amplitudes from the point of view of unitarity at low energy.

Second, from the computational point of view several difficulties are encountered.

(i) Loop corrections become large at moderate energies. For example in π–π scattering one encounters the ρ resonance. Then it becomes necessary to apply a summation method to the perturbation series. Calculations have been performed using the method of Padé approximants.

(ii) Since the σ mass is larger than $2m_\pi$, the σ particle is unstable (it 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 unphysical 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, but the validity of the expansion is then limited to energies smaller than $4m_\sigma^2$, that is, very 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 (13.79), only limited results have been obtained beyond the simple predictions which rely on the geometry of the model and are, therefore, mostly contained in WT identities as we explain below.

13.6.4 Ward–Takahashi identities

We have described the difficulties one encounters when one tries to derive consequences from a phenomenological chiral action. However, some relations are valid beyond perturbation theory: the WT identities which are direct consequences of the broken symmetry.

Unfortunately, equation (13.32) shows that the WT identities always involve correlation functions with one π-field at zero momentum. Therefore, they would lead to relations between observables only if the π-meson 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. This extrapolation is model-dependent, and the results are only reliable if the predictions at zero pion mass are already in qualitative agreement with experiment. Let us again first discuss the purely bosonic sector.

Boson sector. The non-trivial part of the WT identities corresponds to the transformations

$$\delta\boldsymbol{\pi}(x) = -\boldsymbol{\omega}\sigma(x)\,, \quad \delta\sigma(x) = \boldsymbol{\omega}\cdot\boldsymbol{\pi}(x). \tag{13.88}$$

Calling $\mathbf{J}(x)$ the source for the π-field and $H(x)$ the source for the σ-field, we can write the WT identities for the generating functional of connected correlation functions $\mathcal{W}(\mathbf{J}, H)$:

$$\int \mathrm{d}x \left[J_i(x)\frac{\delta}{\delta H(x)} - \big(c + H(x)\big)\frac{\delta}{\delta J_i(x)}\right]\mathcal{W} = 0\,. \tag{13.89}$$

It is convenient to introduce some additional notation to take into account the residual $O(3)$ symmetry. We set

$$\begin{aligned}
W^{(2)}_{ij}(p) &= \delta_{ij}D_\pi(p),\\
W^{(2)}(p) &= D_\sigma(p),\\
\tilde{W}^{(3)}_{ij}(p_1,p_2;p_3) &= \delta_{ij}D_\pi(p_1)D_\pi(p_2)D_\sigma(p_3)C(p_1,p_2;p_3),\\
\left[\tilde{W}^{(4)}_{ijkl}(p_1,p_2,p_3,p_4)\right]_{\mathrm{amp}} &= \delta_{ij}\delta_{kl}A(p_1,p_2,p_3,p_4) + \delta_{ik}\delta_{jl}A(p_1,p_3,p_2,p_4)\\
&\quad + \delta_{il}\delta_{kj}A(p_1,p_4,p_3,p_2),
\end{aligned}$$

with the conventions that indices correspond to π-fields, and in mixed π–σ correlation functions the arguments of the π-fields are written first.

Differentiating with respect to J_j, and setting the sources to zero, we obtain the equivalent of equation (13.44):

$$v = \langle\sigma\rangle = cD_\pi(0) \equiv c/\mu^2, \tag{13.90}$$

where we have denoted by μ^2 the value of the inverse of the π propagator at zero momentum which is now different from the pion mass squared m_π^2.

Differentiating once with respect to J_j and H, we obtain

$$\delta_{ij}\tilde{W}^{(2)}(p) - \tilde{W}^{(2)}_{ij}(p) = c\tilde{W}^{(3)}_{ij}(0,p;-p),$$

and thus using equation (13.90)

$$D_\pi^{-1}(p) - D_\sigma^{-1}(p) = vC(0,p;-p). \tag{13.91}$$

Setting $p=0$, and now denoting by m_σ^2 the value of the inverse σ propagator at zero momentum,

$$m_\sigma^2 = D_\sigma^{-1}(0), \tag{13.92}$$

we get, in particular,

$$\mu^2 - m_\sigma^2 = vC(0,0;0). \tag{13.93}$$

Differentiating thrice with respect to J we obtain a relation between three- and four-point correlation functions:

$$c\tilde{W}^{(4)}_{ijkl}(0,p_2,p_3,p_4) = \delta_{ij}\tilde{W}^{(3)}_{kl}(p_3,p_4;p_2) + \ 2 \text{ terms}. \tag{13.94}$$

It follows that

$$vA(0,p_2,p_3,p_4) = C(p_3,p_4;p_2)D_\sigma(p_2)D_\pi^{-1}(p_2). \tag{13.95}$$

First, for $p_2^2 = -m_\pi^2$, the equation reduces to Adler's consistency condition

$$A(0,p_2(p_2^2 = -m_\pi^2),p_3,p_4) = 0\,. \tag{13.96}$$

Moreover, setting $p_3 = 0$ in (13.95) and eliminating the function C between (13.91) and (13.95) we find

$$v^2 A(0,p,0,-p) = D_\pi^{-1}(p)\left[D_\sigma(p)D_\pi^{-1}(p) - 1\right]. \tag{13.97}$$

The first term in the r.h.s. has a double zero at the pion mass. Therefore, taking the derivative with respect to p^2 we recover Weinberg's relation:

$$v^2\,\frac{\partial}{\partial p^2}\left(A(0,p,0,-p) + D_\pi^{-1}(p)\right)\Big|_{p^2=-m_\pi^2} = 0\,. \tag{13.98}$$

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 (13.85) satisfies both conditions (13.96,13.98).

Another constraint on the π–π scattering amplitude is obtained, for example, by setting all momenta to zero in (13.97):

$$v^2 A(0,0,0,0) = \mu^2(\mu^2/m_\sigma^2 - 1). \tag{13.99}$$

This equation, however, involves an independent free parameter m_σ. Again one verifies that expression (13.85) satisfies equation (13.99) in the tree approximation.

Fermion sector. The infinitesimal transformations of the fermion fields, corresponding to the equations (13.88) are

$$\delta\psi = \tfrac{1}{2}i\gamma_S\boldsymbol{\tau}\cdot\boldsymbol{\omega}\psi\,,\qquad \delta\bar\psi = \tfrac{1}{2}i\bar\psi\gamma_S\boldsymbol{\tau}\cdot\boldsymbol{\omega}\,. \tag{13.100}$$

We denote by $\bar\eta$ and η the sources for the fermion fields. The generating functional $\mathcal{W}(\eta,\bar\eta,J,H)$ of connected correlation functions then satisfies the WT identity:

$$\int \mathrm{d}x \left\{\frac{i}{2}\left[\bar\eta(x)\gamma_S\boldsymbol{\tau}\frac{\delta}{\delta\bar\eta(x)} - \eta(x)\gamma_S\boldsymbol{\tau}\frac{\delta}{\delta\eta(x)}\right] - \mathbf{J}(x)\frac{\delta}{\delta H(x)} \right.$$
$$\left. + (H(x)+c)\frac{\delta}{\delta\mathbf{J}(x)}\right\}\mathcal{W}(\eta,\bar\eta,J,H) = 0\,. \tag{13.101}$$

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 1PI functions:

$$v\tilde{\Gamma}^{(3)}_{\pi NN}(0;p,-p) = \frac{i\tau}{2}\left\{\gamma_S, \tilde{\Gamma}^{(2)}_{NN}(p)\right\}_+ . \qquad (13.102)$$

The index + in the r.h.s. 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 (13.86). It has a physical interpretation in terms of the weak current under the name of Goldberger–Treiman relation. The r.h.s. is known when the nucleons are on mass-shell. The l.h.s. can be approximately 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 (13.81), 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

$$\frac{G_A}{G_V} \simeq g_{\pi NN}\frac{f_\pi}{m_N} . \qquad (13.103)$$

Replacing by experimental numbers one finds 1.22 for the l.h.s. and 1.36 for the r.h.s., a notable improvement over the tree approximation (13.87).

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 (13.96)) involving amplitudes for the emission of one unphysical pion at zero momentum. To determine completely the low energy π–N scattering amplitude, it is, however, 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 well with experimental results.

Bibliographical Notes

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 and G. Jona-Lasinio, *Phys. Rev.* 122 (1961) 345; 124 (1961) 246; J. Goldstone, A. Salam and S. Weinberg, *Phys. Rev.* 127 (1962) 965; S. Bludman and A. Klein, *Phys. Rev.* 131 (1963) 2364; G. Jona-Lasinio, *Nuovo Cimento* 34 (1964) 1790.

For a general survey

G.S. Guralnik, C.R. Hagen and T.W. Kibble, *Advances in Physics*, R.E. Marshak and R. Cool eds. (Interscience, New York 1968).

The Proceedings of the Cargèse Summer Institute 1970 contains several contributions devoted to breaking of symmetries, WT identities and renormalizations

Cargèse Lectures in Physics vol. 5, D. Bessis ed. (Gordon and Breach, New York 1971).

Among them let us quote more specifically

K. Symanzik, who gives a general discussion of renormalization of symmetry breaking terms; A.N. Vasil'ev (in French) who shows how WT identities can be derived from

the functional integral representation of correlation functions; B.W. Lee who discusses Chiral Dynamics.

These authors have been a direct source of inspiration for the presentation we have given here, which is an update of

J. Zinn-Justin in *Trends in Elementary Particle Physics (Lectures Notes in Physics 37)*, Bonn 1974, H. Rollnik and K. Dietz eds. (Springer-Verlag, Berlin 1975).

See also the lectures of Coleman

S. Coleman in *Laws of Hadronic Matter*, Erice 1973, A. Zichichi ed. (Academic Press, New York 1975).

The linear σ-model has been proposed in

M. Gell-Mann and M. Lévy, *Nuovo Cimento* 16 (1960) 705.

Its properties have been studied in

B.W. Lee, *Nucl. Phys.* B9 (1969) 649; J.-L. Gervais and B.W. Lee, *Nucl. Phys.* B12 (1969) 627; J.L. Basdevant and B.W. Lee, *Phys. Rev.* D2 (1970) 1680.

This work is reviewed in

B.W. Lee, *Chiral Dynamics* (Gordon and Breach, New York 1972).

For a review of the application of Padé approximants to the summation of perturbation theory for phenomenological lagrangians see, for example,

J. Zinn-Justin, *Phys. Rep.* 1C (1970) 55.

$SU(2) \times SU(2)$ symmetry breaking has been discussed in

S. Weinberg, *Phys. Rev.* 166 (1968) 1568.

Chiral Dynamics is directly connected to Current Algebra. Among the large number of articles and text books let us just quote

S. Adler and R. Dashen, *Current Algebras and Applications to Particle Physics* (Benjamin, New York, 1968); S. Weinberg in *Lectures on Elementary Particle Physics and Quantum Field Theory*, Brandeis 1970, S. Deser, M. Grisaru and H. Pendleton eds. (MIT Press, Cambridge 1970); S.B. Treiman, R. Jackiw and D.J. Gross, *Lectures on Current Algebra and its Applications* (Princeton University Press, Princeton 1972); C. Itzykson and J.-B. Zuber's text book as quoted in the introduction; M.L. Goldberger and S.B. Treiman, *Phys. Rev.* 110 (1958) 1178; S.L. Adler, *Phys. Rev.* 139B (1965) 1638; 140B (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 and B. Renner, *Phys. Rev.* 175 (1968) 2195.

For the improved energy–momentum tensor (Appendix A13.3) see

C.G. Callan, S. Coleman and R. Jackiw, *Ann. Phys. (NY)* 59 (1970) 42.

It has been noticed that at zero of the RG β-function both dilatation and conformal invariance are restored in

B. Schroer, *Lett. Nuovo Cimento* 2 (1971) 867.

Conformal invariance and its spectacular consequences in two dimensions are thoroughly discussed in the Proceedings of Les Houches Summer Institute 1988 *Fields, Strings and Critical Phenomena,* E. Brézin and J. Zinn-Justin eds. (Elsevier, Amsterdam 1989). See in particular the lectures by

J. Cardy, *Conformal Invariance and Statistical Mechanics*; P. Ginsparg, *Applied Conformal Field Theory.*

APPENDIX A13
CURRENTS AND NOETHER'S THEOREM

In this appendix, in order to discuss properties of the relativistic classical equations of motion, we adopt the covariant notation of real time field theory with a metric tensor $g_{\mu\nu}$ and a metric of signature $(+ - - \cdots)$. The formal transition between euclidean and real time will be achieved by setting $x_d = ix_0 \equiv it$. Summation over successive upper and lower indices will be implied (see Chapter 22).

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 $\mathcal{S}$,

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \, \mathcal{L}\big(\phi(x), \partial_\mu \phi(x)\big), \tag{A13.1}$$

is

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial[\partial_\mu \phi(x)]} - \frac{\partial \mathcal{L}}{\partial \phi(x)} = 0 \tag{A13.2}$$

(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:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial[\partial_\mu \Lambda(x)]} - \frac{\partial \mathcal{L}}{\partial \Lambda(x)} = 0\,. \tag{A13.3}$$

We define a current $J^\mu(x)$, functional of $\phi(x)$, by

$$J^\mu(x) = \left.\frac{\partial \mathcal{L}}{\partial[\partial_\mu \Lambda(x)]}\right|_{\Lambda(x)=0}, \tag{A13.4}$$

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 (*A*13.3) as

$$\partial_\mu J^\mu(x) = \frac{\partial \mathcal{L}}{\partial \Lambda(x)}, \tag{A13.5}$$

which is Noether's theorem.

If, in addition, the lagrangian is invariant under space-independent group transformations, $\partial\mathcal{L}/\partial\Lambda$ vanishes and thus the current J_μ is conserved:

$$\partial_\mu J^\mu(x) = 0\,. \tag{A13.6}$$

In classical field theory the space integral of the time-component of the current is a charge $\mathcal{Q}^{\alpha}(t \equiv x_0)$:

$$\mathcal{Q}^{\alpha}(t) = \int \mathrm{d}^{d-1}x \, J_0^{\alpha}(x). \qquad (A13.7)$$

By differentiating with respect to t and using the current conservation equation (A13.6) one finds

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}^{\alpha}(t) = \int \mathrm{d}^{d-1}x \sum_{\mu=1}^{d-1} \partial_{\mu} J_{\mu}(x) = 0 \, .$$

The charges $\mathcal{Q}^{\alpha}(t)$ are constants of the classical motion.

Example. If the lagrangian density has the form

$$\mathcal{L} = \tfrac{1}{2}\partial_{\mu}\phi_i \partial^{\mu}\phi_i - V\left[\phi(x)\right] \qquad (A13.8)$$

(in real time covariant notation) and if the infinitesimal group transformations are

$$\delta\phi_i(x) = t^{\alpha}_{ij}\Lambda^{\alpha}(x)\phi_j(x), \qquad (A13.9)$$

the current $J^{\alpha}_{\mu}(x)$ is given by

$$J^{\alpha}_{\mu}(x) = t^{\alpha}_{ij}\partial_{\mu}\phi_i(x)\phi_j(x). \qquad (A13.10)$$

A13.2 Euclidean Quantum Field Theory

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 will appear either in the coupling at leading order of matter to gauge fields (see Chapters 18,19) or as polynomials in the fields (operators in the sense of Chapter 12) satisfying some identities which we will derive and, therefore, having special renormalization properties.

We, therefore, consider the generating functional $\mathcal{Z}(J)$,

$$\mathcal{Z}(J) = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi) + \int J_i(x)\phi_i(x)\mathrm{d}x\right], \qquad (A13.11)$$

in which the action is invariant under group transformations whose infinitesimal form is given by equation (A13.9) when $\Lambda(x)$ is a constant.

In what follows *dimensional* regularization is assumed.

We now perform a change of variables in the integral (A13.11) of the form of a transformation (13.10). We define the euclidean current $J^{\alpha}_{\mu}(x)$ by equation (A13.4) in terms of the euclidean action density. If $\mathcal{S}(\phi)$ is symmetric, the variation of the action reads

$$\delta\mathcal{S}(\phi) = \int \partial_{\mu}\Lambda^{\alpha}(x) J^{\alpha}_{\mu}(x)\mathrm{d}x \, . \qquad (A13.12)$$

Identifying the coefficient of $\Lambda^{\alpha}(x)$, we obtain

$$\int [\mathrm{d}\phi] \left[\partial_{\mu}J^{\alpha}_{\mu}(x) - J_i(x)t^{\alpha}_{ij}\phi_j(x)\right] \exp\left[-\mathcal{S}(\phi) + \int J_i(x)\phi_i(x)\mathrm{d}x\right] = 0 \, . \qquad (A13.13)$$

This identity can be written as

$$\partial_\mu^x \mathcal{Z}_{J_\mu^\alpha(x)} = J_i(x) t_{ij}^\alpha \frac{\delta \mathcal{Z}}{\delta J_j(x)}, \qquad (A13.14)$$

where $\mathcal{Z}_{J_\mu^\alpha(x)}$ is the generating functional of correlation functions with a $J_\mu^\alpha(x)$ operator insertion.

The same equation is valid for connected correlation functions. After Legendre transformation we find

$$\partial_\mu^x \Gamma_{J_\mu^\alpha(x)} = -\frac{\delta \Gamma}{\delta \varphi_i(x)} t_{ij}^\alpha \varphi_j(x). \qquad (A13.15)$$

Equations (A13.14,A13.15) are the analogues for correlation functions of the current conservation equation (A13.6). Integrated over all space, they yield, not surprisingly, equations (13.17–13.20), that is, the WT identities of the symmetry.

From the point of view of renormalization, equation (A13.15) tells us that the insertion of $\partial_\mu J_\mu^\alpha(x)$ in a renormalized correlation function is finite.

In a simple renormalizable ϕ_4^4-like field theory, covariance then implies that the same must be true for the current $J_\mu^\alpha(x)$. This result is non-trivial since from expression (A13.10) we see 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 the RG equations.

A13.3 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 22.1). 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

$$\partial_\mu \phi(x) \mapsto \partial_\mu \phi(x+\varepsilon) + \partial_\mu \varepsilon^\nu \partial_\nu \phi(x+\varepsilon).$$

To calculate the variation we then change variables $x + \varepsilon = y$. Translation invariance implies that the action density depends on x through the field ϕ but not explicitly. Therefore, the only new effect is to change the measure of integration:

$$\mathrm{d}y^\mu = \mathrm{d}x^\mu + \partial_\nu \varepsilon^\mu \mathrm{d}x^\nu .$$

If we now compare the new action with the initial one (A13.1) we see 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, we obtain the identity

$$\partial_\mu T_\nu^\mu(x) = 0\,, \qquad (A13.16)$$

in which the *energy–momentum tensor* $T_\nu^\mu(x)$ is defined by

$$T_\nu^\mu(x) = \frac{\partial \mathcal{L}}{\partial[\partial_\mu \phi(x)]} \partial_\nu \phi(x) - \delta_\nu^\mu \mathcal{L}\left[\phi(x)\right]. \qquad (A13.17)$$

It is convenient to also introduce the tensor $T_{\mu\nu}$:

$$T_{\mu\nu}(x) = g_{\mu\lambda} T^{\lambda}_{\nu}(x), \tag{A13.18}$$

in which $g_{\mu\nu}$ is the Minkowski metric tensor. In the example of the lagrangian ($A13.8$),

$$T_{\mu\nu}(x) = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu}\left[\tfrac{1}{2}(\partial_\rho\phi)(\partial^\rho\phi) - V(\phi)\right], \tag{A13.19}$$

and thus $T_{\mu\nu}$ is symmetric.

To the tensor $T_{\mu\nu}(x)$ correspond constants of the classical motion P_μ, energy and momentum, obtained by integrating the time components (with respect to one index) of $T_{\mu\nu}$ over space:

$$P_\mu = \int \mathrm{d}^{d-1}x\, T_{0\mu}(x) \tag{A13.20}$$

with $x_0 \equiv t$,

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\mu = 0\,. \tag{A13.21}$$

We noted 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 General Relativity (for details see Chapter 22 and the corresponding references).

Also 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

$$\delta x^\mu = \Lambda^\mu_\nu(x) x^\nu, \tag{A13.22}$$

correspond to the choice

$$\varepsilon^\mu = \Lambda^\mu_\nu x^\nu. \tag{A13.23}$$

The corresponding currents $M^{\mu\nu\rho}$ are then

$$M^{\mu\nu\rho}(x) = T^{\mu\nu} x^\rho - T^{\mu\rho} x^\nu. \tag{A13.24}$$

Dilatation invariance. We again consider, as an example, the the ϕ^4 field theory in four dimensions:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x) - \frac{1}{2}m^2\phi^2(x) - \frac{1}{4!}g\phi^4(x). \tag{A13.25}$$

In the absence of the mass term, the action is scale-invariant, that is, invariant in the substitution

$$\phi(x) \mapsto \phi_\lambda(x) = \lambda\phi(\lambda x). \tag{A13.26}$$

For what concerns the variation of the argument, dilatation corresponds to taking ε^μ of the form

$$\varepsilon^\mu = x^\mu \lambda(x). \tag{A13.27}$$

We thus expect the dilatation current S^μ to involve $x^\nu T^\mu_\nu$. A short calculation leads to

$$S^\mu(x) = x^\nu \left[T^\mu_\nu(x) + \frac{1}{6}\left(\partial^2\delta^\mu_\nu - \partial_\mu\partial^\nu\right)\phi^2(x)\right]. \tag{A13.28}$$

In the presence of a mass term, the current $S^\mu(x)$ is not conserved. Instead,

$$\partial_\mu S^\mu(x) = m^2\phi^2(x). \tag{A13.29}$$

We now introduce the tensor $\tilde{T}_{\mu\nu}(x)$:

$$\tilde{T}_{\mu\nu}(x) = T_{\mu\nu}(x) + \frac{1}{6}\left(\partial^2 g_{\mu\nu} - \partial_\mu\partial_\nu\right)\phi^2(x). \tag{A13.30}$$

The tensor $\tilde{T}_{\mu\nu}$ can be used as energy–momentum tensor instead of $T_{\mu\nu}$: it is a polynomial in the field, symmetric as a tensor, and satisfies the conservation equation

$$\partial_\mu \tilde{T}^\mu_\nu = 0\,. \tag{A13.31}$$

In terms of $\tilde{T}^\mu_\nu(x)$, equation (A13.28) then reads

$$S^\mu(x) = x^\nu \tilde{T}^\mu_\nu(x), \tag{A13.32}$$

and the divergence of the dilatation current is

$$\partial_\mu S^\mu = \tilde{T}^\mu_\mu\,. \tag{A13.33}$$

In dilatation-invariant theories, the trace of the "improved" energy–momentum tensor $\tilde{T}^\mu_\nu$ vanishes.

A13.4 Energy–Momentum Tensor and Euclidean Field Theory

Performing the infinitesimal change of variables

$$\phi(x) = \phi'\left(x + \varepsilon(x)\right), \tag{A13.34}$$

in the functional integral, one can derive WT identities for the insertion of the energy–momentum tensor (also called the *stress tensor*). The variation of the action with a source is

$$\delta\left[\int J\phi\,\mathrm{d}x - \mathcal{S}(\phi)\right] = \varepsilon_\nu(x)\left[J(x)\partial_\nu\phi(x) + \partial_\mu T_{\mu\nu}(x)\right]. \tag{A13.35}$$

It follows that

$$\partial^x_\mu \mathcal{Z}_{T_{\mu\nu}(x)} + J(x)\partial^x_\nu \frac{\delta\mathcal{Z}}{\delta J(x)} = 0\,. \tag{A13.36}$$

Integrating this identity over space yields

$$\int \mathrm{d}x\, J(x)\partial_\nu \frac{\delta\mathcal{Z}}{\delta J(x)} = 0\,, \tag{A13.37}$$

which expresses the translation invariance of correlation functions.

After Legendre transformation, one finds

$$\partial^x_\mu \Gamma_{T_{\mu\nu}(x)} + \frac{\delta\Gamma}{\delta\varphi(x)}\partial_\nu\varphi(x) = 0\,. \tag{A13.38}$$

Again, we conclude that the insertion of the operator $\partial_\mu T_{\mu\nu}(x)$ in a renormalized correlation function is finite. However, this does not imply that the insertion of $T_{\mu\nu}$ itself is finite. In the ϕ^4_4 field theory 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 $T_{\mu\nu}$:

$$(T_{\mu\nu})_{\mathrm{r}} = T_{\mu\nu} + A(\delta_{\mu\nu}\nabla^2 - \partial_\mu\partial_\nu)(\phi^2)_{\mathrm{r}} . \tag{A13.39}$$

Note that the renormalized energy–momentum tensor automatically has a non-vanishing trace, and it can no longer be improved since the coefficient A 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. It is, nevertheless, possible to derive WT identities involving the divergence of the dilatation current. By integrating them over space, one obtains the CS equations derived in Chapter 10.

A13.5 Dilatation and Conformal Invariance

We now consider a general euclidean action, $\mathcal{S}$, invariant under translation, rotation and dilatation. We perform the infinitesimal change of variables

$$x_\mu \longmapsto x_\mu + \varepsilon_\mu(x). \tag{A13.40}$$

Translation invariance implies that the variation of the action involves only the partial derivatives of $\varepsilon_\mu(x)$:

$$\delta\mathcal{S} = \int \mathrm{d}^d x\, T_{\mu\nu}(x)\partial_\mu\varepsilon_\nu(x). \tag{A13.41}$$

Rotation invariance implies that $\delta\mathcal{S}$ vanishes for

$$\varepsilon_\mu = \Lambda_{\mu\nu}x_\nu\,, \tag{A13.42}$$

in which $\Lambda_{\mu\nu}$ is an arbitrary antisymmetric matrix. Therefore, the integral of the stress tensor must be symmetric:

$$\int \mathrm{d}^d x (T_{\mu\nu} - T_{\nu\mu}) = 0\,.$$

Dilatation invariance corresponds to

$$\varepsilon_\mu = \lambda x_\mu\,, \tag{A13.43}$$

and implies the vanishing of the integral of the trace of the stress tensor:

$$\int \mathrm{d}^d x\, T_{\mu\mu} = 0\,.$$

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:

$$T_{\mu\nu} = T_{\nu\mu}\,, \quad T_{\mu\mu} = 0\,.$$

It then follows that the variation of the action also vanishes for any function ε_μ which satisfies

$$\partial_\mu \varepsilon_\nu + \partial_\nu \varepsilon_\mu - \tfrac{1}{2} d\delta_{\mu\nu}\partial\cdot\varepsilon = 0\,, \tag{A13.44}$$

where d is the dimension of euclidean space. The group of transformations which satisfy equation (A13.44) is larger than the product of transformations which we have considered so far: it is the whole *conformal* group. Indeed, let us calculate the variation of a line element of the form

$$(\mathrm{d}s)^2 = g(x)\mathrm{d}x_\mu \mathrm{d}x_\mu\,, \tag{A13.45}$$

which corresponds to a conformally flat metric.

We find

$$\delta\left[(\mathrm{d}s)^2\right] = \mathrm{d}x_\mu \mathrm{d}x_\mu \partial_\rho g(x)\varepsilon_\rho + \mathrm{d}x_\mu \mathrm{d}x_\nu\left(\partial_\mu\varepsilon_\nu + \partial_\nu\varepsilon_\mu\right)g(x). \tag{A13.46}$$

We now see that equation (A13.44) is the necessary and sufficient condition for the line element to retain the form (A13.45). By definition, the transformations which preserve the form of the metric (A13.45) are conformal transformations.

From equation (A13.44) it follows that

$$\left[\delta_{\mu\nu}\nabla^2 + (d-2)\partial_\mu\partial_\nu\right]\partial\cdot\varepsilon = 0\,. \tag{A13.47}$$

For $d > 2$ the equation implies that all second derivatives of $\partial\cdot\varepsilon$ vanish. Returning then to equation (A13.44) one shows easily that all third derivatives of ε_μ also vanish. Solutions of degree 0 correspond to translations. Solutions of degree 1 correspond to rotations and dilatations. The additional solutions of equation (A13.44) are second degree polynomials of the form

$$\varepsilon_\mu = a_\mu x^2 - 2x_\mu a\cdot x\,. \tag{A13.48}$$

They correspond to special conformal transformations. The integrated form of these transformations is

$$x'_\mu = \frac{x_\mu + a_\mu x^2}{1 + 2a\cdot x + a^2x^2}. \tag{A13.49}$$

The conformal group is isomorphic to $SO(d+1,1)$. Imposing conformal invariance on correlation functions determines, in particular, two- and three-point functions.

In dimension $d = 2$ the situation is completely different: the set of equations (A13.44) 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 consequences are much more striking and lead to the classification of a whole class of conformally invariant field theories.

Of course, we have seen in Chapter 10 that the scale invariance of the classical theory is broken at the quantum level. However, as we shall discuss in Chapter 25, there exist situations in which the RG β-function vanishes, at least for some values of the coupling constants. Then both the dilatation invariance and, therefore, the conformal invariance are restored.

Remark. The condition that the action should depend only on the field and its first derivatives can be illustrated by a simple counter-example. Consider the free action $\mathcal{S}(\phi)$

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left(\nabla^2\phi(x)\right)^2.$$

The propagator in Fourier space is $1/p^4$. The theory is obviously translation, rotation and scale invariant. However, one verifies that it is not conformal invariant.

14 THE NON-LINEAR σ-MODEL: AN EXAMPLE OF A NON-LINEAR SYMMETRY

We now consider models possessing global symmetries non-linearly realized on the fields. This implies, in particular, that under an infinitesimal group transformation, the variation of the field is a non-linear function of the field itself. Since such models have non-trivial geometric properties, we first extensively discuss the simplest example, the non-linear σ-model, a model with an $O(N)$ symmetry, the field being an N-vector of fixed length.

A simple analysis reveals that in the non-linear σ-model, in the tree approximation, the $O(N)$ symmetry is always spontaneously broken, unlike what happens in a ϕ^4-like theory with the same symmetry: the action describes the interactions of $N-1$ massless fields, the Goldstone modes.

Power counting shows that the model is renormalizable in two dimensions. Therefore, the field is dimensionless and we face a problem already mentioned in Section 9.3: although the degree of divergence of Feynman diagrams is bounded, an infinite number of counter-terms is generated because all correlation functions are divergent. We prove in this chapter 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.

Note that since the fields are massless, in two dimensions IR divergences appear in the perturbative expansion and an IR regulator is thus required.

In Section 14.8 we discuss the renormalization of composite operators. Finally, in Section 14.9 we indicate how the results can be recovered from another, linear, representation of the model where the condition that the field is a vector of fixed length is enforced by a Lagrange multiplier.

In Chapter 15 we shall show how the arguments generalize to all models defined on homogeneous spaces. Actually the $O(N)$ non-linear σ-model belongs to a class of models constructed on special homogeneous spaces, the symmetric spaces, which as Riemannian manifolds, admit a unique metric. We shall study them in more detail both as classical and quantum field theories.

14.1 The Non-Linear σ-Model: Definition

The field manifold. We consider an N-vector field $\boldsymbol{\phi}(x)$ which satisfies an $O(N)$ invariant constraint,

$$\boldsymbol{\phi}^2(x) = 1\,. \tag{14.1}$$

The field $\boldsymbol{\phi}$ 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 $\mathbf{u}$ a fixed N-vector:

$$\mathbf{u} = (1, 0, \dots, 0), \tag{14.2}$$

and, thus, $\mathbf{u}^2 = 1$. The field $\boldsymbol{\phi}$ can be parametrized as

$$\boldsymbol{\phi}(x) = \mathbf{g}(x)\mathbf{u}\,, \tag{14.3}$$

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 $O(N)$ which 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 r.h.s. of equation (14.3) is left unchanged. The relation (14.3) thus exhibits the isomorphism between the coset (homogeneous) space $O(N)/O(N-1)$ and the sphere S_{N-1}.

Below, we need a parametrization of the field $\boldsymbol{\phi}$ in terms of independent variables. A convenient parametrization of the sphere (14.1) is

$$\boldsymbol{\phi}(x) = \{\sigma(x), \boldsymbol{\pi}(x)\}, \tag{14.4}$$

in which $\boldsymbol{\pi}(x)$ is an $(N-1)$-component field and the field $\sigma(x)$ a function of $\boldsymbol{\pi}(x)$ through equation (14.1). The equation can be solved locally, for example, if $\sigma(x)$ is positive:

$$\sigma(x) = \left(1 - \boldsymbol{\pi}^2(x)\right)^{1/2}. \tag{14.5}$$

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

$$\delta\pi_i = \omega_i\left(1 - \boldsymbol{\pi}^2(x)\right)^{1/2}, \tag{14.6}$$

in which ω_i are constants, infinitesimal parameters of the transformation. The transformation law of the σ-field is then a consequence of the transformation (14.6) of the $\boldsymbol{\pi}$-field:

$$\delta\sigma(x) \equiv \delta\left(1 - \boldsymbol{\pi}^2(x)\right)^{1/2} = -\boldsymbol{\omega}\cdot\boldsymbol{\pi}(x). \tag{14.7}$$

The action and functional integral. The most general $O(N)$ symmetric action containing at most two derivatives is, up to a multiplicative constant (in a theory invariant under space translations)

$$\mathcal{S}(\boldsymbol{\phi}) = \frac{1}{2}\int \mathrm{d}^d x\, \partial_\mu\boldsymbol{\phi}(x)\cdot\partial_\mu\boldsymbol{\phi}(x). \tag{14.8}$$

Indeed, due to the constraint (14.1), any symmetric derivative-free term reduces to a constant (and $\boldsymbol{\phi}\cdot\partial_\mu\boldsymbol{\phi}$ vanishes).

In terms of the field $\boldsymbol{\pi}(x)$, the action (14.8) can be cast into another geometric form

$$\mathcal{S}(\boldsymbol{\pi}) = \frac{1}{2}\int \mathrm{d}^d x\, G_{ij}\left(\boldsymbol{\pi}(x)\right)\partial_\mu\pi_i(x)\partial_\mu\pi_j(x), \tag{14.9}$$

in which $G_{ij}(\boldsymbol{\pi})$ is a metric tensor on the sphere:

$$G_{ij} = \delta_{ij} + \frac{\pi_i\pi_j}{1-\boldsymbol{\pi}^2}. \tag{14.10}$$

In the form (14.9) the action is covariant under a reparametrization of the sphere.

We have seen in Section 3.2 that the quantization of actions of the form (14.9) introduces an additional ill-defined (since infinite) determinant. From the equations (3.23–3.26) we infer a representation of the generating functional of correlation functions:

$$\mathcal{Z}(\mathbf{J}) = \int \left[\frac{\mathrm{d}\boldsymbol{\pi}(x)}{\left(1-\boldsymbol{\pi}^2(x)\right)^{1/2}}\right] \exp\left[-\frac{1}{g}\left(\mathcal{S}(\boldsymbol{\pi}) - \int \mathrm{d}^d x\, \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\right)\right], \tag{14.11}$$

in which g is the coupling constant of the quantum field theory.

We had noticed that in the case of Riemannian manifolds a non-trivial measure is needed for geometric reasons: here we, indeed, find the $O(N)$-invariant functional measure $\mathrm{d}\boldsymbol{\pi}/\sqrt{1-\boldsymbol{\pi}^2}$ on the sphere (see Section 3.4).

This formal interpretation of the determinant, however, does not eliminate the difficulty. We have to find a method to deal with this infinite contribution to the action:

$$\prod_x \left(1-\boldsymbol{\pi}^2(x)\right)^{-1/2} \sim \exp\left[-\frac{1}{2}\delta^d(0)\int \mathrm{d}^d x\, \ln\left(1-\boldsymbol{\pi}^2(x)\right)\right]. \tag{14.12}$$

As we have explained in Section 3.2, this difficulty is directly related to the problem of operator ordering which appears in the quantization.

14.2 Perturbation Theory. Power Counting

For $g \to 0$ and in zero source ($J = 0$) the functional integral is dominated by saddle points, minima of the classical action (14.8):

$$|\partial_\mu \boldsymbol{\phi}(x)| = 0 \quad \Rightarrow \quad \boldsymbol{\phi}(x) = \boldsymbol{\phi}_0\ ,$$

where $\boldsymbol{\phi}_0$ 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 already mentioned 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, postponing a more complete analysis to Chapters 23–33 in which the theory of phase transitions is discussed. Below, we first examine the contribution of one saddle point. In the parametrization (14.4) we choose the minimum $\boldsymbol{\pi}(x) = 0$ or $\boldsymbol{\phi}(x) = \mathbf{u}$.

14.2.1 Formal perturbation theory

We have introduced in the functional integral (14.11) a parameter g which plays the formal role of $\hbar$ and, therefore, orders perturbation theory. For g small, the fields $\boldsymbol{\pi}(x)$ which contribute to the functional integral, then, must be such that

$$|\partial_\mu \boldsymbol{\pi}(x)| \sim \sqrt{g}\ ,$$

and since we expand around $\boldsymbol{\pi}(x) = 0$, the field itself must satisfy

$$|\boldsymbol{\pi}(x)| \sim \sqrt{g}\ . \tag{14.13}$$

Values of $\boldsymbol{\pi}(x)$ of order 1 give exponentially small contributions to the functional 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 (14.5) ($\sigma(x) > 0$) are irrelevant in perturbation theory and in addition, in the functional integral, we can freely integrate over $\boldsymbol{\pi}(x)$ from $+\infty$ to $-\infty$, disregarding the constraint

$$|\boldsymbol{\pi}(x)| \le 1 .$$

Perturbation theory, then, again relies on the evaluation of simple gaussian integrals.

We can now discuss formal perturbation theory, setting aside temporarily the question of UV or IR (low momentum) divergences. We rewrite the functional integral (14.11):

$$\mathcal{Z}(\mathbf{J}) = \int [\mathrm{d}\boldsymbol{\pi}] \exp\left[-\int \mathrm{d}^d x\, \mathcal{L}(\boldsymbol{\pi}, \mathbf{J})\right],$$

$$\mathcal{L}(\boldsymbol{\pi}, \mathbf{J}) = \frac{1}{2g}\left[(\partial_\mu \boldsymbol{\pi})^2 + \frac{(\boldsymbol{\pi}\cdot\partial_\mu\boldsymbol{\pi})^2}{1-\boldsymbol{\pi}^2}\right] + \frac{1}{2}\delta^d(0)\ln\left(1-\boldsymbol{\pi}^2(x)\right) - \frac{1}{g}\mathbf{J}(x)\cdot\boldsymbol{\pi}(x). \quad (14.14)$$

Note that the measure term has no $1/g$ factor and starts contributing only at one-loop order. Since $\boldsymbol{\pi}$ is of order $\sqrt{g}$, it is convenient to rescale the field:

$$\boldsymbol{\pi} \mapsto \boldsymbol{\pi}\sqrt{g} .$$

After this rescaling, the action density $\mathcal{L}$ becomes

$$\mathcal{L}(\boldsymbol{\pi}, \mathbf{J}) = \frac{1}{2}\left[(\partial_\mu \boldsymbol{\pi})^2 + g\frac{(\boldsymbol{\pi}\cdot\partial_\mu\boldsymbol{\pi})^2}{1-g\boldsymbol{\pi}^2}\right] + \frac{1}{2}\delta^d(0)\ln\left(1-g\boldsymbol{\pi}^2\right) - \frac{1}{\sqrt{g}}\mathbf{J}(x)\cdot\boldsymbol{\pi}(x). \quad (14.15)$$

Expression (14.15) 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 easy to verify that at any finite order in perturbation theory and for a given correlation function, only a finite number of vertices contribute. Formally, the measure term yields additional vertices without derivatives.

The propagator $\Delta_{ij}(p)$ of the $\boldsymbol{\pi}$-field is

$$\Delta_{ij}(p) = \frac{\delta_{ij}}{p^2}. \quad (14.16)$$

In the tree approximation, the $\boldsymbol{\pi}$-field is massless. Returning to the analysis of Section 13.4, we understand 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 (14.1). This constraint formally freezes the fluctuations of $\boldsymbol{\phi}^2(x)$ and sends, in the classical limit, the σ mass to infinity.

Note that these specific properties are independent of the special choice (14.4) of parametrization of $\boldsymbol{\phi}(x)$.

14.2.2 Power counting

The general analysis has been given in Chapter 9. The form of the propagator shows that the dimension $[\boldsymbol{\pi}]$ of the $\boldsymbol{\pi}$-field is

$$[\boldsymbol{\pi}] = \tfrac{1}{2}(d-2). \tag{14.17}$$

Therefore, the dimension of a vertex containing $2n$ $\boldsymbol{\pi}$-fields is

$$\left[\partial^2 \boldsymbol{\pi}^{2n}\right] = n(d-2) + 2. \tag{14.18}$$

As a consequence

(i) for $d < 2$ the theory is super-renormalizable by power counting;
(ii) for $d = 2$ it is just renormalizable;
(iii) for $d > 2$ the theory is not renormalizable.

We, therefore, 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

$$(\partial_\mu \boldsymbol{\pi} \cdot \boldsymbol{\pi})^2 \left(\boldsymbol{\pi}^2\right)^n, \qquad (\partial_\mu \boldsymbol{\pi})^2 \left(\boldsymbol{\pi}^2\right)^n, \qquad \left(\boldsymbol{\pi}^2\right)^n.$$

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 to first exhibit a regularization scheme which preserves the $O(N)$ symmetry and then derive a set of WT identities which express the consequences of the symmetry for correlation functions.

14.3 Regularization

In constructing a regularized version of the non-linear σ-model, one has to be careful 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 (14.9) are related to the quadratic part. A simple method is to start from the description of the model in terms of the $\boldsymbol{\phi}$-field because the action (14.8) is formally a free field action.

14.3.1 Perturbative regularizations

Momentum cut-off regularization. A natural idea is to use the regularization (13.12) that works for linear symmetries. The action $\mathcal{S}(\boldsymbol{\phi})$ is replaced by $\mathcal{S}_\Lambda(\boldsymbol{\phi})$:

$$\mathcal{S}_\Lambda(\boldsymbol{\phi}) = \frac{1}{2}\int \mathrm{d}^d x\, \partial_\mu \boldsymbol{\phi}(x) \cdot \prod_{r=1}^{r_{\max}} (1 - \nabla^2/M_r^2)\partial_\mu \boldsymbol{\phi}(x), \tag{14.19}$$

in which the masses M_r are proportional to the cut-off Λ. Expressing, then, $\boldsymbol{\phi}(x)$ in terms of $\boldsymbol{\pi}(x)$, 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 (9.19) which gives the superficial degree of divergence of a diagram γ,

$$\delta(\gamma) = 2L - sI + \sum_\alpha k_\alpha v_\alpha \,,$$

and eliminating the number of internal lines I through the topological relation (9.21),

$$L = I - \sum_\alpha v_\alpha + 1 \,,$$

we find

$$\delta(\gamma) = (2-s)\,L + s + \sum_\alpha (k_\alpha - s)\,v_\alpha \,, \tag{14.20}$$

in which we recall that L is the number of loops and k_α the number of derivatives at vertex α.

As stated above, the worst case is $k_\alpha = s$:

$$\delta(\gamma) \le (2-s)\,L + s \,.$$

As a consequence, for $L \ge 2$ it is sufficient to take $s \ge 6$ to regularize all diagrams. However, the one-loop diagrams have a behaviour independent of s, and thus cannot be regularized. This property is not independent of the other limitations of momentum cut-off (or Pauli–Villars's) regularization, that it cannot regularize the divergent measure term. Indeed, we show later that the one-loop divergences generated by the interaction are needed to cancel the divergences coming from the measure.

Momentum regularization is mainly useful in the study of non-linear models coupled to chiral fermions.

Dimensional regularization. Dimensional regularization preserves the $O(N)$ symmetry of the action. Furthermore, as a consequence of the formal rule

$$\int \mathrm{d}^d k = (2\pi)^d \delta^d(0) = 0 \,,$$

the measure term can be ignored, and, therefore, perturbation theory has no large momentum divergences for $d < 2$. Owing to his technical simplicity, this is the regularization we shall in general use for practical calculations. One theoretical drawback is that the role of the measure is hidden. Therefore, for the theoretical discussion of the renormalization of the non-linear σ-model we shall consider both dimensional and lattice regularizations.

14.3.2 Lattice regularization and statistical mechanics

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 $\boldsymbol{\phi}$-field. We then replace derivatives by finite differences, that is, in the notation of Section 9.6,

$$\partial_\mu \boldsymbol{\phi}(x) \longmapsto \nabla^{\mathrm{lat}}_\mu \boldsymbol{\phi}(x) = \left[\boldsymbol{\phi}(x + an_\mu) - \boldsymbol{\phi}(x)\right]/a,$$

in which x now belongs to a hypercubic lattice of lattice spacing a, and n_μ is the unit vector in the μ direction.

then the most singular interaction has s derivatives. Using equation (9.19) which gives the superficial degree of divergence of a diagram γ,

$$\delta(\gamma) = 2L - sI + \sum_{\alpha} k_\alpha v_\alpha \,,$$

and eliminating the number of internal lines I through the topological relation (9.21),

$$L = I - \sum_{\alpha} v_\alpha + 1 \,,$$

we find

$$\delta(\gamma) = (2-s)\,L + s + \sum_{\alpha} (k_\alpha - s)\, v_\alpha \,, \tag{14.20}$$

in which we recall that L is the number of loops and k_α the number of derivatives at vertex α.

As stated above, the worst case is $k_\alpha = s$:

$$\delta(\gamma) \le (2-s)\,L + s \,.$$

As a consequence, for $L \ge 2$ it is sufficient to take $s \ge 6$ to regularize all diagrams. However, the one-loop diagrams have a behaviour independent of s, and thus cannot be regularized. This property is not independent of the other limitations of momentum cut-off (or Pauli–Villars's) regularization, that it cannot regularize the divergent measure term. Indeed, we show later that the one-loop divergences generated by the interaction are needed to cancel the divergences coming from the measure.

Momentum regularization is mainly useful in the study of non-linear models coupled to chiral fermions.

Dimensional regularization. Dimensional regularization preserves the $O(N)$ symmetry of the action. Furthermore, as a consequence of the formal rule

$$\int \mathrm{d}^d k = (2\pi)^d \delta^d(0) = 0 \,,$$

the measure term can be ignored, and, therefore, perturbation theory has no large momentum divergences for $d < 2$. Owing to his technical simplicity, this is the regularization we shall in general use for practical calculations. One theoretical drawback is that the role of the measure is hidden. Therefore, for the theoretical discussion of the renormalization of the non-linear σ-model we shall consider both dimensional and lattice regularizations.

14.3.2 Lattice regularization and statistical mechanics

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 9.6,

$$\partial_\mu \phi(x) \longmapsto \nabla^{\text{lat}}_\mu \phi(x) = \left[\phi(x + a n_\mu) - \phi(x)\right]/a,$$

in which x now belongs to a hypercubic lattice of lattice spacing a, and n_μ is the unit vector in the μ direction.

Finally, to implement condition (14.1), we integrate over $\boldsymbol{\phi}(x)$ with the invariant measure on the sphere. The regularized functional integral has the form

$$\mathcal{Z}(\mathbf{J}) = \int \prod_{x\in a\mathbb{Z}^d} \delta\left(\boldsymbol{\phi}^2(x) - 1\right) \mathrm{d}\boldsymbol{\phi}(x) \exp\left[-\frac{1}{g}\mathcal{S}(\boldsymbol{\phi}, \mathbf{J})\right] \tag{14.21}$$

with

$$\mathcal{S}(\boldsymbol{\phi}, \mathbf{J}) = \frac{1}{2}\sum_{x,\mu} \left[\nabla_\mu \boldsymbol{\phi}(x)\right]^2 - \sum_x \mathbf{J}(x)\cdot\boldsymbol{\phi}(x). \tag{14.22}$$

Using the parametrization (14.4), we can express the lattice field $\boldsymbol{\phi}(x)$ in terms of $\boldsymbol{\pi}(x)$. We obtain a regularized form of the functional integral (14.11). In particular, the measure term now generates well-defined interactions:

$$\frac{1}{2}\delta^d(0)\int \mathrm{d}^d x \,\ln\left(1 - \boldsymbol{\pi}^2(x)\right) \longmapsto \frac{1}{2}\sum_x \ln\left(1 - \boldsymbol{\pi}^2(x)\right).$$

In the expression (14.21) we recognize the partition function of a classical spin lattice model with a nearest-neighbour ferromagnetic interaction 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 this model will be discussed in Chapters 23–33. In particular, in Chapter 31 we will use the perturbative expansion of the field theory model to study ferromagnetic order at low temperature. Expression (14.21) not only provides a regularization of perturbation theory but also allows the use of various non-perturbative methods to study the non-linear σ-model. Moreover, it is the only regularization which allows a discussion of the role of the measure in perturbation theory (see Sections 14.5.1,14.6).

14.4 Infrared (IR) Divergences

Since in a massless theory the propagator behaves like $1/p^2$, perturbation theory is divergent at low momentum (IR) in dimension 2, the dimension in which the non-linear σ-model is renormalizable (see Section 31.2.2 for explicit calculations). As we will argue later, this divergence is directly related to the absence of spontaneous breaking of continuous symmetries in two dimensions.

To generate a finite perturbation theory, it is necessary to introduce an IR cut-off. Since the absence of mass is a consequence of the spontaneous breaking of the $O(N)$ symmetry in the classical limit, it is necessary, to give a mass to the $\boldsymbol{\pi}$-field, to break the symmetry explicitly. We can for example introduce an explicit mass term. However, the study of symmetry breaking mechanisms in Chapter 13 suggests a more convenient method which consists in adding to the action (14.9) a constant source h for the σ-field (a magnetic field in the statistical model of classical spins):

$$\mathcal{S}(\boldsymbol{\pi}) \longmapsto \mathcal{S}(\boldsymbol{\pi}) - h\int \sigma(x)\,\mathrm{d}^d x\,, \qquad h > 0. \tag{14.23}$$

A first consequence of this modification is that the minimum of the action is no longer degenerate. Instead, we now have to maximize the source term and this implies $\boldsymbol{\pi} = 0$ at the minimum.

Second, if we expand σ in powers of $\boldsymbol{\pi}$,

$$\sigma = \left(1 - \boldsymbol{\pi}^2\right)^{1/2} = 1 - \tfrac{1}{2}\boldsymbol{\pi}^2 + O\left((\boldsymbol{\pi}^2)^2\right), \tag{14.24}$$

and collect the quadratic terms in the action, we find the new $\boldsymbol{\pi}$-field propagator $\Delta_{ij}(p)$

$$\Delta_{ij}(p) = \delta_{ij}\frac{g}{p^2 + h}. \tag{14.25}$$

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 the 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. We shall discuss in Chapter 37 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 and, 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

$$\mathbf{p} = (2\pi/L)\,\mathbf{n}\,, \quad \mathbf{n} \in \mathbb{Z}^d.$$

In this case, some care is needed to handle properly 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. Let us here mention briefly, without proof, an interesting result whose significance will be discussed later with the physics (in the sense of statistical mechanics) of the non-linear σ-model. We have explained that in two dimensions correlation functions are IR divergent and we have introduced an IR cut-off in the form of a source term giving a mass term for the $\boldsymbol{\pi}$-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 noted in Section 14.2 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, order by order in perturbation theory, that $O(N)$ invariant correlation functions have in two dimensions a finite IR limit, that is, for example, a limit when in the notation (14.23) the breaking parameter h goes to zero, or any IR cut-off is removed (see also Chapter 31).

14.5 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 have been discussed in Chapter 13 and are now obvious.

Second, if we expand σ in powers of $\boldsymbol{\pi}$,

$$\sigma = \left(1 - \boldsymbol{\pi}^2\right)^{1/2} = 1 - \tfrac{1}{2}\boldsymbol{\pi}^2 + O\left((\boldsymbol{\pi}^2)^2\right), \tag{14.24}$$

and collect the quadratic terms in the action, we find the new $\boldsymbol{\pi}$-field propagator $\Delta_{ij}(p)$

$$\Delta_{ij}(p) = \delta_{ij}\frac{g}{p^2 + h}. \tag{14.25}$$

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 the 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. We shall discuss in Chapter 37 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 and, 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

$$\mathbf{p} = (2\pi/L)\,\mathbf{n}\,, \quad \mathbf{n} \in \mathbb{Z}^d.$$

In this case, some care is needed to handle properly 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. Let us here mention briefly, without proof, an interesting result whose significance will be discussed later with the physics (in the sense of statistical mechanics) of the non-linear σ-model. We have explained that in two dimensions correlation functions are IR divergent and we have introduced an IR cut-off in the form of a source term giving a mass term for the $\boldsymbol{\pi}$-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 noted in Section 14.2 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, order by order in perturbation theory, that $O(N)$ invariant correlation functions have in two dimensions a finite IR limit, that is, for example, a limit when in the notation (14.23) the breaking parameter h goes to zero, or any IR cut-off is removed (see also Chapter 31).

14.5 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 have been discussed in Chapter 13 and are now obvious.

14.5.1 WT identities

We consider infinitesimal transformations of the form

$$\delta\pi_i(x) = \omega_i\sigma(x) \equiv \omega_i\sqrt{1-\boldsymbol{\pi}^2(x)}\,. \tag{14.26}$$

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:

$$\delta\left[\int \mathrm{d}^d x\, \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\right] = \int \mathrm{d}^d x\, \boldsymbol{\omega}\cdot\mathbf{J}(x)\sigma(x), \qquad \delta\left[h\int \mathrm{d}^d x\, \sigma(x)\right] = -h\int \mathrm{d}^d x\, \boldsymbol{\omega}\cdot\boldsymbol{\pi}(x).$$

A new composite operator $\sigma(x) = (1-\boldsymbol{\pi}^2(x))^{1/2}$ is generated by the variation of the source term. Following the general strategy explained already in the case of quadratic symmetry breaking terms, we introduce a source for this operator. Then, since the variation under a transformation (14.26) of this new operator as well as the variation of the breaking term 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 $\int \boldsymbol{\pi}^2(x)\mathrm{d}^d x$. We call $H(x)$ the source for σ, and $\mathcal{Z}(\mathbf{J},H)$ the generating functional:

$$\mathcal{Z}(\mathbf{J},H) = \int\left[\frac{\mathrm{d}\boldsymbol{\pi}(x)}{\sqrt{1-\boldsymbol{\pi}^2(x)}}\right]\exp\left\{\frac{1}{g}\left[-\mathcal{S}(\boldsymbol{\pi},H) + \int \mathrm{d}^d x\, \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\right]\right\} \tag{14.27}$$

with

$$\mathcal{S}(\boldsymbol{\pi},H) = \mathcal{S}(\boldsymbol{\pi}) - \int \mathrm{d}^d x\, H(x)\sigma(x)\,. \tag{14.28}$$

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 (14.26), we obtain the equation

$$0 = \int\left[\frac{\mathrm{d}\boldsymbol{\pi}}{\sqrt{1-\boldsymbol{\pi}^2}}\right]\int \mathrm{d}^d x\left[\mathbf{J}\cdot\delta\boldsymbol{\pi} + H(x)\delta\sigma\right]\exp\left[\frac{1}{g}\left(-\mathcal{S}(\boldsymbol{\pi},H) + \int \mathrm{d}^d x\, \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\right)\right],$$

or, explicitly,

$$0 = \int\left[\frac{\mathrm{d}\boldsymbol{\pi}}{\sqrt{1-\boldsymbol{\pi}^2}}\right]\int \mathrm{d}^d x\left[\sigma(x)\mathbf{J}(x) - H(x)\boldsymbol{\pi}(x)\right]\exp\left[\frac{1}{g}\left(-\mathcal{S}(\boldsymbol{\pi},H) + \int \mathrm{d}^d x\, \mathbf{J}\cdot\boldsymbol{\pi}\right)\right]. \tag{14.29}$$

We now replace $\boldsymbol{\pi}(x)$ by $g(\delta/\delta\mathbf{J}(x))$ and $\sigma(x)$ by $g(\delta/\delta H(x))$. Equation (14.29) then becomes

$$\int \mathrm{d}^d x\left(\mathbf{J}(x)\frac{\delta}{\delta H(x)} - H(x)\frac{\delta}{\delta\mathbf{J}(x)}\right)\mathcal{Z}(\mathbf{J},H) = 0\,. \tag{14.30}$$

This first order linear differential equation makes no reference to the non-linear character of the transformation (14.26). 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:

$$\int \mathrm{d}^d x \left(\mathbf{J}(x) \frac{\delta}{\delta H(x)} - H(x) \frac{\delta}{\delta \mathbf{J}(x)} \right) \mathcal{W}(\mathbf{J},H) = 0 \,. \tag{14.31}$$

We now perform a Legendre transformation. In contrast with the linear case, however, $\sigma(x)$ is here a function of the $\boldsymbol{\pi}(x)$ field, and, therefore, the Legendre transformation applies only to the source $\mathbf{J}(x)$ and not to $H(x)$:

$$\mathcal{W}(\mathbf{J},H) + \Gamma(\boldsymbol{\pi},H) = \int \mathrm{d}^d x \, \mathbf{J}(x) \cdot \boldsymbol{\pi}(x), \qquad \boldsymbol{\pi}(x) = \frac{\delta \mathcal{W}}{\delta \mathbf{J}(x)} \cdot \tag{14.32}$$

We again use identity (7.73), since $H(x)$ is an external parameter:

$$\left. \frac{\delta \mathcal{W}}{\delta H(x)} \right|_{\mathbf{J}} = - \left. \frac{\delta \Gamma}{\delta H(x)} \right|_{\boldsymbol{\pi}} . \tag{14.33}$$

The Legendre transform of equation (14.31) is, therefore,

$$\int \mathrm{d}^d x \left(\frac{\delta \Gamma}{\delta \boldsymbol{\pi}(x)} \frac{\delta \Gamma}{\delta H(x)} + H(x) \boldsymbol{\pi}(x) \right) = 0 \,. \tag{14.34}$$

This is the basic equation from which the general form of the counter-terms which render the theory finite can be derived.

Remark. Equation (14.34) is quadratic in Γ. This is an essential difference with the case of linearly realized symmetries. Gauge theories will provide another example sharing this property. Actually, one can show that if one uses the strategy explained above, that is, adds sources for all new composite operators generated in the group transformation, then the WT identities derived for the 1PI functional are at most quadratic in Γ.

14.5.2 Master equation

It is easy to verify that the initial action satisfies equation (14.34), either directly, or by performing a loop expansion of equation (14.34) 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*

$$\int \mathrm{d}^d x \left(\frac{\delta \mathcal{S}}{\delta H(x)} \frac{\delta \mathcal{S}}{\delta \boldsymbol{\pi}(x)} + H(x) \boldsymbol{\pi}(x) \right) = 0 \,. \tag{14.35}$$

We then perform an infinitesimal change of variables in the functional integral (14.27) of the form

$$\boldsymbol{\pi} = \boldsymbol{\pi}' + \delta \boldsymbol{\pi} \quad \text{with} \quad \delta \boldsymbol{\pi} = \frac{\delta \mathcal{S}(\boldsymbol{\pi}',H)}{\delta H(x)} \boldsymbol{\omega} \,. \tag{14.36}$$

In the dimensional regularization scheme we can omit the measure term and the jacobian because $\mathcal{S}$ is local (for other regularizations see the remark below).

The variations of the action and the source term are (we omit the primes on the dummy variable $\boldsymbol{\pi}$)

$$\delta\left\{\exp\left[\frac{1}{g}\left(-\mathcal{S}(\boldsymbol{\pi},H)+\int \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\mathrm{d}^d x\right)\right]\right\}$$
$$=\frac{\omega}{g}\cdot\left[H(x)\frac{\delta}{\delta\mathbf{J}(x)}-\mathbf{J}(x)\frac{\delta}{\delta H(x)}\right]\exp\left[\frac{1}{g}\left(-\mathcal{S}(\boldsymbol{\pi},H)+\int \mathbf{J}(x)\cdot\boldsymbol{\pi}(x)\mathrm{d}^d x\right)\right]. \quad (14.37)$$

Therefore, the master equation (14.35), alone, implies equation (14.30) for $\mathcal{Z}(\mathbf{J},H)$ and, thus, equation (14.34) for $\Gamma(\boldsymbol{\pi},H)$.

The measure. The argument given above is valid only for dimensional regularization where the measure term vanishes identically. We now extend it to the case of lattice regularization by showing that the invariant measure for the transformations (14.36) is

$$\prod_x \mathrm{d}\boldsymbol{\pi}(x)\left(\frac{\delta\mathcal{S}}{\delta H(x)}\right)^{-1}. \quad (14.38)$$

First, the change of variables (14.36) generates a jacobian

$$\mathcal{J}=1+\omega_i\int \mathrm{d}x\,\frac{\delta^2\mathcal{S}}{\delta\pi_i(x)\delta H(x)}\,.$$

Correspondingly, the variation of the measure term is

$$\left(\frac{\delta\mathcal{S}}{\delta H(x)}\right)^{-1}\mapsto\left(\frac{\delta\mathcal{S}}{\delta H(x)}\right)^{-1}\left\{1-\left(\frac{\delta\mathcal{S}}{\delta H(x)}\right)^{-1}\frac{\delta^2\mathcal{S}}{\delta\pi_i(x)\delta H(x)}\delta\pi_i(x)\right\},$$

which, using the explicit form (14.36), exactly cancels the jacobian. Note, finally, that the initial measure $[\mathrm{d}\boldsymbol{\pi}(1-\boldsymbol{\pi}^2)^{-1/2}]$ has the form (14.38).

14.6 Renormalization

Before explaining the technical details, we describe the various steps of the proof of the renormalizability of the model.

First, the stability of equation (14.35) under renormalization will be established: this means that if the action at tree order satisfies equation (14.35), then it is possible to renormalize the theory in such a way that the renormalized action still satisfies equation (14.35). Then equation (14.35) will be solved. It is important to realize that the equation does not explicitly refer to the transformation law (14.26). 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 (14.35) 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 $\boldsymbol{\phi}$ now belongs to a sphere of renormalized radius:

$$\boldsymbol{\phi}^2(x)=\boldsymbol{\pi}^2(x)+\sigma^2(x)=1/Z\ .$$

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 1PI functional $\Gamma(\boldsymbol{\pi}, H)$:

$$\Gamma(\boldsymbol{\pi}, H) = \sum_{n=0}^{\infty} \Gamma_n g^n .$$

We then insert this expansion into equation (14.34). The functional Γ_0 is simply the initial action and satisfies by itself equation (14.34). The one-loop functional Γ_1 satisfies

$$\int \mathrm{d}^d x \left(\frac{\delta \Gamma_0}{\delta \boldsymbol{\pi}(x)} \frac{\delta \Gamma_1}{\delta H(x)} + \frac{\delta \Gamma_0}{\delta H(x)} \frac{\delta \Gamma_1}{\delta \boldsymbol{\pi}(x)} \right) = 0 . \tag{14.39}$$

This is a linear partial differential equation for Γ_1 which can be written symbolically as

$$\mathcal{D}_i \Gamma_1(\boldsymbol{\pi}, H) = 0 \tag{14.40}$$

with

$$\mathcal{D}_i = \int \mathrm{d}^d x \left(\frac{\delta \Gamma_0}{\delta \pi_i(x)} \frac{\delta}{\delta H(x)} + \frac{\delta \Gamma_0}{\delta H(x)} \frac{\delta}{\delta \pi_i(x)} \right) . \tag{14.41}$$

Let us calculate the commutator $[\mathcal{D}_i, \mathcal{D}_j]$. One first verifies that the space dependence plays no role and can thus be omitted. Then,

$$\begin{aligned}[]
[\mathcal{D}_i, \mathcal{D}_j] = {} & \left(\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial \pi_j \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial \pi_j} \right) \frac{\partial}{\partial H} + \left(\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial H \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial H} \right) \frac{\partial}{\partial \pi_j} \\
& - (i \leftrightarrow j).
\end{aligned}$$

Differentiating equation (14.35) with respect to π_j and H, respectively, we obtain

$$\begin{aligned}
\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial \pi_j \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial \pi_j} + H \delta_{ij} = 0 , \\
\frac{\partial \Gamma_0}{\partial \pi_i} \frac{\partial^2 \Gamma_0}{\partial H \partial H} + \frac{\partial \Gamma_0}{\partial H} \frac{\partial^2 \Gamma_0}{\partial \pi_i \partial H} + \pi_i = 0 .
\end{aligned}$$

We, thus, find

$$[\mathcal{D}_i, \mathcal{D}_j] = \int \mathrm{d}^d x \left(\pi_j(x) \frac{\delta}{\delta \pi_i(x)} - \pi_i(x) \frac{\delta}{\delta \pi_j(x)} \right) .$$

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 (14.39) is compatible (see also Section 13.1.1).

Renormalization. We now examine the large cut-off behaviour (or the behaviour when d approaches 2). Equation (14.39) is satisfied for all values of the regularizing parameter. We conclude that the divergent part Γ_1^{div} of Γ, defined in any minimal subtraction scheme, also satisfies equation (14.39):

$$\mathcal{D}_i \Gamma_1^{\text{div}} = 0 .$$

By adding to the action in the tree approximation $\mathcal{S}(\boldsymbol{\pi}, H)$, $-g\Gamma_1^{\text{div}}(\boldsymbol{\pi}, H)$ 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 (14.35) exactly. Indeed, at order 0, equation (14.35) is verified since $\mathcal{S}(\boldsymbol{\pi}, H)$ satisfies it. At order 1 equation (14.35) implies equation (14.39) 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 this 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 (14.35) exactly, and such that Γ_1, ...,Γ_{n-1} have been rendered finite. Then, as we have shown in Section 14.5.2, the generating functional $\Gamma(\boldsymbol{\pi}, H)$ renormalized up to order $(n-1)$ also satisfies equation (14.34). We write equation (14.34) symbolically:

$$\Gamma * \Gamma = K\,, \tag{14.42}$$

The nth order $(n > 0)$ in a loop expansion of equation (14.42) then takes the form

$$\sum_{p=0}^{n} \Gamma_p * \Gamma_{n-p} = 0 \tag{14.43}$$

and, therefore,

$$\Gamma_0 * \Gamma_n + \Gamma_n * \Gamma_0 = -\sum_{p=1}^{n-1} \Gamma_p * \Gamma_{n-p}\,. \tag{14.44}$$

The induction hypothesis implies that the r.h.s. is finite. The divergent part of the equation thus satisfies

$$\Gamma_0 * \Gamma_n^{\text{div}} + \Gamma_n^{\text{div}} * \Gamma_0 = 0\,. \tag{14.45}$$

The form of the equation is independent of n. We then define $\mathcal{S}_n$, the renormalized action at order n, by

$$\mathcal{S}_n = \mathcal{S}_{n-1} - g^n\Gamma_n^{\text{div}} + \sum_{n+1}^{\infty} g^p \delta\mathcal{S}_n^{(p)}\,. \tag{14.46}$$

It follows that

$$\begin{aligned}\mathcal{S}_n * \mathcal{S}_n - K &= \left(\mathcal{S}_{n-1} - g^n\Gamma_n^{\text{div}}\right) * \left(\mathcal{S}_{n-1} - g^n\Gamma_n^{\text{div}}\right) - K + O\left(g^{n+1}\right) \\ &= -g^n\left(\mathcal{S}_{n-1} * \Gamma_n^{\text{div}} + \Gamma_n^{\text{div}} * \mathcal{S}_{n-1}\right) + O\left(g^{n+1}\right).\end{aligned} \tag{14.47}$$

Since at this order in the r.h.s., $\mathcal{S}_{n-1}$ can be replaced by Γ_0, that is, $\mathcal{S}(\boldsymbol{\pi}, H)$, equation (14.45) then implies

$$\mathcal{S}_n * \mathcal{S}_n - K = O\left(g^{n+1}\right). \tag{14.48}$$

Hence, $\mathcal{S}_n$ satisfies equation (14.35) 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 (14.35) identically.

This completes the derivation. The renormalized 1PI functional satisfies equation (14.34), while the complete renormalized action satisfies equation (14.35).

To determine the form of the renormalized action, we now have to solve equation (14.35) taking into account 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 14.5.1 that the invariant measure for the transformations (14.36) is

$$\prod_x \mathrm{d}\boldsymbol{\pi}(x)\left[\frac{\delta \mathcal{S}}{\delta H(x)}\right]^{-1}.$$

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 in comparison with the vertices coming from the action. For instance, at one-loop order, they contribute under the form of their tree approximation. Therefore, if, once Γ_n is rendered finite, we modify the measure term by a divergent term of order g^n, to take into account the nth order renormalization, this will affect Γ_{n+1}, Γ_{n+2}, ... which are not yet renormalized, but leave Γ_n unchanged. It is, therefore, possible to introduce the field renormalization into the measure without changing the arguments given above about renormalization.

Actually, the measure term cancels at next order the potential quadratic divergences which could appear according to power counting. It prevents, at the same time, the generation of a mass by loop corrections

14.7 The Renormalized Action: Solution to the Master Equation

We now first solve the master equation (14.35) for the renormalized action and then show how the renormalizations appear order by order in perturbation theory.

14.7.1 The master equation

The general form of the renormalized action is obtained by solving equation (14.35):

$$\int \mathrm{d}^d x \left(\frac{\delta \mathcal{S}_\mathrm{r}}{\delta \boldsymbol{\pi}(x)}\frac{\delta \mathcal{S}_\mathrm{r}}{\delta H(x)} + H(x)\boldsymbol{\pi}(x)\right) = 0 .$$

Power counting tells us that the dimension $[\boldsymbol{\pi}]$ of $\boldsymbol{\pi}$ and $[H]$ of H are in two dimensions,

$$[\boldsymbol{\pi}] = 0 , \quad [H] = 2 .$$

Since the action density has dimension 2, the action has the general form

$$\mathcal{S}(\boldsymbol{\pi}, H) = \mathcal{S}_\mathrm{r}(\boldsymbol{\pi}) - \int \mathrm{d}^d x\, \sigma_\mathrm{r}\big(\boldsymbol{\pi}(x)\big) H(x), \tag{14.49}$$

a term quadratic in H having at least dimension 4. Moreover, the coefficient $\sigma_\mathrm{r}(\boldsymbol{\pi})$ is dimensionless and thus a derivative-free function of $\boldsymbol{\pi}(x)$ while $\mathcal{S}_\mathrm{r}(\boldsymbol{\pi})$ has dimension 2 and contains terms with at most two derivatives.

From the coefficient of $H(x)$ in equation (14.35) we obtain

$$\sigma_\mathrm{r}(\boldsymbol{\pi})\frac{\delta\sigma_\mathrm{r}}{\delta\boldsymbol{\pi}(x)}+\boldsymbol{\pi}(x)=0\,. \tag{14.50}$$

Since $\sigma_\mathrm{r}(\boldsymbol{\pi})$ is derivative-free and $O(N-1)$-invariant, the solution of equation (14.50) is simply

$$\sigma_\mathrm{r}^2\left(\boldsymbol{\pi}^2\right)+\boldsymbol{\pi}^2(x)=Z^{-1}. \tag{14.51}$$

This shows that $\sigma_\mathrm{r}(\boldsymbol{\pi})$ is the renormalized σ-field and Z is the field renormalization constant. We now write the equation for $H(x)=0$:

$$\int \mathrm{d}^d x\,\frac{\delta \mathcal{S}_\mathrm{r}(\boldsymbol{\pi})}{\delta\boldsymbol{\pi}(x)}\sigma_\mathrm{r}\left(\boldsymbol{\pi}(x)\right)=0\,. \tag{14.52}$$

This equation tells us that $\mathcal{S}_\mathrm{r}(\boldsymbol{\pi})$ is invariant under an infinitesimal transformation of the form

$$\delta\boldsymbol{\pi}(x)=\boldsymbol{\omega}\,\sigma_\mathrm{r}(\boldsymbol{\pi}),$$

or solving equation (14.51):

$$\delta\boldsymbol{\pi}(x)=\boldsymbol{\omega}\left(Z^{-1}-\boldsymbol{\pi}^2(x)\right)^{1/2}. \tag{14.53}$$

This is the renormalized form of the non-linear part of the $O(N)$ transformations. Equations (14.51) and (14.52) show that the renormalized functional $\mathcal{S}_\mathrm{r}(\boldsymbol{\pi})$ is $O(N)$-invariant, but the radius of the sphere has been renormalized. The renormalized action can, therefore, be written as

$$\mathcal{S}_\mathrm{r}(\boldsymbol{\pi},H)=\frac{1}{2Z_g}\int\left[(\partial_\mu\boldsymbol{\pi})^2+(\partial_\mu\sigma_\mathrm{r})^2\right]\mathrm{d}^d x-\int h\,\sigma_\mathrm{r}(x)\,\mathrm{d}^d x \tag{14.54}$$

with the relation

$$\sigma_\mathrm{r}(x)=\left[Z^{-1}-\boldsymbol{\pi}^2(x)\right]^{1/2}. \tag{14.55}$$

We have proven a rather remarkable result: although the interaction in the non-linear σ-model is non-polynomial, the theory can be renormalized with only two renormalization constants. We note that by giving a mass to the $\boldsymbol{\pi}$-field, through a term of the form $\int\sigma(x)\mathrm{d}^d x$, we have introduced no new renormalization constant. This would not have been the case for a term of the form $\int\boldsymbol{\pi}^2(x)\mathrm{d}^d x$, 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, of course, has no equivalent in the non-linear theory, and the structure of the mass renormalization. In particular, no spontaneous mass is generated.

14.7.2 Linearized WT identities

In our discussion of the renormalization of the model we have admitted without proof 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 solve explicitly the equations (14.39) or (14.45) satisfied by the divergent part of Γ, we first discuss the general solution of an equation of the form

$$\int \mathrm{d}^d x \left(\frac{\delta \mathcal{S}}{\delta \pi_i(x)} \frac{\delta}{\delta H(x)} + \frac{\delta \mathcal{S}}{\delta H(x)} \frac{\delta}{\delta \pi_i(x)} \right) \mathcal{O}(\boldsymbol{\pi}, H) = 0\,, \tag{14.56}$$

in which $\mathcal{O}(\boldsymbol{\pi}, H)$ is an arbitrary $O(N-1)$ symmetric local operator and $\mathcal{S}$ is the bare action:

$$\mathcal{S} = \int \mathrm{d}^d x \left\{ \tfrac{1}{2} \left[(\partial_\mu \boldsymbol{\pi})^2 + (\partial_\mu \sigma)^2 \right] - H(x)\sigma(x) \right\}, \tag{14.57}$$

with $\sigma(x) = \left(1 - \boldsymbol{\pi}^2(x)\right)^{1/2}$. Equation (14.56) reads, explicitly,

$$\int \mathrm{d}^d x \left\{ \left[-\nabla^2 \pi_i(x) + \alpha(x)\pi_i(x) \right] \frac{\delta}{\delta H(x)} - \sigma(x) \frac{\delta}{\delta \pi_i(x)} \right\} \mathcal{O}(\boldsymbol{\pi}, H) = 0\,, \tag{14.58}$$

in which we have defined

$$\alpha(x) = \frac{1}{\sigma(x)} \left[H(x) + \Delta\sigma(x) \right]. \tag{14.59}$$

The operator $\alpha(x)$ is an affine function of $H(x)$. We change variables in equation (14.58), $H \mapsto \alpha$, and consider $\mathcal{O}(\boldsymbol{\pi}, H)$ as a functional $\tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)$. A short but careful calculation then leads to a new equation:

$$\int \mathrm{d}^d x \, \sigma(x) \frac{\delta \tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)}{\delta \pi_i(x)} = 0\,, \tag{14.60}$$

which shows that the operator $\tilde{\mathcal{O}}(\boldsymbol{\pi}, \alpha)$ is $O(N)$ symmetric at $\alpha(x)$ fixed.

This result has two applications: it completes our proof and, as we discuss below, it yields the renormalized form of a general $O(N)$-invariant operator.

From power counting we know that Γ_n^{div} has dimension 2. It has, therefore, terms of degree 0 and 1 in α. According to the previous result it has the form

$$\Gamma_n^{\text{div}} = \frac{1}{2} a_n \int \left\{ \left[\partial_\mu \boldsymbol{\pi}(x) \right]^2 + \left[\partial_\mu \sigma(x) \right]^2 \right\} \mathrm{d}^d x + \frac{1}{2} b_n \int \alpha(x) \, \mathrm{d}^d x\,. \tag{14.61}$$

The first term can be absorbed into a coupling constant renormalization:

$$g \mapsto g\,(1 + g^n a_n)\,. \tag{14.62}$$

We now calculate the variation of the action when the radius of the sphere is renormalized. The variation $\delta\sigma$ of $\sigma(x)$ is

$$\delta\sigma(x) = \left[1 - \delta Z - \boldsymbol{\pi}^2(x) \right]^{1/2} - \left[1 - \boldsymbol{\pi}^2(x) \right]^{1/2} = -\tfrac{1}{2} \delta Z / \sigma(x).$$

It follows that

$$\delta \left\{ \int \mathrm{d}^d x \left[\tfrac{1}{2} \left[\partial_\mu \sigma(x) \right]^2 - H(x)\sigma(x) \right] \right\} = \frac{1}{2} \delta Z \int \alpha(x) \, \mathrm{d}^d x\,. \tag{14.63}$$

Therefore, the second term can be absorbed in a field renormalization:

$$\delta Z = b_n g^n. \tag{14.64}$$

This completes altogether our proof of the renormalization of the non-linear σ-model. Nevertheless, for completeness we outline in Section 14.9 a different derivation, whose basic idea is to return to a linear formulation of the symmetry.

14.8 Renormalization of Composite Operators

General $O(N)$-invariant operators. By solving the equation (14.56) in its most general form, we have also obtained the structure of renormalized $\boldsymbol{\phi}(x)$ correlation functions with one insertion of an $O(N)$ invariant-operator of arbitrary dimension. Indeed, to generate such insertions we can add to the action a source term of the form $\int \mathrm{d}^d x\, K(x)\mathcal{O}(\boldsymbol{\pi}(x))$. Since $\mathcal{O}(\boldsymbol{\pi})$ is $O(N)$ invariant, equation (14.35) holds for the complete action which includes this new term. An expansion of equation (14.35) at first order in $K(x)$ leads to equation (14.56) 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 $[\mathcal{O}]$ and $O(N)$ invariant at $\alpha(x)$ fixed.

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. To each different irreducible representation is associated a new renormalization constant. For example, a mass term $\int \mathrm{d}^d x\, \boldsymbol{\pi}^2(x)$ is a component of the symmetric traceless tensor

$$\int \mathrm{d}^d x \left[\phi_i(x)\phi_j(x) - \tfrac{1}{N}\delta_{ij}\boldsymbol{\phi}^2(x)\right],$$

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 are related to $\boldsymbol{\pi}(x)$ by

$$\boldsymbol{\theta}(x) = \boldsymbol{\pi}(x) f\left(\boldsymbol{\pi}^2\right).$$

We rewrite this expansion in terms of the $(i, 0\ldots 0)$ 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:

$$\boldsymbol{\theta}(x) = \boldsymbol{\pi}(x)\sum c_\ell P_\ell^N\left(\sigma(x)\right).$$

Each non-vanishing coefficient c_ℓ introduces an independent renormalization constant. 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:

$$\boldsymbol{\theta}(x) = \boldsymbol{\theta}_{\mathrm{r}}(x) Z\left[\boldsymbol{\theta}_{\mathrm{r}}^2(x)\right].$$

14.9 A Linear Representation

For all 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 specially easy since only one $O(N)$-invariant additional field, which we denote by $\alpha(x)$, is required to implement the constraint (14.1):

$$\boldsymbol{\phi}^2(x) = 1.$$

The integral representation (14.27) of the generating functional $\mathcal{Z}(\mathbf{J})$ can also be written as

$$\mathcal{Z}(\mathbf{J}) = \int [\mathrm{d}\alpha\, \mathrm{d}\boldsymbol{\phi}] \exp\left[-\frac{1}{g}\left(\mathcal{S}(\boldsymbol{\phi},\alpha) - \int \mathrm{d}^d x\, \mathbf{J}(x)\cdot\boldsymbol{\phi}(x)\right)\right], \tag{14.65}$$

where the integration contour for α is parallel to the imaginary axis, and

$$\mathcal{S}(\boldsymbol{\phi},\alpha) = \frac{1}{2}\int \mathrm{d}^d x \left\{\left[\partial_\mu \boldsymbol{\phi}(x)\right]^2 + \alpha(x)\left[\boldsymbol{\phi}^2(x) - 1\right]\right\}. \tag{14.66}$$

Note that, here, $\mathbf{J}(x)$ is an N-component source for the N-component field $\boldsymbol{\phi}(x)$.

We choose an extremum of the potential as the starting point for perturbation theory:

$$\alpha(x) = 0\,, \qquad \sigma(x) \equiv \phi_1(x) = 1\,, \qquad \phi_i(x) = 0 \quad \text{for } i > 1\,. \tag{14.67}$$

The propagator of the $(N-1)$ remaining components $\boldsymbol{\pi}(x)$ is still proportional to $1/p^2$. The component $\sigma(x)$ is coupled to $\alpha(x)$. At leading order, the connected two-point functions are

$$\begin{pmatrix} \langle\sigma\sigma\rangle & \langle\sigma\alpha\rangle \\ \langle\alpha\sigma\rangle & \langle\alpha\alpha\rangle \end{pmatrix}_{\mathrm{c}} (p) = \begin{pmatrix} 0 & g \\ g & gp^2 \end{pmatrix}. \tag{14.68}$$

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

$$[\boldsymbol{\phi}] = 0\,, \qquad [\alpha] = 2\,.$$

The general form of the renormalized action $\mathcal{S}_{\mathrm{r}}(\boldsymbol{\phi},\alpha)$ follows:

$$\begin{aligned}\mathcal{S}_{\mathrm{r}}(\boldsymbol{\phi},\alpha) = \frac{1}{2}\int \mathrm{d}^d x \Big\{ & B_1\left(\boldsymbol{\phi}^2(x)\right)\left[\partial_\mu\boldsymbol{\phi}(x)\right]^2 + B_2\left(\boldsymbol{\phi}^2(x)\right)\left(\boldsymbol{\phi}\cdot\partial_\mu\boldsymbol{\phi}\right)^2 \\ & + B_3\left(\boldsymbol{\phi}^2(x)\right)\alpha(x) - B_4\left(\boldsymbol{\phi}^2(x)\right)\Big\},\end{aligned} \tag{14.69}$$

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 we are only interested in the $\boldsymbol{\phi}$-field correlation functions (and not the α-field), we can explicitly integrate over the α-field. This fixes the value of $\boldsymbol{\phi}^2(x)$ to a solution of the equation

$$B_3\left(\boldsymbol{\phi}^2(x)\right) = 0\,. \tag{14.70}$$

This equation has a solution order by order in perturbation theory:

$$\boldsymbol{\phi}^2(x) = Z^{-1}(g) = 1 + O(g)\,. \tag{14.71}$$

After integration, B_1, B_2 and B_3 become pure constants and $\boldsymbol{\phi}\cdot\partial_\mu\boldsymbol{\phi}$ vanishes identically.

We, therefore, recover 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, we understand that the insertions of operators of dimension larger than two will eventually generate terms of degree larger than one in $\alpha(x)$. The integration over α then no longer implies the strict constraint (14.71).

Bibliographical Notes

For early references on the σ-model see

M. Gell-Mann and M. Lévy, *Nuovo Cimento* 16 (1960) 705; S. Weinberg, *Phys. Rev.* 166 (1968) 1568; J. Schwinger, *Phys. Rev.* 167 (1968) 1432; W.A. Bardeen and B.W. Lee, *Nuclear and Particle Physics*, Montreal 1967, B. Margolis and C.S. Lam eds. (Benjamin, New York 1969); K. Meetz, *J. Math. Phys.* 10 (1969) 589.

The quantization was discussed in

I.S. Gerstein, R. Jackiw, B.W. Lee and S. Weinberg, *Phys. Rev.* D3 (1971) 2486; J. Honerkamp and K. Meetz, *Phys. Rev.* D3 (1971) 1996; J. Honerkamp, *Nucl. Phys.* B36 (1972) 130; A.A. Slavnov and L.D. Faddeev, *Theor. Math. Phys.* 8 (1971) 843.

Pauli–Villars's regularization for chiral models has been proposed by

A.A. Slavnov, *Nucl. Phys.* B31 (1971) 301.

The model was studied as a limit of the linear σ-model in

D. Bessis and J. Zinn-Justin, *Phys. Rev.* D5 (1972) 1313.

The renormalization group properties were discussed in

A.M. Polyakov, *Phys. Lett.* 59B (1975) 79; E. Brézin and J. Zinn-Justin, *Phys. Rev. Lett.* 36 (1976) 691; *Phys. Rev.* B14 (1976) 3110; W.A. Bardeen, B.W. Lee and R.E. Shrock, *Phys. Rev.* D14 (1976) 985.

The renormalizability in two dimensions was proved in

E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D14 (1976) 2615; B14 (1976) 4976.

The Elitzur conjecture has been proved in

F. David, *Commun. Math. Phys.* 81 (1981) 149.

15 GENERAL NON-LINEAR MODELS IN TWO DIMENSIONS

In this chapter we describe the formal properties and discuss the renormalization of a class of geometric models: models based on homogeneous spaces. Homogeneous spaces are associated with non-linear realizations of group representations and these models are natural generalizations of the non-linear σ-model considered in Chapter 14. 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. This is the system of coordinates that we have used in the discussion of the non-linear σ-model and again use in the first part of this chapter because the renormalization properties are simpler and the physical interpretation of correlation functions more direct. We then also examine some properties of these models in a generic parametrization. The renormalization problem is solved by the introduction of a symmetry (generally called BRS symmetry) with anticommuting (Grassmann) parameters which, later, will play an essential role in the renormalization of gauge theories.

In a second part of the chapter we study the more specific properties of models corresponding to a special class of homogeneous spaces: symmetric spaces. The non-linear σ-model of Chapter 14 provides the simplest example. These models are characterized by the uniqueness of the metric and thus of the classical action. A consequence of the classical field equations, in two dimensions 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 RG functions at one-loop order, and find the first examples of UV asymptotic freedom.

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.

Note, finally, that 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.

15.1 Homogeneous Spaces and Goldstone Modes

Rather than giving a purely mathematical construction, we shall try to motivate the study of these models by some physical arguments.

15.1.1 Spontaneous symmetry breaking: Goldstone modes

Notation. We consider a field $\phi_i(x)$ transforming under a linear representation of a group G. We use the notation of Chapter 13 for the infinitesimal group transformation: t^α are the generators of the Lie algebra $\mathcal{L}(G)$ in some matrix representation, and ω_α the infinitesimal group parameters of the transformation:

$$\delta\phi_i(x) = t^\alpha_{ij}\omega_\alpha\phi_j(x). \tag{15.1}$$

Since we consider only compact groups, we can assume that the representation is orthogonal, and that the generators of $\mathcal{L}(G)$ are $N \times N$ antisymmetric matrices.

Goldstone modes. Our starting point is, as in Chapter 13, a G-invariant action $\mathcal{S}(\phi)$ of the form

$$\mathcal{S}(\phi) = \int \mathrm{d}x \left\{ \frac{1}{2} [\partial_\mu \phi(x)]^2 + V\big(\phi(x)\big) \right\}. \tag{15.2}$$

We assume that in the classical (tree) approximation the model exhibits the phenomenon of spontaneous symmetry breaking (SSB), that is, $V(\phi)$ has non-G-invariant global minima. We distinguish one of them, ϕ^c, around which we expand perturbation theory, and call H the isotropy or little group (the stabilizer) of ϕ^c. We have shown in Section 13.4 that under these circumstances a number of field components of $\phi(x)$, which correspond to the Goldstone modes of the broken symmetry, are massless.

If we are only interested in the long distance behaviour (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 functional integral can be neglected. In this limit, the remaining massless components of the field $\phi(x)$ can be entirely parametrized in terms of a matrix $\mathbf{R}(g(x))$ of the representation of G, acting on the vector ϕ_c

$$\phi_i(x) = R_{ij}\big(g(x)\big)\phi^c_j\,, \qquad g(x) \in G\,. \tag{15.3}$$

Note that if we multiply $g(x)$ on the right by an element $h(x)$ of H, since ϕ^c_j is left invariant by the group H, $\phi_i(x)$ is not modified:

$$\phi_i(x) = R_{ij}\big(g(x)h(x)\big)\phi^c_j = R_{ik}\big(g(x)\big)R_{kj}\big(h(x)\big)\phi^c_j = R_{ij}\big(g(x)\big)\phi^c_j\,.$$

This shows that $\phi_i(x)$ is really a function of the elements of the coset 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 $\mathcal{L}(H)$, $\{t^\alpha\}$, $\alpha > l$, and the complementary set, which we denote by $\mathcal{L}(G/H)$ of generators $\{t^\alpha\}$, $1 \le \alpha \le l$, which is such that

$$\textstyle\sum_{\alpha=1}^{l} c_\alpha t^\alpha_{ij}\phi^c_j = 0 \quad\Rightarrow\quad c_\alpha = 0\,.$$

The matrix $\mathbf{R}$ can be canonically parametrized in terms of fields $\xi_a(x)$ as

$$\mathbf{R}\big(\xi_a(x)\big) = \exp\left(\sum_{a=1}^{l} \xi_a(x) t^a\right). \tag{15.4}$$

15.1.2 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. The action $\mathcal{S}(\phi)$ can then be written as

$$\mathcal{S}(\phi) = \frac{1}{2}\int \mathrm{d}x\, \phi^c_j \partial_\mu R^{-1}_{jk}\big(g(x)\big)\partial_\mu R_{ki}\big(g(x)\big)\phi^c_i\,. \tag{15.5}$$

By expanding action (15.5) for $g(x)$ close to the identity we indeed verify that all remaining fields are massless.

Notation. It will also be convenient to use a bra and ket notation to indicate vectors, denoting by $|0\rangle$ the vector $\boldsymbol{\phi}^c$. Equation (15.3) then takes the form

$$|\phi(x)\rangle = \mathbf{R}(g(x))\,|0\rangle\,, \tag{15.6}$$

in which $|\phi(x)\rangle$ is a notation for the field $\phi_i(x)$. With this notation, the classical action $\mathcal{S}(\boldsymbol{\phi})$ can be rewritten in terms of $\mathbf{R}^{-1}\partial_\mu\mathbf{R}$:

$$\mathcal{S}(\boldsymbol{\phi}) = -\frac{1}{2}\int \mathrm{d}x\,\langle 0|\left(\mathbf{R}^{-1}\partial_\mu\mathbf{R}\right)^2|0\rangle\,. \tag{15.7}$$

Any matrix of the form $\mathbf{R}^{-1}\partial\mathbf{R}$ (a pure gauge) belongs to the Lie algebra of G.

15.1.3 Metric and action in general coordinates

In this section, we use a notation and conventions adapted to Riemannian geometry (see Chapter 22 for details).

We denote by φ^i an arbitrary set of coordinates on the manifold (coset space) G/H. The action (15.5) can then be written in the form

$$\mathcal{S}(\varphi) = \frac{1}{2}\int \mathrm{d}x\, g_{ij}\left(\varphi(x)\right)\partial_\mu\varphi^i(x)\partial_\mu\varphi^j(x), \tag{15.8}$$

where g_{ij} is a metric on G/H considered as a Riemannian manifold. We know from the discussion of Section 15.1.1 that the matrix $\mathbf{R}$ when acting on the vector $|0\rangle$ can be parametrized in the form (15.4). The variables ξ^a in equation (15.4) thus provide an example of a set of coordinates on the coset space G/H.

Comparing the expressions (15.8) and (15.7), we can relate the metric to other geometric objects. The current $\mathbf{R}^{-1}\partial_\mu\mathbf{R}$ can be written as

$$\mathbf{R}^{-1}\partial_\mu\mathbf{R} = \mathbf{R}^{-1}\frac{\partial\mathbf{R}}{\partial\varphi^i}\partial_\mu\varphi^i(x). \tag{15.9}$$

In what follows ∂_i with roman indices means derivative with respect to φ^i and ∂_μ with greek indices means derivative with respect to x_μ.

We can expand $\mathbf{R}^{-1}\partial_\mu\mathbf{R}$ on a basis of generators of the Lie algebra of G (more details can be found in Appendix A15.1):

$$\mathbf{R}^{-1}\frac{\partial\mathbf{R}}{\partial\varphi^i} \equiv \mathbf{R}^{-1}\partial_i\mathbf{R} = L_i^\alpha(\varphi)t^\alpha. \tag{15.10}$$

and, therefore,

$$\mathbf{R}^{-1}\partial_\mu\mathbf{R} = L_i^\alpha(\varphi)t^\alpha\partial_\mu\varphi^i(x). \tag{15.11}$$

An expression of the metric tensor in terms of $L_i^\alpha(\varphi)$ follows:

$$g_{ij}(\varphi) = L_i^a(\varphi)L_j^b(\varphi)\mu_{ab} \tag{15.12}$$

with the definition

$$\mu_{ab} = -\left\langle 0\left|t^a t^b\right|0\right\rangle. \tag{15.13}$$

The latin letters (a,b) indicate that the indices run only over values corresponding to generators belonging to $\mathcal{L}(G/H)$.

One verifies that the quantities L_i^α are independent of the representation (Section A15.1). Therefore, the dependence in 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:

$$\langle 0 \left| t^a \tau^\gamma t^b \right| 0 \rangle = \langle 0 \left| t^a \left[\tau^\gamma , t^b \right] \right| 0 \rangle = \langle 0 \left| \left[t^a , \tau^\gamma \right] t^b \right| 0 \rangle ,$$

for $t^a, t^b \in \mathcal{L}(G/H)$, $t^\gamma \in \mathcal{L}(H)$ and, therefore,

$$\mu_{ac} f_{c\gamma b} - f_{a\gamma c} \mu_{cb} = 0 . \tag{15.14}$$

The interpretation of this equation is simple: the vectors $t^a |0\rangle$ transform under an (in general reducible) representation of H which has the matrices $f_{a\gamma b}$ as generators.

The symmetric matrix $\boldsymbol{\mu}$ commutes with all generators $\mathbf{f}^\gamma$ ($[\mathbf{f}^\gamma]_{ab} = f_{a\gamma b}$) 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 $t^a |0\rangle$ (Schur's lemma).

This result exhausts all consequences of the G-symmetry.

15.1.4 Quantization and perturbation theory near dimension 2

So far we have only examined the classical theory. To quantize it, we can start from the euclidean action (15.5), derive a quantum mechanical hamiltonian and use canonical quantization. It is faster to begin with action (15.2) and freeze the massive degrees of freedom as explained above. The result is the same and yields a functional representation:

$$\mathcal{Z} = \int \left[\det{}^{1/2} g_{ij}(\varphi) \mathrm{d}(\varphi) \right] \exp \left[-\frac{1}{2} \int \mathrm{d}x \, g_{ij} \left(\varphi(x) \right) \partial_\mu \varphi^i(x) \partial_\mu \varphi^j(x) \right] . \tag{15.15}$$

The measure of the φ-integration is 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 action (15.5), we derive in the next sections a set of WT identities corresponding to the non-linearly realized symmetry under the group G. Dimensional or lattice regularization respecting the G-symmetry is implied in what follows. The choice of a group-invariant regularization corresponds also to a choice of quantization of the classical hamiltonian consistent with the symmetry.

15.2 WT Identities and Renormalization in Linear Coordinates

For renormalization problems it is convenient to choose a special parametrization, which consists in embedding the homogeneous space into a euclidean space, as we have explained in Chapter 14.

Remark. Note that the vector space $\mathcal{V}$ spanned by the family of vectors of the form (15.3) may have lower dimension than the original space of representation to which $\boldsymbol{\phi}$ 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* $\boldsymbol{\phi}$ to its components in $\mathcal{V}$.

15.2.1 Linear coordinates

We consider the vector space $\mathcal{V}$ spanned by the vectors of form (15.3), as well as the subspace $\mathcal{V}'$ of dimension l spanned by the vectors $t^{\alpha}_{ij}\phi^c_j$. We choose an orthogonal basis in $\mathcal{V}$ such that the vectors of $\mathcal{V}'$ have only the first l components non-vanishing. We then distinguish the l first components of the vector $\boldsymbol{\phi}$ of equation (15.3), calling them $\pi_a(x)$, and the others called $\sigma_i(x)$:

$$\boldsymbol{\phi}(x) = \begin{cases} \pi_a(x) , & 1 \le a \le l \,; \\ \sigma_i(x) , & l < i \,. \end{cases} \tag{15.16}$$

The fields $\sigma_i(x)$ and $\pi_a(x)$ are functions of the fields $\xi_a(x)$. If we expand equation (15.4) in powers of ξ,

$$\phi_i(x) = \phi^c_i + \sum_{a=1}^{l} \xi_a(x) t^a_{ij} \phi^c_j + O\left(\xi^2\right) , \tag{15.17}$$

we see, comparing expressions (15.3) and (15.16), that $\pi_a(x)$ and $\xi_a(x)$ are canonically related:

$$\pi_a(x) = \xi_b(x) t^b_{aj} \phi^c_j + O\left(\xi^2\right) , \tag{15.18}$$

that is, this relation can be inverted to express the fields ξ_a as functions of the fields π_a. On the other hand, the fields $\sigma_i(x)$ are of the form

$$\sigma_i(x) = \phi^c_i + O\left(\xi^2\right) ,$$

and can, therefore, 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 $\{t^\alpha\}$ mix the fields $\pi_a(x)$ and the $\sigma_i(x)$ which are functions of π_a:

$$\delta\pi_a = \left[t^\alpha_{ab}\pi_b + t^\alpha_{aj}\sigma_j(\pi)\right] \omega_\alpha \,. \tag{15.19}$$

Note that since the σ_i are functions of π_a, the transformation laws

$$\delta\sigma_i(\pi) = \left[t^\alpha_{ib}\pi_b + t^\alpha_{ij}\sigma_j(\pi)\right] \omega_\alpha \,, \tag{15.20}$$

are now consequences of equation (15.19) and the functional form of the $\sigma_i(\pi)$.

15.2.2 Correlation functions, WT identities, renormalization

We consider the generating functional of correlation functions:

$$\mathcal{Z}(J) = \int \left[\det{}^{1/2} g_{ij}(\pi) \mathrm{d}(\pi)\right] \exp\left[-\mathcal{S}(\boldsymbol{\phi}) + \int \mathrm{d}x\, J_i(x)\phi_i(x)\right] , \tag{15.21}$$

the integrand being expressed in terms of these special coordinates. Sources have been added for all components of $\boldsymbol{\phi}$ in $\mathcal{V}$ for the reasons which have already been explained in the preceding chapter: when we try to derive WT identities expressing the consequences of the symmetry for correlation functions, the composite σ_a operators 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 at fixed constant values of the σ sources.

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 action (15.5).

WT identities. We perform an infinitesimal transformation (15.19) in the functional integral (15.21). 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:

$$\int \mathrm{d}x\, J_i(x) t^\alpha_{ij} \frac{\delta \mathcal{Z}(J)}{\delta J_j(x)} = 0\,, \tag{15.22}$$

and, thus,

$$\int \mathrm{d}x\, J_i(x) t^\alpha_{ij} \frac{\delta W}{\delta J_j(x)} = 0\,. \tag{15.23}$$

Again the difference appears in the Legendre transformation. Let us now call $J_a(x)$, $1 \le a \le l$, the sources for the fields π_a, and H_i, $l < i \le n$, the sources for the composite fields σ_i. The Legendre transform has to be performed only on J_a:

$$\Gamma(\pi, H) + \mathcal{W}(J, H) = \int \mathrm{d}x\, \pi_a(x) J_a(x), \tag{15.24}$$

$$\pi_a(x) = \frac{\delta W}{\delta J_a(x)} \Leftrightarrow J_a(x) = \frac{\delta \Gamma}{\delta \pi_a(x)}. \tag{15.25}$$

Since, as explained several times,

$$\frac{\delta \Gamma}{\delta H_i(x)} = -\frac{\delta \mathcal{W}}{\delta H_i(x)},$$

identity (15.23) implies for the 1PI functional Γ

$$\int \mathrm{d}x \left\{ \frac{\delta \Gamma}{\delta \pi_a(x)} \left[t^\alpha_{ab}\pi_b(x) - t^\alpha_{aj} \frac{\delta \Gamma}{\delta H_j(x)} \right] + H_i(x) \left[t^\alpha_{ib}\pi_b(x) - t^\alpha_{ij} \frac{\delta \Gamma}{\delta H_j(x)} \right] \right\} = 0\,. \tag{15.26}$$

Renormalization. This identity has a quadratic form as in the case of the σ-model. The arguments of Section 14.6 apply and prove the stability of the identity under renormalization. The same identity is thus fulfilled by the renormalized action $\mathcal{S}_\mathrm{r}$:

$$\int \mathrm{d}x \left\{ \frac{\delta \mathcal{S}_\mathrm{r}}{\delta \pi_a(x)} \left[t^\alpha_{ab}\pi_b(x) - t^\alpha_{aj} \frac{\delta \mathcal{S}_\mathrm{r}}{\delta H_j(x)} \right] + H_i(x) \left[t^\alpha_{ib}\pi_b(x) - t^\alpha_{ij} \frac{\delta \mathcal{S}_\mathrm{r}}{\delta H_j(x)} \right] \right\} = 0\,. \tag{15.27}$$

Power counting tells us that $\mathcal{S}_\mathrm{r}(\pi, H)$ and $H_a(x)$ have dimension 2. Therefore, again $\mathcal{S}_\mathrm{r}$ is linear in the source $H_a(x)$. Let us write $\mathcal{S}_\mathrm{r}$ as

$$\mathcal{S}_\mathrm{r}(\pi, H) = -\int \sigma_{\mathrm{r}i}(\pi) H_i(x) \mathrm{d}x + \Sigma_\mathrm{r}(\pi), \tag{15.28}$$

in which the functions $\sigma_{\mathrm{r}i}(\pi)$ are derivative-free and Σ_r has dimension 2, that is, has at most two derivatives. The term linear in H in equation (15.27) yields

$$\frac{\delta \sigma_{\mathrm{r}i}(\pi)}{\delta \pi_a(x)} \left[t^\alpha_{ab}\pi_b(x) + t^\alpha_{aj}\sigma_{\mathrm{r}j}\left(\pi(x)\right) \right] = t^\alpha_{ib}\pi_b(x) + t^\alpha_{ij}\sigma_{\mathrm{r}j}\left(\pi(x)\right). \tag{15.29}$$

These partial differential equations for the function $\sigma_{\mathrm{r}i}(\pi)$ imply that if the fields π_a have the transformation law

$$\delta_\alpha \pi_a = t^\alpha_{ab}\pi_b + t^\alpha_{aj}\sigma_{\mathrm{r}j}(\pi), \tag{15.30}$$

then, as a consequence,

$$\delta_\alpha \sigma_i(\pi) = t^\alpha_{ib}\pi_b + t^\alpha_{ij}\sigma_{\mathrm{r}j}(\pi)\,. \tag{15.31}$$

Thus, the field $\boldsymbol{\phi}$ with component $(\pi_a, \sigma_{\mathrm{r}i})$ transforms under a linear representation of the group G. We now write the equation obtained by setting H equal to zero:

$$\int \mathrm{d}x \frac{\delta \Sigma_{\mathrm{r}}}{\delta \pi_a(x)} \left[t^\alpha_{ab}\pi_b(x) + t^\alpha_{aj}\sigma_{\mathrm{r}j}(x)\right] = 0\,. \tag{15.32}$$

This equation tells us that $\Sigma_{\mathrm{r}}(\pi)$ 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 written in terms of the field $\boldsymbol{\phi}(x)$, and then eliminating the fields $\sigma_{\mathrm{r}i}(x)$. Also, we have shown in Section 15.1.3 that the action, expressed in terms of the metric tensor, can be parametrized in terms of the matrix $\boldsymbol{\mu}$ (equation (15.13)). Here, we find that even if in the tree approximation we begin with a special choice of the matrix $\boldsymbol{\mu}$, solution of equation (15.14), we obtain after renormalization the most general solution of this equation.

15.2.3 Field renormalizations

Here, we briefly comment about the solutions of equation (15.29). For renormalization purposes, we are looking for *generic* solutions $\sigma_i(\pi)$ expandable in powers of π:

$$\sigma_i = S_i + S_i^{a_1}\pi_{a_1} + \tfrac{1}{2}S_i^{a_1 a_2}\pi_{a_1}\pi_{a_2} + \cdots\,, \tag{15.33}$$

in which S_i, $S_i^{a_1}\ldots$ are constants.

The equation of order zero in π is

$$S_i^a t^\alpha_{aj} S_j = t^\alpha_{ij} S_j\,. \tag{15.34}$$

If t^α belongs to the Lie algebra $\mathcal{L}(H)$, t^α_{aj} vanishes since π and σ belong to different representations and, therefore,

$$t^\alpha_{ij} S_j = 0 \quad \text{for all} \quad t^\alpha \in \mathcal{L}(H). \tag{15.35}$$

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 $t^\alpha_{aj}S_j$ spans the π subspace. Therefore, equation (15.34) implies

$$S_i^a t^\alpha_{aj} S_j = t^\alpha_{ij} S_j\,, \qquad 1 \le \alpha \le l\,. \tag{15.36}$$

The scalar product of the vectors $\mathbf{S}_i$ with all vectors of a complete basis is known, therefore, all $\mathbf{S}_i$ are determined. In particular, if the π subspace contains no vector invariant under the action of the group H, the r.h.s. of equation (15.36) and, therefore, also all vectors $\mathbf{S}_i$ 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

$$S_i^{ab_1\ldots b_k}\ldots t^\alpha_{aj}S_j = \text{ known quantities}\,.$$

The coefficients $S_i^{ab_1\ldots b_k}$ are completely determined since considered as vectors $\mathbf{S}_i^{b_1\ldots b_k}$ 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. We have, therefore, enumerated all the renormalization constants of the model.

15.3 Renormalization in Arbitrary Coordinates, BRS Symmetry

We have shown in Section 15.2 how a special choice of parametrization leads to a rather simple discussion of the symmetry properties and renormalization of correlation functions.

We have indicated in Chapter 14, in the case of the non-linear σ-model, that with another choice we would have been forced to introduce 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 function of the renormalized field.

However, as we have indicated in Appendix A7.2, in some situations, only parametrization-independent quantities (related to geometric properties of the manifold or to the S-matrix) are physical. It is, therefore, 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 which 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. We therefore introduce a new method which, in some generalized form, will also be relevant to the renormalization of gauge theories and which is based on infinitesimal group transformations with anticommuting parameters.

15.3.1 Infinitesimal group transformations

We consider a non-linear realization of the representation of a group G acting on a field $\varphi^i(x)$. We write the infinitesimal group transformations corresponding to parameters ω^α in the form

$$\delta_\omega\varphi^i(x) = D^i_\alpha\left(\varphi(x)\right)\omega^\alpha\,. \qquad (15.37)$$

We assume that the functions $D^i_\alpha(\varphi)$ are smooth, that is, infinitely differentiable, at $\varphi = 0$. We write the action in the form (15.8):

$$\mathcal{S}(\varphi) = \frac{1}{2}\int \mathrm{d}x\, g_{ij}\left(\varphi(x)\right)\partial_\mu\varphi^i(x)\partial_\mu\varphi^j(x). \qquad (15.38)$$

The invariance of the action $\mathcal{S}(\varphi)$ under the transformations (15.37) implies

$$\int \mathrm{d}x\, D^i_\alpha\bigl(\varphi(x)\bigr)\frac{\delta\mathcal{S}(\varphi)}{\delta\varphi^i(x)}=0\,, \tag{15.39}$$

and, therefore, for the metric tensor $g_{ij}(\varphi)$,

$$D^i_\alpha\frac{\partial g_{jk}}{\partial\varphi^i}+g_{ik}\frac{\partial D^i_\alpha}{\partial\varphi^j}+g_{ij}\frac{\partial D^i_\alpha}{\partial\varphi^k}=0\,. \tag{15.40}$$

Solving this equation is equivalent to finding all possible metric tensors on a given homogeneous space, compatible with the group structure. The Appendix A15.2 contains some details about the nature and structure of this equation.

The functions $D^i_\alpha(\varphi)$ satisfy identities which can be obtained either by direct calculation or by noting that the differential operators $\mathcal{D}_\alpha$,

$$\mathcal{D}_\alpha=D^i_\alpha(\varphi)\frac{\partial}{\partial\varphi^i}, \tag{15.41}$$

acting on functions of φ, form themselves a representation of the Lie algebra (see also Section 13.1.1) and, therefore, have commutation relations of the form

$$[\mathcal{D}_\alpha,\mathcal{D}_\beta]=f^\gamma_{\alpha\beta}\mathcal{D}_\gamma\,. \tag{15.42}$$

Note that these commutation relations are compatibility conditions for equation (15.39) considered as a set of linear differential equations for $\mathcal{S}(\varphi)$. Calculating explicitly the commutator in terms of the functions $D^i_\alpha(\varphi)$ we obtain

$$D^j_\alpha(\varphi)\frac{\partial D^i_\beta(\varphi)}{\partial\varphi^j}-D^j_\beta(\varphi)\frac{\partial D^i_\alpha(\varphi)}{\partial\varphi^j}=f^\gamma_{\alpha\beta}D^i_\gamma(\varphi). \tag{15.43}$$

This is the form of the commutation relations of the Lie algebra which is useful in the discussion of non-linear representations. Moreover, in Chapter 21 we shall encounter equations which are formally identical and play an essential role in the discussion of the renormalization of gauge theories.

15.3.2 WT identities

We assume that we have introduced a group-invariant regularization and consider the generating functional $\mathcal{Z}(J,K)$:

$$\mathcal{Z}(J,K)=\int[\mathrm{d}\rho(\varphi)]\exp\left\{-\mathcal{S}(\varphi)+\int\mathrm{d}x\,\left[K(x)A\bigl(\varphi(x)\bigr)+J_i(x)\varphi^i(x)\right]\right\}. \tag{15.44}$$

The measure $[\mathrm{d}\rho(\varphi)]$ is the group-invariant measure. Note that we have added to the action, sources not only for the field $\varphi^i(x)$ but also for a local derivative-free function of $\varphi(x)$, $A(\varphi(x))$. The function $A(\varphi)$ has to satisfy only 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 (15.39), we introduce a set of anticommuting constants C^α and $\bar\varepsilon$, that is, all belonging to a *Grassmann algebra*, and perform a transformation

$$\varphi^i(x) \mapsto \varphi^i(x) + \bar\varepsilon D^i_\alpha\big(\varphi(x)\big)C^\alpha . \tag{15.45}$$

The action and the measure are invariant. The variation of the source terms in expression (15.44) is

$$\delta\left\{\int \mathrm{d}x\,\left[J_i(x)\varphi^i(x) + K(x)A\left(\varphi(x)\right)\right]\right\} = \bar\varepsilon\int \mathrm{d}x\left[J_i(x) + K(x)\frac{\delta A}{\delta\varphi^i(x)}\right]D^i_\alpha(\varphi)C^\alpha . \tag{15.46}$$

This variation involves two composite operators: $D^i_\alpha\left(\varphi(x)\right)C^\alpha$, because the transformation is non-linear, and $\left(\delta A/\delta\varphi^i\right)D^i_\alpha(\varphi)C^\alpha$. In accordance with our general strategy, we introduce for them two anticommuting sources $\Lambda_i(x)$ and $L(x)$. Note that if we assign a fermion number $+1$ to C^α and -1 to $\Lambda_i(x)$ and $L(x)$, this fermion number is conserved. We then calculate the variation of these operators under the transformation (15.45):

$$\delta\left[D^i_\alpha(\varphi)C^\alpha\right] = \bar\varepsilon\frac{\delta D^i_\alpha}{\delta\varphi^j}D^j_\beta C^\beta C^\alpha . \tag{15.47}$$

Since C^α and C^β anticommute, we can antisymmetrize the coefficient of $C^\alpha C^\beta$. Then using the commutation relations (15.43) we find

$$\delta\left[D^i_\alpha(\varphi)C^\alpha\right] = \tfrac{1}{2}\bar\varepsilon f^\gamma_{\alpha\beta}C^\alpha C^\beta D^i_\gamma(\varphi). \tag{15.48}$$

It is possible to cancel this variation by performing a simultaneous transformation on the sources C^α of the form

$$\delta C^\alpha = -\tfrac{1}{2}\bar\varepsilon f^\alpha_{\beta\gamma}C^\beta C^\gamma . \tag{15.49}$$

One verifies that the combined transformation (15.45,15.49) is nilpotent of vanishing square. Its geometric origin will be explained in Section 16.4 when we discuss *BRS symmetry*.

We now calculate the variation of $\left(\delta A/\delta\varphi^i\right)D^i_\alpha(\varphi)C^\alpha$ under the combined transformation (15.45,15.49):

$$\delta\left[\frac{\delta A}{\delta\varphi^i}D^i_\alpha(\varphi)C^\alpha\right] = \bar\varepsilon\frac{\delta^2 A}{\delta\varphi^i\delta\varphi^j}D^j_\beta D^i_\alpha C^\beta C^\alpha + \frac{\delta A}{\delta\varphi_i}\delta\left[D^i_\alpha(\varphi)C_\alpha\right] = 0 . \tag{15.50}$$

The first term vanishes because the coefficient of $C^\alpha C^\beta$ is symmetric in (α,β), and the second term vanishes by construction. The algebra thus closes because the combined transformation (15.45,15.49) is nilpotent. We obtain an identity for

$$\begin{aligned}\mathcal{Z}(J,K,C,\Lambda,L) = \int [\mathrm{d}\rho(\varphi)]\exp\Bigg[&-\mathcal{S}(\varphi) + \int \mathrm{d}x\,\left(J_i\varphi^i + KA(\varphi)\right)\\ &+ \int \mathrm{d}x\left(\Lambda_i + L\frac{\delta A}{\delta\varphi^i}\right)D^i_\alpha(\varphi)C^\alpha\Bigg],\end{aligned} \tag{15.51}$$

which, according to the previous analysis, has the form

$$\mathcal{Z}(J,K,C+\delta C,\Lambda,L)=\left[1+\bar{\varepsilon}\int \mathrm{d}x\left(J_i(x)\frac{\delta}{\delta\Lambda_i(x)}+K(x)\frac{\delta}{\delta L(x)}\right)\right]\mathcal{Z}(J,K,C,\Lambda,L), \tag{15.52}$$

and, therefore,

$$\left[\int \mathrm{d}x\left(J_i\frac{\delta}{\delta\Lambda_i}+K\frac{\delta}{\delta L}\right)+\frac{1}{2}f^{\alpha}_{\beta\gamma}C^{\beta}C^{\gamma}\frac{\partial}{\partial C^{\alpha}}\right]\mathcal{Z}=0\,. \tag{15.53}$$

A similar identity for $\mathcal{W}=\ln\mathcal{Z}$ follows. After Legendre transformation with respect to J_i we then obtain the WT identities for Γ:

$$\int \mathrm{d}x\left(\frac{\delta\Gamma}{\delta\varphi^i}\frac{\delta\Gamma}{\delta\Lambda_i}+K\frac{\delta\Gamma}{\delta L}\right)+\frac{1}{2}f^{\alpha}_{\beta\gamma}C^{\beta}C^{\gamma}\frac{\partial\Gamma}{\partial C^{\alpha}}=0\,. \tag{15.54}$$

15.3.3 The renormalized action

Arguments, which by now should be familiar to the reader, allow one to show that equation (15.54) is stable under renormalization and is also satisfied by the effective renormalized action $\mathcal{S}_{\mathrm{r}}$:

$$\int \mathrm{d}x\left(\frac{\delta\mathcal{S}_{\mathrm{r}}}{\delta\varphi^i}\frac{\delta\mathcal{S}_{\mathrm{r}}}{\delta\Lambda_i}+K\frac{\delta\mathcal{S}_{\mathrm{r}}}{\delta L}\right)+\frac{1}{2}f^{\alpha}_{\beta\gamma}C^{\beta}C^{\gamma}\frac{\partial\mathcal{S}_{\mathrm{r}}}{\partial C^{\alpha}}=0\,. \tag{15.55}$$

To solve the equation, we note that, as a consequence of fermion number conservation, Λ, L and C can appear only in the combinations (ΛC) and (LC). In two dimensions, the sources K, (ΛC) and (LC) have dimension 2, and, therefore, $\mathcal{S}_{\mathrm{r}}$ is a linear function of these sources with derivative-free coefficients:

$$\mathcal{S}_{\mathrm{r}}(\varphi,K,C,\Lambda,L)=\mathcal{S}_{\mathrm{r}}(\varphi)-\int \mathrm{d}x\left[KA_{\mathrm{r}}(\varphi)+\Lambda_i D^i_{\mathrm{r}\alpha}(\varphi)C^{\alpha}+LM_{\alpha}(\varphi)C^{\alpha}\right]. \tag{15.56}$$

We first write the equation obtained by identifying the coefficient of $\Lambda_i(x)$ in equation (15.55):

$$\frac{\partial D^i_{\mathrm{r}\alpha}}{\partial\varphi^j}D^j_{\mathrm{r}\beta}C^{\alpha}C^{\beta}-\frac{1}{2}f^{\alpha}_{\beta\gamma}C^{\beta}C^{\gamma}D^i_{\mathrm{r}\alpha}=0\,. \tag{15.57}$$

The antisymmetric coefficient of $C^{\alpha}C^{\beta}$ must vanish:

$$\frac{\partial D^i_{\mathrm{r}\alpha}}{\partial\varphi^j}D^j_{\mathrm{r}\beta}-\frac{\partial D^i_{\mathrm{r}\beta}}{\partial\varphi^j}D^j_{\mathrm{r}\alpha}=f^{\gamma}_{\alpha\beta}D^i_{\mathrm{r}\gamma}\,. \tag{15.58}$$

The functions $D^i_{\mathrm{r}\alpha}(\varphi)$ are thus associated with a non-linear representation of the group G.

We now identify the coefficient of K:

$$\frac{\delta A_{\mathrm{r}}}{\delta\varphi^i}D^i_{\mathrm{r}\alpha}(\varphi)C^{\alpha}-M_{\alpha}C^{\alpha}=0\quad\Rightarrow\quad M_{\alpha}(\varphi)=\frac{\delta A_{\mathrm{r}}}{\delta\varphi^i}D^i_{\mathrm{r}\alpha}(\varphi), \tag{15.59}$$

and the coefficient of L:

$$\frac{\delta M_\alpha}{\delta\varphi^i} D^i_{\mathrm{r}\beta} - \frac{\delta M_\beta}{\delta\varphi^i} D^i_{\mathrm{r}\alpha} = f^\gamma_{\alpha\beta} M_\gamma \,. \tag{15.60}$$

The latter equation is already implied by the two equations (15.58,15.59). We conclude that $A_\mathrm{r}(\varphi)$ is in general an arbitrary function of φ.

Finally, the last equation, independent of the different sources, is

$$\frac{\delta \mathcal{S}_\mathrm{r}}{\delta\varphi^i}(\varphi) D^i_{\mathrm{r}\alpha}(\varphi) = 0 \,. \tag{15.61}$$

The renormalized action is invariant under the non-linear transformations of the group G generated by $D^i_{\mathrm{r}\alpha}(\varphi)$.

We do not need to discuss again thoroughly the solutions of the WT identities which we have reduced to equations (15.58,15.61). The latter equation implies a renormalized form of equation (15.40), 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 $D^{\mathrm{r}i}_\alpha(\varphi)$ and implies that $\mathcal{S}_\mathrm{r}(\varphi)$ 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. The method based on adding sources for composite operators should be used when applicable.

15.4 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 Appendix A15.4.2 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 Appendix A15.4.

Field theory models in two dimensions, constructed on symmetric spaces, have special properties both on the classical level and after quantization, which we examine below. The non-linear σ-model 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 on 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 g of a group $\mathfrak{G}$ of the form (see Appendix A15.4 for details)

$$g = g_2^{-1} g_0 g_1 \,,$$

where g_0, g_1, g_2 are elements of $\mathfrak{G}$ which satisfy some conditions. We can then distinguish two cases:

(i) $g_0 \equiv 1$ and g_1, g_2 are two independent elements of $\mathfrak{G}$. The automorphism is $g_1 \leftrightarrow g_2$. 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 theory models are called *principal chiral models.* They are related to the chiral models studied in Section 13.6.

(ii) The element g_0 is a fixed element different from the identity and satisfies

$$g_0^* g_0 = \epsilon \mathbf{1} \,, \qquad \epsilon = \pm 1 \,,$$

in which the star operation is an involutive automorphism of the group $\mathfrak{G}$ which may be trivial (typically for unitary groups it can also be the complex conjugation). The elements g_1 and g_2 are related by

$$g_2 = \left(g_1^{-1}\right)^* ,$$

which implies that the elements g have the form

$$g = \left(f^{-1}\right)^* g_0 f \,, \tag{15.62}$$

where the elements f vary freely in the group $\mathfrak{G}$. As a consequence,

$$g^* g = \epsilon \mathbf{1} \,. \tag{15.63}$$

Then $G \equiv \mathfrak{G}$, the automorphism of G is $g \mapsto g_0^{-1} g^* g_0$ and H is thus the subgroup of elements which satisfy $h = g_0^{-1} h^* g_0$.

Note that below we extend by continuity the * operation to the Lie algebra.

15.5 The Classical Action. Conservation Laws

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:

$$\mathcal{S}(\mathbf{g}) = \frac{1}{2} \int \mathrm{d}x \; \mathrm{tr} \left[\partial_\mu \mathbf{g}(x) \partial_\mu \mathbf{g}^{-1}(x)\right] , \tag{15.64}$$

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 rewritten in terms of the associated current $\mathbf{A}_\mu$ which belongs to the Lie algebra $\mathcal{L}(\mathfrak{G})$:

$$\mathbf{A}_\mu = \mathbf{g}^{-1}(x) \partial_\mu \mathbf{g}(x). \tag{15.65}$$

The field $\mathbf{A}_\mu$ can be considered as a connection or gauge field (these concepts will be discussed in Chapters 18,19,22), of a special kind, called a pure gauge. It is characterized by the vanishing of the corresponding curvature tensor $\mathbf{F}_{\mu\nu}$:

$$\mathbf{F}_{\mu\nu} \equiv \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu] = 0 \,, \tag{15.66}$$

as a short calculation shows.

The action then takes the form

$$\mathcal{S}(\mathbf{g}) = -\frac{1}{2} \int \mathrm{d}x \; \mathrm{tr}\, \mathbf{A}_\mu^2 \,. \tag{15.67}$$

To the action (15.64) corresponds a classical field equation. In all cases, a general variation of $\mathbf{g}(x)$ can be written as

$$\mathbf{g}(x) \mapsto (1 - \varepsilon^*(x))\, \mathbf{g}(x)\, (1 + \varepsilon(x))\,, \tag{15.68}$$

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 star automorphism in order to preserve the condition (15.63):

$$\mathbf{g}^*\mathbf{g} = s\mathbf{1}\,. \tag{15.69}$$

The corresponding variation of $\mathbf{A}_\mu$ is

$$\delta\mathbf{A}_\mu = \mathbf{D}_\mu \left(\varepsilon - \mathbf{g}^{-1}\varepsilon^*\mathbf{g}\right),$$

where we have introduced the *covariant derivative* associated with the current $\mathbf{A}_\mu$, which acts on an element ω of the Lie algebra as

$$\mathbf{D}_\mu\omega = \partial_\mu\omega + [\mathbf{A}_\mu, \omega]. \tag{15.70}$$

The variation of the action then takes the form

$$\delta S = \operatorname{tr}\left[\partial_\mu\mathbf{A}_\mu \left(\varepsilon - \mathbf{g}^{-1}\varepsilon^*\mathbf{g}\right)\right].$$

In the case of chiral models we can restrict ourselves for example to $\varepsilon^* = 0$. In the other cases, the relation (15.69) implies

$$\partial_\mu\mathbf{A}_\mu^* = -\mathbf{g}\partial_\mu\mathbf{A}_\mu\mathbf{g}^{-1}.$$

In both cases, we thus find the classical field equation

$$\partial_\mu\mathbf{A}_\mu = 0\,, \tag{15.71}$$

which expresses the conservation of the current related to the $\mathfrak{G}$-symmetry .

Non-local conserved currents. The equations (15.65,15.71) imply, in two dimensions, the existence of an infinite number of non-local conserved currents.

We now define a covariant derivative $\mathbf{D}_\mu$ by

$$\mathbf{D}_\mu = \partial_\mu\mathbf{1} + \mathbf{A}_\mu\,. \tag{15.72}$$

We note that the explicit form of the covariant derivative is different from the expression (15.70)—indeed, as also discussed in Section 22.2, this form depends on the representation.

We consider the linear partial differential equations for matrices χ:

$$\mathbf{D}_\mu\chi(x) = \kappa\varepsilon_{\mu\nu}\partial_\nu\chi(x), \tag{15.73}$$

in which χ is a function of x and the spectral parameter κ.

Let us show that the linear system (15.73) is compatible. Setting

$$\Delta_\mu = \mathbf{D}_\mu - \kappa\varepsilon_{\mu\rho}\partial_\rho\,, \tag{15.74}$$

we have to calculate the commutator

$$[\Delta_\mu, \Delta_\nu] = [\mathbf{D}_\mu, \mathbf{D}_\nu] - \kappa \left(\varepsilon_{\mu\rho} \left[\partial_\rho, \mathbf{D}_\nu\right] + \varepsilon_{\nu\sigma} \left[\mathbf{D}_\mu, \partial_\sigma\right]\right). \tag{15.75}$$

In two dimensions, since $\mu \neq \nu$, in the last term of the r.h.s. only $\rho = \nu$ and $\sigma = \mu$ give a non-vanishing contribution:

$$[\Delta_\mu, \Delta_\nu] = [\mathbf{D}_\mu, \mathbf{D}_\nu] - \kappa\varepsilon_{\mu\nu} \left[\partial_\rho, \mathbf{D}_\rho\right]. \tag{15.76}$$

The field equation (15.71) implies

$$[\partial_\mu, \mathbf{D}_\mu] = \partial_\mu \mathbf{A}_\mu = 0\,. \tag{15.77}$$

The commutator $[\mathbf{D}_\mu, \mathbf{D}_\nu]$ is the curvature $\mathbf{F}_{\mu\nu}$ of equation (15.66) which vanishes:

$$[\mathbf{D}_\mu, \mathbf{D}_\nu] = \mathbf{F}_{\mu\nu} = 0\,. \tag{15.78}$$

The conclusion is

$$[\Delta_\mu, \Delta_\nu] = 0\,, \tag{15.79}$$

and the linear system (15.73) is compatible.

We now define the current $\mathbf{J}_\mu(x, \kappa)$:

$$\mathbf{J}_\mu(x, \kappa) = \mathbf{D}_\mu \chi(x, \kappa). \tag{15.80}$$

As a consequence of equation (15.73), the current is conserved:

$$\partial_\mu \mathbf{J}_\mu = 0\,. \tag{15.81}$$

The solution of equation (15.73) has an expansion in powers of κ of the form

$$\chi(x, \kappa) = \mathbf{1} + \sum_{n=1}^{\infty} \chi_n(x) \kappa^{-n}. \tag{15.82}$$

The corresponding expansion of $\mathbf{J}_\mu$ then generates an infinite number of conserved currents:

$$\mathbf{J}_\mu(x, \kappa) = \mathbf{A}_\mu(x) + \sum_{n=1}^{\infty} \mathbf{J}_\mu^n(x) \kappa^{-n}. \tag{15.83}$$

The interesting question, which we will 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 more details.

15.6 Quantum Theory: Perturbation Theory and RG Functions

In the quantized theory we consider the integral representation of the generating functional $\mathcal{Z}(\mathbf{j})$:

$$\mathcal{Z}(\mathbf{j}) = \int [\mathrm{d}\mathbf{f}(x)] \exp\left\{-\frac{1}{\lambda}\left[\mathcal{S}(\mathbf{g}) - \int \mathrm{d}x \, \mathrm{tr}\, \mathbf{g}(x)\mathbf{j}(x)\right]\right\}, \tag{15.84}$$

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 (15.62):

$$\mathbf{g} = \left(\mathbf{f}^{-1}\right)^* \mathbf{g}_0 \mathbf{f} ,$$

and $\mathrm{d}\mathbf{f}$ is the de Haar measure for the group $\mathfrak{G}$. The parameter λ is the coupling constant.

We can then parametrize $\mathbf{f}(x)$ in a form analogous to (15.4), in terms of independent field variables $\xi_a(x)$, coefficients of the generators belonging to $\mathcal{L}(G/H)$. We can then expand perturbation theory around a finite value of the source $\mathbf{j}(x)$ to provide perturbation theory with an IR cut-off.

Actually, for practical calculations it is convenient to return to the field representation (15.4,15.6):

$$|\phi(x)\rangle = \mathbf{R}(\xi)\,|0\rangle = \exp\left[\sum_a t^a \xi_a(x)\right]|0\rangle , \tag{15.85}$$

valid for general homogeneous spaces. With these notations the generating functional $\mathcal{Z}(J)$ reads

$$\mathcal{Z}(J) = \int [\mathrm{d}\rho(\xi)] \exp\left\{-\frac{1}{\lambda}\int \mathrm{d}^d x \left[-\tfrac{1}{2}\langle 0|\left(\mathbf{R}^{-1}\partial_\mu \mathbf{R}\right)^2|0\rangle - \langle J(x)|\,\mathbf{R}\,|0\rangle\right]\right\}. \tag{15.86}$$

To calculate in perturbation theory, we expand the action in powers of $\xi_a(x)$. The geometric part of the Feynman diagram calculation then involves the evaluation of averages of the form $\langle 0|\, t^{a_1} t^{a_2} \ldots t^{a_k}\,|0\rangle$.

15.6.1 RG functions at one-loop order

We will eventually discuss the properties of these models from the point of view of the renormalization group. Therefore, we calculate here the renormalization constants and the renormalization group functions at one-loop order. Moreover, these calculations will illustrate some of the preceding considerations.

Preliminary remarks. We normalize the vector $|0\rangle$:

$$\langle 0\,|0\rangle = 1 . \tag{15.87}$$

To evaluate the one-loop diagrams, we also need a few tensors:

$$\langle 0|\, t^a t^b\,|0\rangle = -\delta_{ab} . \tag{15.88}$$

This relation is a consequence (up to the normalization) of the property that in the case of symmetric spaces the vectors $t^a|0\rangle$ form an irreducible representation of the subgroup H.

In the same way, because the structure constants are antisymmetric and the vector $|0\rangle$ is the unique vector having H for stabilizer group:

$$t^a t^a |0\rangle = -\mu |0\rangle . \tag{15.89}$$

In equation (15.89) and in the equations that follow, repeated indices means summation over indices running from 1 to l, values which correspond to the generators of G/H.

Combining equations (15.87–15.89) we find the value of μ:

$$t^a t^a |0\rangle = -l |0\rangle . \tag{15.90}$$

We need two tensors with three indices. Since the generators are represented by anti-symmetric matrices,

$$\langle 0| t^a t^b t^c |0\rangle + \langle 0| t^c t^b t^a |0\rangle = 0 .$$

A few commutations and the properties of symmetric spaces lead to:

$$\langle 0| t^a t^b t^c |0\rangle = 0 . \tag{15.91}$$

A different useful quantity is obtained by replacing t^b by a generator τ^β of H:

$$\langle 0| t^a \tau^\beta t^c |0\rangle = -f_{a\beta c} . \tag{15.92}$$

We still have to evaluate three tensors with four indices. Combining commutation relations and previous relations we find

$$\langle 0| t^a t^a t^b t^b |0\rangle = l^2 , \tag{15.93}$$
$$\langle 0| t^a t^b t^a t^b |0\rangle = l^2 - \sum_{\gamma > l} f_{\gamma ab} f_{\gamma ab} , \tag{15.94}$$
$$\langle 0| t^a t^b t^b t^a |0\rangle = \langle 0| t^a t^b t^a t^b |0\rangle .$$

Using the antisymmetry of the structure constants and the special properties of symmetric spaces it is easy to verify that

$$\sum_{\alpha > l} \sum_{a=1}^{l} f_{\alpha ab} f_{\alpha ab'} = \tfrac{1}{2} \delta_{bb'} C , \tag{15.95}$$

in which C is the Casimir of the group G. Summing over $b = b'$ we find an expression for the second term in the r.h.s. of equation (15.94). Therefore,

$$\langle 0| t^a t^b t^a t^b |0\rangle = l\,(l - C/2) . \tag{15.96}$$

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 $\mathcal{S}_\mathrm{r}$:

$$\mathcal{S}_\mathrm{r} = \frac{1}{\lambda Z_\lambda} \int \mathrm{d}^d x \left[\tfrac{1}{2} \langle \partial_\mu \phi | \partial_\mu \phi \rangle - Z_\phi^{-1/2} Z_\lambda \langle J(x) | \phi(x) \rangle \right] , \tag{15.97}$$

for the special constant source

$$|J(x)\rangle = m^2 \,|0\rangle\,. \tag{15.98}$$

Dimensional regularization will be used so that the measure term can be omitted. The action has to be expanded up to order $\xi^4(x)$. Using

$$\mathrm{e}^{-\mathbf{t}\cdot\boldsymbol{\xi}}\,\partial_\mu\,\mathrm{e}^{\mathbf{t}\cdot\boldsymbol{\xi}} = \mathbf{t}\cdot\partial_\mu\boldsymbol{\xi} + \frac{1}{2}\left[\mathbf{t}\cdot\partial_\mu\boldsymbol{\xi}, \mathbf{t}\cdot\boldsymbol{\xi}\right] + \frac{1}{6}\left[\left[\partial_\mu\boldsymbol{\xi}\cdot\mathbf{t}, \boldsymbol{\xi}\cdot\mathbf{t}\right], \boldsymbol{\xi}\cdot\mathbf{t}\right] + O\left(\xi^4\right), \tag{15.99}$$

which for symmetric spaces leads to

$$\langle\partial_\mu\phi\,|\partial_\mu\phi\rangle = \left(\partial_\mu\boldsymbol{\xi}\right)^2 + \frac{1}{3}\sum_\gamma\left(f_{\gamma ab}\xi_a\partial_\mu\xi_b\right)^2 + O\left(\xi^5\right), \tag{15.100}$$

we can write the renormalized action up to order ξ^4:

$$\begin{aligned}\mathcal{S}_{\mathrm{r}}(\xi) = \frac{1}{\lambda}\int \mathrm{d}^d x\Biggl\{&\frac{1}{Z_\lambda}\left[\frac{1}{2}\left(\partial_\mu\boldsymbol{\xi}\right)^2 + \frac{1}{6}\sum_\gamma\left(f_{\gamma ab}\xi_a\partial_\mu\xi_b\right)^2\right] \\ &- Z_\phi^{-1/2} m^2\left(1 - \frac{1}{2}\boldsymbol{\xi}^2 + \frac{1}{24}\langle 0|\,t^a t^b t^c t^d\,|0\rangle\,\xi_a\xi_b\xi_c\xi_d\right)\Biggr\} + O\left(\xi^5\right).\end{aligned} \tag{15.101}$$

A short calculation yields $\mathcal{W}(m^2)$ up to order λ:

$$\begin{aligned}\mathcal{W}(m^2) = Z_\phi^{-1/2}\frac{m^2}{\lambda} + \frac{l}{2}\int \mathrm{d}^d q\,\ln\left(1 + \frac{m^2}{q^2}Z_\phi^{-1/2}Z_\lambda\right) - \frac{\lambda}{8}m^2 l\,(l - C)\left(\int\frac{\mathrm{d}^d q}{q^2 + m^2}\right)^2 \\ + O\left(\lambda^2\right).\end{aligned} \tag{15.102}$$

Setting $d = 2+\varepsilon$, we make a Laurent expansion for ε small and define the renormalization constants by minimal subtraction:

$$Z_\lambda = 1 + \frac{C}{4\pi\varepsilon}\lambda + O\left(\lambda^2\right), \tag{15.103}$$

$$Z_\phi = 1 + \frac{l}{2\pi\varepsilon}\lambda + \frac{l\,(2l + C)}{(4\pi\varepsilon)^2}\lambda^2 + O\left(\lambda^3\right). \tag{15.104}$$

The coupling constant RG function $\beta(\lambda)$ and the field RG function $\eta(\lambda)$ are, then,

$$\beta(\lambda) = \varepsilon\lambda\left(1 + \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln Z_\lambda\right)^{-1} = \varepsilon\lambda - \frac{C}{4\pi}\lambda^2 + O\left(\lambda^3\right), \tag{15.105}$$

$$\eta(\lambda) = \beta(\lambda)\frac{\mathrm{d}\ln Z_\phi}{\mathrm{d}\lambda} = \frac{l}{2\pi}\lambda + O\left(\lambda^4\right). \tag{15.106}$$

In contrast with field theories like $\lambda\phi_4^4$, 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 (asymptotic freedom) of this remarkable property will be discussed in Chapter 31.

15.6.2 One-loop β-function and background field method

If one is interested only in the coefficients of the perturbative expansion of the β-function, one can shorten the calculation by using the background field method whose principles have been explained in Appendix A7.2. As an exercise and to show some advantages of this 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 more details see Chapter 37) in $(d-1)$ dimensions the coordinates x_μ vary in the interval

$$0 \le x_\mu \le L_\perp\,, \qquad 1 \le \mu \le d-1\,,$$

and we impose periodic boundary conditions on the field:

$$\phi(z, x_1, \ldots, 0, \ldots, x_{d-1}) = \phi(z, x_1, \ldots, L_\perp, \ldots, x_{d-1}).$$

We have called z the last coordinate, the imaginary time coordinate, which varies in the interval

$$0 \le z \le L\,,$$

and for which we impose fixed "twisted" boundary conditions:

$$|\phi_\theta(z=0, \mathbf{x})\rangle = |0\rangle\,, \qquad |\phi_\theta(z=L, \mathbf{x})\rangle = \mathrm{e}^{\boldsymbol{\theta}}\,|0\rangle\,, \tag{15.107}$$

in which $\boldsymbol{\theta}$ is a linear combination of generators belonging to $\mathcal{L}(G/H)$:

$$\boldsymbol{\theta} = \sum t^a \theta_a\,, \qquad t^a \in \mathcal{L}(G/H). \tag{15.108}$$

As a consequence, momenta in Fourier space are quantized:

$$p_\mu = \frac{2\pi}{L_\perp} n_\mu\,, \quad n_\mu \in \mathbb{Z}^{d-1}, \qquad p_z = \frac{\pi}{L} m\,, \quad m \in \mathbb{Z}\,.$$

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

$$|\phi_\theta(z, \mathbf{x})\rangle = \mathrm{e}^{z\boldsymbol{\theta}/L}\,|\phi(z, \mathbf{x})\rangle\,. \tag{15.109}$$

The new field then satisfies

$$|\phi(0, \mathbf{x})\rangle = |\phi(L, \mathbf{x})\rangle = |0\rangle\,, \tag{15.110}$$

which in the parametrization

$$|\phi(z, \mathbf{x})\rangle = \exp\left[t^a \xi_a(z, \mathbf{x})\right]|0\rangle\,, \tag{15.111}$$

is equivalent to

$$\xi_a(0, \mathbf{x}) = \xi_a(L, \mathbf{x}) = 0\,. \tag{15.112}$$

The renormalized action then reads ($\partial_z \equiv \partial/\partial z$):

$$\mathcal{S}_{\mathrm{r}}(\xi) = \frac{1}{2\lambda Z_\lambda} \int \mathrm{d}z\, \mathrm{d}^{d-1}x \left[\langle \partial_\mu \phi |\, \partial_\mu \phi \rangle + \frac{2}{L} \langle \partial_z \phi |\, \boldsymbol{\theta}\, | \phi \rangle - \frac{1}{L^2} \langle \phi\, | \boldsymbol{\theta}^2 |\, \phi \rangle \right]. \tag{15.113}$$

The calculation of the one-loop contribution involves only the expansion of $\mathcal{S}_{\mathrm{r}}(\xi)$ up to order ξ^2. Using the relations (15.88) and (15.91),

$$\langle 0| t^a t^b |0\rangle = -\delta_{ab}\,, \qquad \langle 0| t^a t^b t^c |0\rangle = 0\,,$$

we find

$$\langle \partial_\mu \phi | \partial_\mu \phi \rangle = (\partial_\mu \boldsymbol{\xi})^2 + O\left(\xi^3\right), \tag{15.114}$$

$$\int \mathrm{d}z\, \langle \partial_z \phi | \boldsymbol{\theta} | \phi \rangle = \int \partial_z \xi_a \theta_a \mathrm{d}z + O\left(\xi^3\right) = O\left(\xi^3\right), \tag{15.115}$$

$$\langle \phi | \boldsymbol{\theta}^2 | \phi \rangle = -\theta_a \theta_a + \tfrac{1}{2} \xi_a \xi_b \left\langle 0 \left| \left[t^a, \left[t^b, \boldsymbol{\theta}^2\right]\right] \right| 0 \right\rangle + O\left(\xi^3\right). \tag{15.116}$$

To evaluate the last term we need

$$V_{abcd} = \tfrac{1}{2} \left\langle 0 \left| \left[t^a, \left[t^b, t^c t^d\right]\right] \right| 0 \right\rangle = \tfrac{1}{2} \sum_{\varepsilon} \left(f_{\varepsilon ad} f_{\varepsilon bc} + f_{\varepsilon ac} f_{\varepsilon bd}\right). \tag{15.117}$$

The action in the gaussian approximation is

$$\mathcal{S}_{\mathrm{r}}(\xi) = \frac{L_\perp^{d-1} \theta_a \theta_a}{2L\lambda Z_\lambda} + \int \mathrm{d}^{d-1}x\, \mathrm{d}z \left[\frac{1}{2} \left(\partial_\mu \boldsymbol{\xi}\right)^2 - \frac{1}{2L^2} \xi_a \xi_b \theta_c \theta_d V_{abcd}\right] + O\left(\xi^3\right). \tag{15.118}$$

The free energy as a function of θ_a is, at one-loop order,

$$\mathcal{W}(\theta) = -\frac{L_\perp^{d-1} \theta_a \theta_a}{2L\lambda Z_\lambda} - \frac{1}{2} \operatorname{tr} \ln \left\{ -\left[\partial_z^2 + \nabla_\perp^2\right] \delta_{ab} - \frac{1}{L^2} V_{abcd} \theta_c \theta_d \right\}, \tag{15.119}$$

where $\nabla_\perp^2$ is the laplacian in $d-1$ dimensions. To compute the renormalization constant Z_λ it is sufficient to expand up to second order in θ:

$$\mathcal{W}(\theta) = -\frac{L_\perp^{d-1} \theta_a \theta_a}{2L\lambda Z_\lambda} + \frac{1}{2L^2} V_{aacd} \theta_c \theta_d \sum_{\substack{p_z = m\pi/L \\ \mathbf{p} = 2\pi \mathbf{n}/L_\perp}} \frac{1}{p_z^2 + \mathbf{p}^2}. \tag{15.120}$$

We have shown that in a symmetric space (equation (15.95))

$$V_{aacd} = \sum_{\varepsilon} f_{\varepsilon ad} f_{\varepsilon ac} = \tfrac{1}{2} C \delta_{cd} > 0\,,$$

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:

$$\sum_{p_z, \mathbf{p}} \frac{1}{p_z^2 + \mathbf{p}^2} \sim \frac{LL_\perp}{(2\pi)^2} \int_{|p|>1} \frac{\mathrm{d}^d p}{\mathbf{p}^2} \sim -\frac{LL_\perp}{2\pi\varepsilon}. \tag{15.121}$$

It follows that Z_λ in the minimal subtraction scheme is

$$Z_\lambda = 1 + \frac{C}{4\pi\varepsilon} \lambda + O\left(\lambda^2\right), \tag{15.122}$$

in agreement with the result (15.103) of Section 15.6.1.

15.7 Generalizations

Non-compact homogeneous spaces. Up to now 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. Obvious conditions 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 compact and even the maximal compact subgroup of G.

Some care has to be taken to define properly the functional integral since the volume of the group manifold is in general infinite so that appropriate boundary conditions have to be imposed.

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 t^a in G/H by it^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 metric with M $+$ signs and N $-$ signs invariant.

The perturbative expansion of the correlation functions of these models is obtained, up to global signs, by changing the sign of the coupling constant in the compact models. This changes, in particular, the sign of the one loop β-function in two dimensions, and thus has important consequences from the renormalization group point of view.

Arbitrary Riemannian manifolds. A much wider class of models has been studied by Friedan. One considers an arbitrary smooth (infinitely differentiable) Riemannian manifold with a smooth positive definite metric $g_{ij}(\varphi)$ and one takes the classical action $\mathcal{S}(\varphi)$:

$$\mathcal{S}(\varphi)=\frac{1}{2}\int \mathrm{d}^d x\, \partial_\mu\varphi^i(x) g_{ij}\left(\varphi(x)\right)\partial_\mu\varphi^j(x). \tag{15.123}$$

The generating functional of correlation functions $\mathcal{Z}(J)$ is then given by

$$\mathcal{Z}(J)=\int [\mathrm{d}\rho(\varphi)]\exp\left[-\mathcal{S}(\varphi)+\int \mathrm{d}x\, J_i(x)\varphi^i(x)\right], \tag{15.124}$$

in which $\mathrm{d}\rho(\varphi)$ has to be a smooth, strictly positive, covariant measure on the manifold, for example (see Chapter 22),

$$\mathrm{d}\rho(\varphi)=\sqrt{g}\,\mathrm{d}\varphi\,,$$

in which g is the determinant of the metric tensor.

As we have discussed in Sections 3.2 and 4.8, in one dimension, the functional integral (15.124) is associated with a hamiltonian of the form of a Laplace operator on the manifold with metric g_{ij} and is related to brownian motion or diffusion processes on the manifold.

In two dimensions, the action (15.123) corresponds to theories renormalizable by power counting, which generalizes theories on homogeneous spaces. However, since the symmetry properties are lost, and, in particular, the Goldstone theorem which forces fields to remain massless, important differences appear after renormalization:

(i) in general the space of all possible metrics is infinite dimensional;

15.7 Generalizations

Non-compact homogeneous spaces. Up to now 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. Obvious conditions 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 compact and even the maximal compact subgroup of G.

Some care has to be taken to define properly the functional integral since the volume of the group manifold is in general infinite so that appropriate boundary conditions have to be imposed.

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 t^a in G/H by it^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 metric with M $+$ signs and N $-$ signs invariant.

The perturbative expansion of the correlation functions of these models is obtained, up to global signs, by changing the sign of the coupling constant in the compact models. This changes, in particular, the sign of the one loop β-function in two dimensions, and thus has important consequences from the renormalization group point of view.

Arbitrary Riemannian manifolds. A much wider class of models has been studied by Friedan. One considers an arbitrary smooth (infinitely differentiable) Riemannian manifold with a smooth positive definite metric $g_{ij}(\varphi)$ and one takes the classical action $\mathcal{S}(\varphi)$:

$$\mathcal{S}(\varphi)=\frac{1}{2}\int \mathrm{d}^d x\, \partial_\mu\varphi^i(x) g_{ij}\left(\varphi(x)\right)\partial_\mu\varphi^j(x). \tag{15.123}$$

The generating functional of correlation functions $\mathcal{Z}(J)$ is then given by

$$\mathcal{Z}(J)=\int [\mathrm{d}\rho(\varphi)]\exp\left[-\mathcal{S}(\varphi)+\int \mathrm{d}x\, J_i(x)\varphi^i(x)\right], \tag{15.124}$$

in which $\mathrm{d}\rho(\varphi)$ has to be a smooth, strictly positive, covariant measure on the manifold, for example (see Chapter 22),

$$\mathrm{d}\rho(\varphi)=\sqrt{g}\,\mathrm{d}\varphi\,,$$

in which g is the determinant of the metric tensor.

As we have discussed in Sections 3.2 and 4.8, in one dimension, the functional integral (15.124) is associated with a hamiltonian of the form of a Laplace operator on the manifold with metric g_{ij} and is related to brownian motion or diffusion processes on the manifold.

In two dimensions, the action (15.123) corresponds to theories renormalizable by power counting, which generalizes theories on homogeneous spaces. However, since the symmetry properties are lost, and, in particular, the Goldstone theorem which forces fields to remain massless, important differences appear after renormalization:

(i) in general the space of all possible metrics is infinite dimensional;

(ii) with a given metric can be associated an infinite number of different covariant measure terms since one can construct an infinite number of scalars like the scalar curvature (see Chapter 22), and a covariant 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 (15.123) of the action by adjusting an infinite number of parameters.

In addition, the renormalized metric will be generically the most general metric on the manifold. In other words, the renormalized action is the most general action allowed by power counting arguments.

Considerations based on covariance (Chapter 22), 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 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. Friedan has in this way obtained the first two terms. We have expressed them in terms of g^{ij} the inverse of the metric tensor g_{ij}, because it naturally orders perturbation theory. In $2+\varepsilon$ dimensions,

$$\beta^{ij}(g) = \varepsilon g^{ij} - \frac{1}{2\pi} R^{ij} - \frac{1}{8\pi^2} R^{iklm} R^{j}_{klm} + O\left(\left(g^{ij}\right)^4\right), \tag{15.125}$$

in which R^{ij} is obtained from the Ricci tensor by raising the indices with g^{ij} and R^{j}_{klm} is the curvature tensor.

Bibliographical Notes

The general group theoretical problem of non-linear realizations is discussed in

S. Coleman, J. Wess and B. Zumino, *Phys. Rev.* 177 (1969) 2239.

See also

S. Gasiorowicz and D. Geffin, *Rev. Mod. Phys.* 41 (1969) 531.

Quantization of chiral models is described in

L.D. Faddeev and A.A. Slavnov, *Theor. Math. Phys.* 8 (1971) 843.

Classical conservation laws and RG functions in symmetric spaces have been discussed in

V.E. Zakharov and A.V. Mikhailov, *Zh. Eksp. Teor. Fiz.* 74 (1978) 1953, *Commun. Math. Phys.* 74 (1980) 21; V.L. Golo and A.M. Perelomov, *Phys. Lett.* 79B (1978) 112; H. Eichenherr, *Nucl. Phys.* B146 (1978) 215; A. D'Adda, P. di Vecchia and M. Lüscher, *Nucl. Phys.* B146 (1978) 63; E. Witten, *Nucl. Phys.* B149 (1979) 285; H. Eichenherr and M. Forger, *Nucl. Phys.* B155 (1979) 381; M. Lüscher, *Nucl. Phys.* B135 (1978) 1; E. Brézin, C. Itzykson, J. Zinn-Justin and J.-B. Zuber, *Phys. Lett.* 82B (1979) 442; H.J. de Vega, *Phys. Lett.* 87B (1979) 233.

For a discussion of quantum conservation laws and S-matrix in symmetric spaces see

E. Abdalla, M.C.B. Abdalla and M. Forger, *Nucl. Phys.* B297 (1988) 374 and references therein.

Additional results concerning non-local currents can be found in

D. Bernard, *Commun. Math. Phys.* 137 (1991) 191; D. Bernard and A. LeClair, *Commun. Math. Phys.* 142 (1991) 99.

Different properties are described and RG functions calculated in

R.D. Pisarski, *Phys. Rev.* D20 (1979) 3358; E. Brézin, S. Hikami and J. Zinn-Justin, *Nucl. Phys.* B165 (1980) 528.

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 and F. Wegner, *Z. Phys.* B38 (1980) 113; S. Hikami, *Prog. Theor. Phys. Suppl.* 84 (1985) 120.

Calculation of the RG β-functions up to three and four loops for a class of models can be found in

S. Hikami, *Phys. Lett.* 98B (1981) 208; W. Bernreuther and F.J. Wegner, *Phys. Rev. Lett.* 57 (1986) 1383; F. Wegner, *Nucl. Phys.* B316 (1989) 663.

For a discussion of general two-dimensional models on Riemannian manifolds see

D.H. Friedan, *Ann. Phys. (NY)* 163 (1985) 318.

APPENDIX A15
HOMOGENEOUS SPACES: A FEW ALGEBRAIC PROPERTIES

This appendix first describes some additional properties of homogeneous spaces when considered as Riemannian manifolds. It assumes some minimal familiarity with the elements of differential geometry presented in Chapter 22. The second part is devoted to some elements of classification of symmetric spaces.

A15.1 Pure Gauge. Maurer–Cartan Equations

Multiplying $\mathbf{R}(\varphi)$ by a group element $\mathbf{R}_g$ on the left we see that $\mathbf{R}^{-1}\partial_i\mathbf{R}$ 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 group G. They can, therefore, be expanded on the generators t^α and we set (equation (15.10))

$$\mathbf{R}^{-1}\partial_i\mathbf{R} = L_i^\alpha(\varphi)t^\alpha. \tag{A15.1}$$

It is easy to verify more directly, by parametrizing the l.h.s. as

$$\mathbf{R} = \mathrm{e}^{t^\alpha\xi^\alpha},$$

that the expansion of L_i^α in powers of ξ involves only commutators of generators of $\mathcal{L}(G)$ and, therefore, $L_i^\alpha(\varphi)$ depends on the parametrization of the group elements, but not on the representation to which $\mathbf{R}(\varphi)$ belongs.

The definition $(A15.1)$ implies that the quantities $L_i^\alpha(\varphi)$ are the components of a vector belonging to the adjoint representation and have the φ^i dependence of a pure gauge (see Chapters 19, 20 for details). We have shown in Section 15.5 that the corresponding curvature vanishes (equation (15.66)). In terms of the components L_i^α one finds

$$\begin{aligned}\partial_i\left(\mathbf{R}^{-1}\partial_j\mathbf{R}\right) - \partial_j\left(\mathbf{R}^{-1}\partial_i\mathbf{R}\right) &= t^\alpha\left[\partial_i L_j^\alpha - \partial_j L_i^\alpha\right],\\ \left[\mathbf{R}^{-1}\partial_i\mathbf{R}, \mathbf{R}^{-1}\partial_j\mathbf{R}\right] &= L_i^\alpha L_j^\beta f_{\alpha\beta}^\gamma t^\gamma.\end{aligned}$$

We, thus, find

$$\partial_i L_j^\alpha - \partial_j L_i^\alpha + f_{\beta\gamma}^\alpha L_i^\beta L_j^\gamma = 0\,. \tag{A15.2}$$

These relations, which express that the curvature corresponding to the gauge connection L_i^α vanishes, are known as the Maurer–Cartan equations.

A15.2 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 (15.37):

$$\varphi^i = \varphi'^i + D_\alpha^i(\varphi')\omega^\alpha\,. \tag{A15.3}$$

The generator $\mathcal{D}_\alpha$ of $\mathcal{L}(G)$, as defined by equation (15.41), then characterizes the corresponding infinitesimal variation of scalars:

$$\mathcal{D}_\alpha S(\varphi) = D_\alpha^i(\varphi)\partial_i S(\varphi)\,. \tag{A15.4}$$

More generally, equation (22.9) defines its action on all tensors on the homogeneous space. For vectors $V_i(\varphi)$ it yields

$$\mathcal{D}_\alpha V_i = D^j_\alpha \partial_j V_i + \partial_i D^j_\alpha V_j \,. \tag{A15.5}$$

This transformation law can be verified by a short calculation in the case of the gauge field $L^a_i(\varphi)$, defined by equation (15.10). As explained in Section 22.1 $\mathcal{D}_\alpha V_i$ is a vector, as expected.

For general tensors the result is

$$\mathcal{D}_\alpha V^{i_1\ldots i_p}_{i_{p+1}\ldots i_n} = D^j_\alpha \partial_j V^{i_1\ldots i_p}_{i_{p+1}\ldots i_n} - \sum_{\ell=1}^{p} \partial_j D^{i_\ell}_\alpha V^{i_1\ldots j\ldots i_p}_{i_{p+1}\ldots i_n} + \sum_{\ell=p+1}^{n} \partial_{i_\ell} D^j_\alpha V^{i_1\ldots i_p}_{i_{p+1}\ldots j\ldots i_n} \,. \tag{A15.6}$$

With this definition $\mathcal{D}_\alpha$ obeys the usual rule of differentiation for products of tensors.

The invariance of the metric, as expressed by equation (15.40), then takes the simple form

$$\mathcal{D}_\alpha g_{jk} = 0 \quad \Leftrightarrow \quad \nabla_i D_{j\alpha} + \nabla_j D_{i\alpha} = 0 \,, \tag{A15.7}$$

where the second equation follows from equation (22.110).

Consistency with parallel transport. The metric tensor defines uniquely a torsion-free parallel transport on the manifold. If the infinitesimal change of variables (A15.3) leaves the metric invariant, it leaves invariant all quantities function only of the metric. With the definition (A15.6) we can write

$$\mathcal{D}_\alpha R^k_{lij} = 0 \,, \tag{A15.8}$$

$$\mathcal{D}_\alpha R_{ij} = 0 \,, \tag{A15.9}$$

$$\mathcal{D}_\alpha R = 0 \,. \tag{A15.10}$$

The Christoffel connection is also invariant, but since it is not a tensor the action of $\mathcal{D}_\alpha$ takes the inhomogeneous form (22.33)

$$\mathcal{D}_\alpha \Gamma^i_{jk} = \partial_j D^l_\alpha \Gamma^i_{jk} - \partial_l D^i_\alpha \Gamma^l_{jk} + \partial_k D^l_\alpha \Gamma^i_{jl} + \partial_j D^l_\alpha \Gamma^i_{lk} + \partial_j \partial_k D^i_\alpha = 0 \,.$$

The compatibility between parallel transport and symmetry can then be expressed by the commutation relation

$$[\mathcal{D}_\alpha, \nabla_i] = 0 \,. \tag{A15.11}$$

To prove this 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 (A15.6). For vectors one finds

$$[\mathcal{D}_\alpha, \nabla_i] V_j = V^k \mathcal{D}_\alpha \Gamma^j_{ki} = 0 \,.$$

Equation (A15.9) implies that R_{ij} is an acceptable metric tensor and equation (A15.10) 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 (A15.9) 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.

A final remark: expression (15.12) for the metric shows that the quantities $L_i^a(\varphi)$ play essentially the role of a vielbein in the case of homogeneous spaces, the only difference being the constant internal metric μ_{ab}.

Remark. In Chapter 15 the coordinates φ^i are themselves fields depending on variables x^μ. Then $\partial_\mu\varphi^i(x)$ 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 (15.8). Moreover, in this situation, another covariant derivative D_μ can be defined, which involves the connection (22.20):

$$D_\mu V^i = \partial_\mu V^i + \Gamma^i_{kj}\partial_\mu\varphi^j V^k . \qquad (A15.12)$$

On functions of $\varphi(x)$ only, 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 $\partial_\mu\varphi^i$.

The definition (A15.12) allows us to write the classical field equation corresponding to the action (15.123) in covariant form:

$$D_\mu\partial_\mu\varphi^i(x) = 0 . \qquad (A15.13)$$

A15.3 Explicit Expressions for the Metric

Finally, let us give several more explicit expressions for the metric.

A15.3.1 Metric tensor and transformation law

Let us show that, for a general homogeneous space, we can find an inverse metric tensor g^{ij} of the form

$$g^{ij}(\varphi) = D^i_\alpha(\varphi) m^{\alpha\beta} D^j_\beta(\varphi), \qquad (A15.14)$$

where $m^{\alpha\beta}$ is a constant symmetric non-singular matrix. The tensor has to satisfy the equivalent of equations (15.40) or (A15.7):

$$D^k_\alpha\partial_k g^{ij} = g^{ik}\partial_k D^j_\alpha + g^{jk}\partial_k D^i_\alpha .$$

We replace g^{ij} by the form (A15.14):

$$m^{\beta\gamma}D^k_\alpha\left(\partial_k D^i_\beta D^j_\gamma + \partial_k D^j_\beta D^i_\gamma\right) = m^{\beta\gamma}D^k_\gamma\left(D^i_\beta\partial_k D^j_\alpha + D^j_\beta\partial_k D^i_\alpha\right),$$

and use the Lie algebra commutation relations (15.43). We then obtain

$$m^{\beta\gamma}f^\delta_{\alpha\beta}\left(D^j_\gamma D^i_\delta + D^i_\gamma D^j_\delta\right) = 0 . \qquad (A15.15)$$

Exchanging $\gamma \leftrightarrow \delta$ in one of the terms, we see that the equation is satisfied if $m^{\alpha\beta}$ is a solution to the numerical equation,

$$m^{\gamma\beta}f^\delta_{\beta\alpha} = f^\gamma_{\alpha\beta}m^{\beta\delta} .$$

A final remark: expression (15.12) for the metric shows that the quantities $L_i^a(\varphi)$ play essentially the role of a vielbein in the case of homogeneous spaces, the only difference being the constant internal metric μ_{ab}.

Remark. In Chapter 15 the coordinates φ^i are themselves fields depending on variables x^μ. Then $\partial_\mu\varphi^i(x)$ 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 (15.8). Moreover, in this situation, another covariant derivative D_μ can be defined, which involves the connection (22.20):

$$D_\mu V^i = \partial_\mu V^i + \Gamma^i_{kj}\partial_\mu\varphi^j V^k \,. \tag{A15.12}$$

On functions of $\varphi(x)$ only, 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 $\partial_\mu\varphi^i$.

The definition (*A*15.12) allows us to write the classical field equation corresponding to the action (15.123) in covariant form:

$$D_\mu\partial_\mu\varphi^i(x) = 0\,. \tag{A15.13}$$

A15.3 Explicit Expressions for the Metric

Finally, let us give several more explicit expressions for the metric.

A15.3.1 Metric tensor and transformation law

Let us show that, for a general homogeneous space, we can find an inverse metric tensor g^{ij} of the form

$$g^{ij}(\varphi) = D^i_\alpha(\varphi)m^{\alpha\beta}D^j_\beta(\varphi), \tag{A15.14}$$

where $m^{\alpha\beta}$ is a constant symmetric non-singular matrix. The tensor has to satisfy the equivalent of equations (15.40) or (*A*15.7):

$$D^k_\alpha\partial_k g^{ij} = g^{ik}\partial_k D^j_\alpha + g^{jk}\partial_k D^i_\alpha .$$

We replace g^{ij} by the form (*A*15.14):

$$m^{\beta\gamma}D^k_\alpha\left(\partial_k D^i_\beta D^j_\gamma + \partial_k D^j_\beta D^i_\gamma\right) = m^{\beta\gamma}D^k_\gamma\left(D^i_\beta\partial_k D^j_\alpha + D^j_\beta\partial_k D^i_\alpha\right),$$

and use the Lie algebra commutation relations (15.43). We then obtain

$$m^{\beta\gamma}f^\delta_{\alpha\beta}\left(D^j_\gamma D^i_\delta + D^i_\gamma D^j_\delta\right) = 0\,. \tag{A15.15}$$

Exchanging $\gamma \leftrightarrow \delta$ in one of the terms, we see that the equation is satisfied if $m^{\alpha\beta}$ is a solution to the numerical equation,

$$m^{\gamma\beta}f^\delta_{\beta\alpha} = f^\gamma_{\alpha\beta}m^{\beta\delta}.$$

In a basis in which the generators 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 $t^a\,|0\rangle$ transforms under an irreducible representation of H, the matrix $\boldsymbol{\mu}$ defined by equation (15.13),

$$\mu_{ab} = -\,\langle 0|\, t^a t^b\,|0\rangle\,, \tag{A15.16}$$

is diagonal

$$\mu_{ab} = \mu\delta_{ab}\,.$$

This provides another proof of the uniqueness of the metric.

The tensor $g^{ij}(\varphi)$ (equation (A15.14)) is a possible inverse metric tensor. In the case of symmetric spaces the unique metric, in a basis in which the generators t^α are orthogonal by the trace, is thus,

$$g^{ij}(\varphi) = D^i_\alpha(\varphi)D^j_\alpha(\varphi), \tag{A15.17}$$

up to the normalization.

A15.3.2 Manifolds embedded in euclidean space

If we know an embedding of a Riemannian manifold $\mathfrak{M}$ in euclidean space, we can describe it with constrained euclidean coordinates, as we have done in the case of homogeneous spaces in Section 15.2.

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:

$$\sigma^s = \sigma^s(\varphi). \tag{A15.18}$$

The metric tensor in this representation is obtained from

$$g_{ij}(\varphi)\mathrm{d}\varphi^i\mathrm{d}\varphi^j = \mathrm{d}\varphi^i\mathrm{d}\varphi^i + \mathrm{d}\sigma^s\mathrm{d}\sigma^s, \tag{A15.19}$$

and, therefore,

$$g_{ij}(\varphi) = \delta_{ij} + \partial_i\sigma^s\partial_j\sigma^s. \tag{A15.20}$$

A short calculation shows that the connection has the simple form

$$\Gamma^i_{jk} = g^{il}\partial_l\sigma^s\partial_j\partial_k\sigma^s, \tag{A15.21}$$

and the curvature tensor is given by

$$R_{ijkl} = \left(\partial_i\partial_k\sigma^s\partial_j\partial_l\sigma^s - \partial_i\partial_l\sigma^s\partial_j\partial_k\sigma^s\right)\left(\mathrm{tr}\left(\mathbf{g}^{-1}\right) - N + 1\right), \tag{A15.22}$$

in which N is the dimension of $\mathfrak{M}$ and $\mathrm{tr}\,\mathbf{g}^{-1}$ the trace of the inverse of the metric:

$$\mathrm{tr}\,\mathbf{g}^{-1} = \mathrm{tr}_{st}\left[\delta_{st} + \partial_k\sigma^s\partial_k\sigma^t\right]^{(-1)}. \tag{A15.23}$$

A15.4 Symmetric Spaces: Classification

In this appendix we examine a few simple properties of symmetric spaces and provide some elements of classification.

A15.4.1 Definition

Let us consider a semi-simple compact Lie group G and assume that we have constructed a non-trivial involutive automorphism of G, which to an element g of G associates an element $\bar{g}$:

$$(\overline{g_1 g_2}) = \bar{g}_1 \bar{g}_2\,, \qquad \text{with } \bar{\bar{g}} = g\,. \tag{A15.24}$$

We consider the coset space G/H obtained by taking for H the subgroup of invariant elements under the automorphism,

$$\overline{H} \equiv H\,. \tag{A15.25}$$

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. Even elements belong by definition 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. We then have the rules

$$\begin{aligned} \bar{t}^a &= -t^a, \qquad & t^a &\in \mathcal{L}(G/H); \\ \bar{\tau}^\alpha &= \tau^\alpha, \qquad & \tau^\alpha &\in \mathcal{L}(H). \end{aligned} \tag{A15.26}$$

It follows that

$$[\tau^\alpha, \tau^\beta] = f_{\alpha\beta\gamma}\tau^\gamma, \tag{A15.27}$$
$$[\tau^\alpha, t^b] = f_{\alpha bc} t^c, \tag{A15.28}$$
$$[t^a, t^b] = f_{ab\gamma}\tau^\gamma. \tag{A15.29}$$

Note that in the case of a compact group only the last set ($A15.29$) of commutation relations is characteristic of a symmetric space, since ($A15.28$) is then a consequence of ($A15.27$) and the antisymmetry of $f_{\alpha bc}$.

Preliminary remarks

(i) We will consider symmetric spaces derived from non-simple groups G, but we want to exclude the possibility that

$$\begin{aligned} G &= G_1 \times G_2\,, \\ H &= H_1 \times H_2 \ \text{ with } \ H_1 \subset G_1\,, H_2 \subset G_2\,, \end{aligned} \tag{A15.30}$$

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, assume a generator τ_δ cannot be obtained. We can then 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 $\mathcal{L}(H)$ which can be obtained as commutators ($A15.29$). We are exactly in the situation we just excluded.

A15.4.2 A basic property

The purpose of this appendix is not to present a complete description 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 let us assume that there exists a subgroup G' of G which contains H, 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 $\left[t''^a, t'^b\right]$ belongs to $\mathcal{L}(H)$ from (A15.29). However, the relations (A15.28) 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, let us 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 (A15.29) 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 constructed the space G/H in the manner described in Section 15.1.1. From the previous results we conclude that the field $\phi(x)$ of equation (15.4),

$$\phi_i(x) = \left[\exp\left(\sum_{\alpha=1}^{l} \xi_a(x) t^a\right)\right]_{ij} \phi_j^c ,$$

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, to each symmetric space is associated a unique classical model. The quantum model will be defined in terms of a unique coupling constant, and the perturbative expansion is equally unique. The general arguments on homogeneous spaces given before tell us that two renormalization constants are sufficient to renormalize the model.

We now describe symmetric spaces corresponding to orthogonal and unitary groups G.

A15.4.3 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 (A15.30), 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

$$(\overline{g_1, g_2}) = (g_2, g_1) \ , \qquad g_1 \in \mathfrak{G}_1 \, , \ \mathfrak{G}_2 \in G_2 \, .$$

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 $g_1^{-1} g_2$.

A15.4.4 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

$$\begin{aligned} g_1^* g_2^* &= (g_1 g_2)^* \, , \\ (g^*)^* &= g \, . \end{aligned}$$

The condition $\bar{\bar{g}} = g$ then has the form

$$g_0^{-1} \left(g_0^{-1}\right)^* g g_0^* g_0 = g \, .$$

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:

$$g_0^* g_0 = \lambda \mathbf{1} \, ,$$

$\mathbf{1}$ being the unit matrix in the defining representation, and λ a phase factor which reduces to $\lambda = \pm 1$ for orthogonal groups.

The subgroup H is defined by the invariant group elements h:

$$\bar{h} = g_0^{-1} h^* g_0 = h \quad \text{or} \quad h^* g_0 = h g_0 \, .$$

A realization of G/H in the group space will be given by group elements of the form

$$g = \left(f^{-1}\right)^* g_0 f \qquad \forall f \in G \, . \tag{A15.31}$$

Indeed, if we multiply f on the left by an element of H, g is not modified.

Note that these group elements satisfy

$$g^* g = \lambda \mathbf{1} \, . \tag{A15.32}$$

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 $\lambda = \pm 1$.

(i) $\lambda = +1$

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 clearly $O(p) \times O(N-p)$.

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) $\lambda = -1$

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 λ is irrelevant. Taking $\lambda = 1$, one sees 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 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 $g_0^* g_0 = \lambda \mathbf{1}$ then implies $\lambda = \pm 1$. If we take $\lambda = +1$, we can diagonalize g_0 by an orthogonal transformation and then set it equal to one 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 $\lambda = -1$, we again see that we must have taken N even

$$N = 2N'.$$

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

$$\begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix} h = h^* \begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix},$$

which is by definition the symplectic subgroup $Sp(N)$ of $U(N)$.

Orthogonal groups. The group G is $O(N)$, the star automorphism is the identity and $\lambda = \pm 1$.

(i) $\lambda = +1$

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 clearly $O(p) \times O(N - p)$.

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) $\lambda = -1$

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 λ is irrelevant. Taking $\lambda = 1$, one sees 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 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 $g_0^* g_0 = \lambda \mathbf{1}$ then implies $\lambda = \pm 1$. If we take $\lambda = +1$, we can diagonalize g_0 by an orthogonal transformation and then set it equal to one 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 $\lambda = -1$, we again see that we must have taken N even

$$N = 2N'.$$

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

$$\begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix} h = h^* \begin{bmatrix} 0 & \mathbf{1}_{N'} \\ -\mathbf{1}_{N'} & 0 \end{bmatrix},$$

which is by definition the symplectic subgroup $Sp(N)$ of $U(N)$.

16 ST AND BRS SYMMETRIES, STOCHASTIC FIELD EQUATIONS

In Section 15.3 we have introduced a transformation depending on anticommuting parameters, to prove the geometric stability of homogeneous spaces under renormalization. There is a set of topics, stochastic field equations, gauge theories, in which similar transformations are met. These new problems have one common feature: they all involve a constraint equation to which, by a set of formal transformations, is associated a quantum action. This action has an anticommuting type symmetry which has no geometric origin but is merely a consequence of these transformations.

In this chapter, we first discuss this mathematical structure from a rather formal point of view, using a notation adapted to a finite number of degrees of freedom. We explain the appearance of Slavnov–Taylor symmetry, which is a conventional non-linear symmetry, in the integral representation of constraint equations. We then show how it leads to a symmetry with anticommuting parameters first discovered in quantized gauge theories by Becchi, Rouet and Stora and, therefore, called BRS symmetry. Its generator has a vanishing square, and generalizes exterior differentiation. We show that this symmetry is remarkably stable against a number of algebraic deformations and this explains its role in the context of stochastic equations. In some cases it can be expressed in compact form by introducing Grassmann coordinates. We show how BRS symmetry can encode the compatibility conditions of a system of linear first order differential equations. We exhibit the special form BRS symmetry takes when the constraint equations apply to group manifolds.

In field theories, examples of constraint or stochastic field equations will be met. For many problems related to perturbation theory, divergences and renormalization, it is more convenient to work with an action and a functional integral rather than with the equation directly, because then standard methods of quantum field theory become available. Therefore, in Section 16.6 we generalize the formalism to an infinite number of degrees of freedom, and discuss the implications of BRS symmetry in the form of WT identities.

A specially important class of stochastic equations corresponds to Langevin equations, extensions to Field Theory of the equations introduced in Chapter 4. They have been proposed to describe the dynamics of critical phenomena (see Chapter 36), or as an alternative method of quantization which could be useful in cases where ordinary methods lead to difficult problems like gauge theories. In all cases divergences appear in perturbative calculations and it is necessary to understand how these equations renormalize. For this purpose we construct the associated action, often called in this context dynamic action. The BRS symmetry of the dynamic action and its consequences in the form of WT identities will be used to prove that under some general conditions the structure of the Langevin equation is stable under renormalization.

The formalism presented here may seem rather heavy when applied to simple models, but it enables us to exhibit general structures with many different applications.

16.1 Slavnov–Taylor (ST) Symmetry

Let φ^α be a set of dynamical variables satisfying a system of equations,

$$E_\alpha(\varphi) = 0\,, \tag{16.1}$$

where 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 $\varphi^\alpha = \varphi^\alpha(E)$. This implies in particular that equation (16.1) 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 $E_{\alpha\beta}$,

$$E_{\alpha\beta} \equiv \partial_\beta E_\alpha\,, \tag{16.2}$$

does not vanish and thus we choose $E_\alpha(\varphi)$ such that it is positive.

Note that it will be convenient throughout this chapter to use the notation $\partial/\partial\varphi^\alpha \mapsto \partial_\alpha$.

For any function $F(\varphi)$ we now derive a first formal expression for $F(\varphi_s)$ which does not involve solving equation (16.1) explicitly. We start from the trivial identity

$$F(\varphi_s) = \int \left\{\prod_\alpha \mathrm{d}E^\alpha\, \delta(E_\alpha)\right\} F\left(\varphi(E)\right),$$

where $\delta(E)$ is Dirac's δ-function. We then change variables $E \mapsto \varphi$. This generates the jacobian $\mathcal{J}(\varphi) = \det \mathbf{E} > 0$. Thus,

$$F(\varphi_s) = \int \left\{\prod_\alpha \mathrm{d}\varphi^\alpha\, \delta\left[E_\alpha(\varphi)\right]\right\} \mathcal{J}(\varphi)\, F(\varphi). \tag{16.3}$$

An invariant measure. The measure $\prod_\alpha \mathrm{d}E_\alpha$ is the invariant measure for the group of translations $E_\alpha \mapsto E_\alpha + \nu_\alpha$. It follows that the measure $\mathrm{d}\rho(\varphi)$,

$$\mathrm{d}\rho(\varphi) = \mathcal{J}(\varphi) \prod_\alpha \mathrm{d}\varphi^\alpha\,, \tag{16.4}$$

is the invariant measure for the translation group realized non-linearly on the new coordinates φ_α (provided ν_α is small enough):

$$\varphi^\alpha \mapsto \varphi'^\alpha \quad \text{with} \quad E_\alpha(\varphi') - \nu_\alpha = E_\alpha(\varphi). \tag{16.5}$$

With the corresponding invariance is associated, in gauge theories, the Slavnov–Taylor symmetry (Chapter 21). The infinitesimal form of the transformation can be written more explicitly:

$$\delta\varphi^\alpha = [E^{-1}(\varphi)]^{\alpha\beta}\nu_\beta\,. \tag{16.6}$$

We shall see later that these rather straightforward identities lead, when applied to field theories, to a number of useful identities.

Reciprocal property. Conversely we can characterize the general form of non-linear representations of the translation group. We write an infinitesimal group transformation,

$$\delta\varphi^\alpha = [M^{-1}(\varphi)]^{\alpha\beta}\nu_\beta\,, \tag{16.7}$$

in which the matrix $M_{\alpha\beta}(\varphi)$ has to be determined. Following the strategy explained in Section 15.3, we impose to the differential operators Δ^α,

$$\Delta^\alpha = [M^{-1}(\varphi)]^{\beta\alpha} \frac{\partial}{\partial\varphi^\beta}, \tag{16.8}$$

to form a representation of the Lie algebra of the translation group, that is, commute:

$$[\Delta^\alpha, \Delta^\beta] = 0 \,. \tag{16.9}$$

This implies

$$[M^{-1}]^{\gamma\alpha}[M^{-1}\partial_\gamma M M^{-1}]^{\delta\beta} = [M^{-1}]^{\gamma\beta}[M^{-1}\partial_\gamma M M^{-1}]^{\delta\alpha} \,.$$

Multiplying by $M_{\alpha\alpha'}M_{\beta\beta'}M_{\delta\delta'}$ and summing over α, β, δ, we find

$$\partial_\beta M_{\delta\alpha} - \partial_\alpha M_{\delta\beta} = 0 \,. \tag{16.10}$$

This implies (for a simply connected φ-manifold, an implicit assumption throughout the whole chapter) that $M_{\alpha\beta}$ has the form

$$M_{\alpha\beta} = \partial_\beta E_\alpha \,. \tag{16.11}$$

We now characterize the invariant measure $\mathcal{J}(\varphi)\mathrm{d}\varphi$ 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 (16.7):

$$\partial_\alpha \mathcal{J}[M^{-1}(\varphi)]^{\alpha\beta} + \mathcal{J}\partial_\alpha[M^{-1}(\varphi)]^{\alpha\beta} = 0 \,. \tag{16.12}$$

This yields a system of partial differential equations for the function $\mathcal{J}(\varphi)$:

$$\partial_\alpha \ln \mathcal{J} = [M^{-1}]^{\gamma\beta}\partial_\gamma M_{\beta\alpha} \,. \tag{16.13}$$

Equation (16.10) is an integrability condition for equation (16.13). We can use it to rewrite equation (16.13):

$$\partial_\alpha \ln \mathcal{J} = \partial_\alpha \ln \det \mathbf{M} \,, \tag{16.14}$$

which has the solution

$$\mathcal{J} = \text{const. } \det \mathbf{M} \,. \tag{16.15}$$

Similar identities will appear again later in this chapter.

16.2 Constraints and BRS Symmetry

In Field Theory the non-linear and as we shall see later non-local character of the transformations (16.5) and (16.6) is the source of many technical difficulties. Remarkably enough the infinitesimal transformations (16.6) can be replaced by a linear anticommuting type transformation at the price of introducing additional variables.

We again start from the identity (16.3) and first replace the δ-function by its Fourier representation:

$$\prod_\alpha \delta\left[E_\alpha(\varphi)\right] = \int \prod_\alpha \frac{\mathrm{d}\bar{\varphi}^\alpha}{2i\pi}\, \mathrm{e}^{-\bar{\varphi}^\alpha E_\alpha(\varphi)}, \tag{16.16}$$

where the $\bar{\varphi}$ integration runs along the imaginary axis. Moreover, a determinant can be written as an integral over Grassmann variables $\bar{c}^\alpha$ and c^α (see Section 1.7):

$$\det \mathbf{E} = \int \prod_\alpha \left(\mathrm{d}\bar{c}^\alpha \mathrm{d}c^\alpha\right) \exp\left(c^\alpha E_{\alpha\beta}\bar{c}^\beta\right). \tag{16.17}$$

The expression (16.3) then becomes

$$F(\varphi_\mathrm{s}) = \mathcal{N} \int \prod_\alpha \left(\mathrm{d}\varphi^\alpha \mathrm{d}\bar{\varphi}^\alpha \mathrm{d}\bar{c}^\alpha \mathrm{d}c^\alpha\right) F(\varphi) \exp\left[-S(\varphi,\bar{\varphi},c,\bar{c})\right], \tag{16.18}$$

in which $\mathcal{N}$ is a constant normalization factor and $S(\varphi,\bar{\varphi},c,\bar{c})$ the function (and element of the Grassmann algebra)

$$S(\varphi,\bar{\varphi},c,\bar{c}) = \bar{\varphi}^\alpha E_\alpha(\varphi) - c^\alpha E_{\alpha\beta}(\varphi)\bar{c}^\beta . \tag{16.19}$$

Somewhat surprisingly, the function S has a new type of symmetry, which is directly related to the invariance of the measure (16.4) under the group of transformations (16.6), and which we describe below.

BRS symmetry. The BRS symmetry, first discovered in the quantization of gauge theories by Becchi, Rouet and Stora (see Chapters 19 and 21), is a Grassmann symmetry in the sense that the parameter ε of the transformation is an anticommuting constant, an additional generator of the Grassmann algebra. The variations of the various dynamic variables are

$$\delta\varphi^\alpha = \varepsilon\bar{c}^\alpha , \qquad \delta\bar{c}^\alpha = 0 , \tag{16.20a}$$

$$\delta c^\alpha = \varepsilon\bar{\varphi}^\alpha , \qquad \delta\bar{\varphi}^\alpha = 0 \tag{16.20b}$$

with

$$\varepsilon^2 = 0 , \qquad \varepsilon\bar{c}^\alpha + \bar{c}^\alpha\varepsilon = 0 , \qquad \varepsilon c^\alpha + \varepsilon c^\alpha = 0 .$$

The transformation is obviously *nilpotent* of vanishing square: $\delta^2 = 0$.

The BRS transformation can be represented by a Grassmann differential operator $\mathcal{D}$, when acting on functions of $\{\varphi,\bar{\varphi},c,\bar{c}\}$:

$$\mathcal{D} = \bar{c}^\alpha \frac{\partial}{\partial\varphi^\alpha} + \bar{\varphi}^\alpha \frac{\partial}{\partial c^\alpha} . \tag{16.21}$$

The nilpotency of the BRS transformation is then expressed by the identity

$$\mathcal{D}^2 = 0 . \tag{16.22}$$

The differential operator $\mathcal{D}$ is a cohomology operator, generalization of the exterior differentiation of differential forms (see also Sections 1.4 and 22.1). In particular, the first term $\bar{c}^\alpha \partial_\alpha$ in the BRS operator is identical to the differentiation of forms in a formalism in which the Grassmann variables $\bar{c}^\alpha$ are introduced as external variables to exhibit the antisymmetry of the corresponding tensors.

Equation (16.22) implies that all quantities of the form $\mathcal{D}Q(\varphi, \bar{\varphi}, c, \bar{c})$, quantities we call *BRS exact*, are BRS invariant. We immediately verify that the function S defined by equation (16.19) is BRS exact:

$$S = \mathcal{D}\left[c^\alpha E_\alpha(\varphi)\right] . \tag{16.23}$$

It follows that S is BRS invariant,

$$\mathcal{D}S = 0 . \tag{16.24}$$

The reciprocal property, the meaning and implications of the BRS symmetry will be discussed in the coming sections.

These properties play an important role, in particular, in the discussion of the renormalization of gauge theories.

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 $\varphi, \bar{\varphi}$. The equation $\mathcal{E}_\alpha(c) = 0$ leads to

$$S = \mathcal{D}[\varphi^\alpha \mathcal{E}_\alpha(c)] = \bar{c}^\alpha \mathcal{E}_\alpha(c) + \varphi^\alpha \frac{\partial \mathcal{E}_\alpha}{\partial c^\beta} \bar{\varphi}^\beta . \tag{16.25}$$

ST and BRS symmetries. A simple way of understanding the relation between the transformations (16.6) and (16.20) is to note that

$$\det \mathbf{E} \left(E^{-1}\right)^{\alpha\beta} = \int \mathrm{d}\bar{c}\, \mathrm{d}c\, c^\beta \bar{c}^\alpha \exp\left(c^\gamma E_{\gamma\delta} \bar{c}^\delta\right) . \tag{16.26}$$

Therefore, if we factorize the integral over c and $\bar{c}$, we can rewrite inside the integral

$$\delta\varphi^\alpha = \left(E^{-1}\right)^{\alpha\beta} \nu_\beta = -\bar{c}^\alpha \left(c^\beta \nu_\beta\right) . \tag{16.27}$$

BRS transformations correspond to $-c^\alpha \nu_\alpha \longmapsto \varepsilon$.

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 BRS symmetry are the same. The BRS symmetry extends the transformations to functions of φ, $\bar{\varphi}$, c and $\bar{c}$. This extension is useful for two reasons:

(i) The transformations (16.20) are linear, while the transformation (16.6) is non-linear. We have already seen in the case of the non-linear σ-model that non-linear transformations could be linearized at the price of introducing auxiliary fields.

(ii) More important, in field theories the transformations (16.20) will be local, in contrast to the transformation (16.6).

These two properties greatly simplify the discussion of renormalization of various field theories.

16.3 Grassmann Coordinates, Gradient Equations

A more compact representation of BRS transformations is obtained by introducing a Grassmann coordinate θ and then the following two functions of θ:

$$\phi^\alpha(\theta) = \varphi^\alpha + \theta \bar{c}^\alpha , \qquad C^\alpha(\theta) = c^\alpha + \theta \bar{\varphi}^\alpha . \tag{16.28}$$

With this notation the transformations (16.20) simply become a translation of θ:

$$\begin{aligned} \delta\phi^\alpha(\theta) &= \varepsilon \frac{\partial \phi^\alpha}{\partial \theta} = \phi^\alpha(\theta + \varepsilon) - \phi^\alpha(\theta), \\ \delta C^\alpha(\theta) &= \varepsilon \frac{\partial C^\alpha}{\partial \theta} = C^\alpha(\theta + \varepsilon) - C^\alpha(\theta). \end{aligned} \tag{16.29}$$

In particular, the BRS operator $\mathcal{D}$ is represented by $\partial/\partial\theta$:

$$\mathcal{D} \mapsto \frac{\partial}{\partial \theta} .$$

We note the expansion

$$C^\alpha(\theta) E_\alpha(\phi(\theta)) = c^\alpha E_\alpha(\varphi) + \theta \left[\bar{\varphi}^\alpha E_\alpha(\varphi) - c^\alpha \frac{\partial E_\alpha}{\partial \varphi^\beta} \bar{c}^\beta \right] . \tag{16.30}$$

Thus,

$$S(\varphi, \bar{\varphi}, c, \bar{c}) = \frac{\partial}{\partial \theta} \left[C^\alpha(\theta) E_\alpha(\phi(\theta)) \right] . \tag{16.31}$$

We recover equation (16.23) in a different notation.

Because in the case of Grassmann variables integration and differentiation are identical operations, an integration over θ also selects the coefficient of θ. Therefore, $S(\varphi, \bar{\varphi}, c, \bar{c})$ can be rewritten as an integral over θ:

$$S(\varphi, \bar{\varphi}, c, \bar{c}) = \int \mathrm{d}\theta \, C^\alpha(\theta) E_\alpha(\phi(\theta)). \tag{16.32}$$

In this expression the BRS 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 (16.28), as in equation (16.32), each integration over θ is associated with a factor $C^\alpha(\theta)$.

Gradient equations. The two Grassmann variables, $\bar{c}^\alpha$ and c^α, that we have introduced in the preceding section play, in general, different roles. There is, however, one special situation in which a symmetry is established between them—when the matrix $E_{\alpha\beta}$ is symmetric:

$$E_{\alpha\beta} = E_{\beta\alpha} \iff \partial_\beta E_\alpha = \partial_\alpha E_\beta . \tag{16.33}$$

Hence, with our general assumptions, $E_\alpha(\varphi)$ is itself a gradient; there exists a function $A(\varphi)$ such that

$$E_\alpha(\varphi) = \partial_\alpha A(\varphi). \tag{16.34}$$

The symmetry between c and $\bar{c}$ generates an additional independent BRS symmetry of generator $\bar{\mathcal{D}}$:

$$\bar{\mathcal{D}} = c^\alpha \frac{\partial}{\partial \varphi^\alpha} + \bar{\varphi}^\alpha \frac{\partial}{\partial \bar{c}^\alpha}\,.$$

It is thus natural to introduce two Grassmann variables $\bar{\theta}$ and θ, and a function $\phi^\alpha(\bar{\theta},\theta)$ (and $\bar{\mathcal{D}} \mapsto \partial/\partial\bar{\theta}$):

$$\phi^\alpha(\bar{\theta},\theta) = \varphi^\alpha + \theta\bar{c}^\alpha + c^\alpha\bar{\theta} + \theta\bar{\theta}\bar{\varphi}^\alpha\,. \tag{16.35}$$

In terms of ϕ the expression (16.32) quite generally reads

$$S(\phi) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\bar{\theta}\frac{\partial\phi^\alpha}{\partial\bar{\theta}} E_\alpha\left[\phi(\bar{\theta},\theta)\right]. \tag{16.36}$$

When the function $E(\varphi)$ has the particular form (16.34), it is possible to integrate by parts over $\bar{\theta}$ and the function $S(\phi)$ then takes the remarkable form

$$S(\phi) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,A\left[\phi(\bar{\theta},\theta)\right] = \bar{\mathcal{D}}\mathcal{D}A(\varphi). \tag{16.37}$$

The two symmetries, which correspond to independent translations of θ and $\bar{\theta}$, are here explicit.

16.4 BRS Symmetry and Compatibility Condition, Group Manifolds

Though the first term $\bar{c}^\alpha\partial_\alpha$ in the BRS operator (16.21) is identical to the differentiation of forms, the Grassmann variables $\bar{c}^\alpha$ here are not simply external variables introduced for convenience to exhibit the antisymmetry of the corresponding tensors. Instead, the $\bar{c}^\alpha$'s are additional dynamical variables. In particular, it may be convenient, as we show below, to change variables and to set

$$\varphi^\alpha = F^\alpha(\varphi'), \qquad \bar{c}^\alpha = U^\alpha_\beta(\varphi')\bar{c}'^\beta\,.$$

It is the second equation which is characteristic of the difference. If we now express the transformations (16.20*a*) in these new variables they take a more complicated form (omitting the primes on the new variables):

$$\delta\varphi^\alpha = \varepsilon D^\alpha_\beta(\varphi)\bar{c}^\beta\,, \qquad \delta\bar{c}^\alpha = -\tfrac{1}{2} f^\alpha_{\beta\gamma}(\varphi)\varepsilon\bar{c}^\beta\bar{c}^\gamma, \tag{16.38}$$

where the functions $D^\alpha_\beta(\varphi)$ and $f^\alpha_{\beta\gamma}(\varphi)$, which is antisymmetric in $\beta \leftrightarrow \gamma$, can be expressed in terms of F and U. The BRS generator $\mathcal{D}$ also takes, in these new variables, a more complicated form but equation (16.22) still holds.

More generally, we can consider situations in which physics is described in terms of some variables A^i, themselves functions of the φ^α. Then BRS transformations will have the form

$$\delta A^i = \varepsilon D^i_\beta(A)\bar{c}^\beta\,, \qquad \delta\bar{c}^\alpha = -\tfrac{1}{2} f^\alpha_{\beta\gamma}(A)\varepsilon\bar{c}^\beta\bar{c}^\gamma. \tag{16.39}$$

The condition (16.22), $\mathcal{D}^2 = 0$, expressed directly in terms of D and f implies two equations obtained by identifying the terms cubic and quadratic in the Grassmann variables $\bar{c}^\alpha$. The properly antisymmetrized coefficients vanish:

$$\left\{f^\delta_{\alpha\beta} f^\epsilon_{\gamma\delta} + D^i_\alpha\partial_i f^\epsilon_{\beta\gamma}\right\}_{\alpha\beta\gamma} = 0\,, \tag{16.40}$$

$$D^j_\alpha\partial_j D^i_\beta - D^j_\beta\partial_j D^i_\alpha = f^\delta_{\alpha\beta} D^i_\delta\,, \tag{16.41}$$

where the global subscript $\alpha\beta\gamma$ in (16.40) means antisymmetrized in the three indices. In equation (16.41) we immediately recognize equation (15.43) of Section 15.3, that is, the commutation relation between generators of a Lie algebra in a non-linear representation. Moreover, if $f^{\alpha}_{\beta\gamma}$ is independent of A, equation (16.40) is simply the Jacobi identity for the structure constants of the Lie algebra.

Compatibility conditions. The extension we find here can be understood in the following way. We consider the set of first order partial differential equations

$$\Delta_\alpha S(A) = 0\,, \tag{16.42}$$

where the Δ_α are the differential operators,

$$\Delta_\alpha = D^i_\alpha(A)\frac{\partial}{\partial A^i}.$$

As discussed in Sections 13.1.1 and 15.3 the system (16.42) is called compatible if the equations $[\Delta_\alpha, \Delta_\beta]S = 0$ are linear combinations of the initial equations (16.42). Then,

$$[\Delta_\alpha, \Delta_\beta] = f^{\gamma}_{\alpha\beta}(A)\Delta_\gamma\,. \tag{16.43}$$

We have encountered up to now only examples where the structure constants $f^{\gamma}_{\alpha\beta}$ were A-independent, but this is not the general situation.

Equation (16.43) itself has an integrability condition, the Jacobi identities since the l.h.s. is a commutator. A short calculation yields the condition (16.40). Therefore, the nilpotency of the BRS operator encodes the compatibility of the linear system (16.42) and the Lie algebra structure. Note, finally, that the equation (16.42) is equivalent to $\mathcal{D}S(A) = 0$. If it has non-trivial solutions, these solutions cannot be cast into the form $S(A) = \mathcal{D}S'(A, c)$, that is, they are not BRS exact.

BRS symmetry and group manifolds. Group manifolds provide a simple example of such a situation. If the variables φ^α parametrize a group element $g(\varphi)$ in some matrix representation, it is convenient to rewrite BRS transformations on $g(\varphi)$ directly. This can be most easily done by noting that with $\phi(\theta)$ (defined by equation (16.28)) we can associate some group element $\mathcal{G}(\theta)$, on which BRS transformations according to equation (16.29) read

$$\delta\mathcal{G}(\theta) = \varepsilon\frac{\partial\mathcal{G}}{\partial\theta}\,. \tag{16.44}$$

However, since $\mathcal{G}(\theta)$ is a group element, it is natural to parametrize it in the form

$$\mathcal{G}(\theta) = \exp(\theta\bar{c})g = (1 + \theta\bar{c})g, \tag{16.45}$$

in which $\bar{c}$ now is a Grassmann matrix belonging to the Lie algebra of the group. In component form, the transformation (16.44) then becomes

$$\delta g = \varepsilon\bar{c}g\,, \tag{16.46}$$

$$\delta\bar{c} = \varepsilon\bar{c}^2. \tag{16.47}$$

Introducing matrices t_α, generators of the Lie algebra, and parametrizing $\bar{c}$ as

$$\bar{c} = \bar{c}^\alpha t_\alpha, \tag{16.48}$$

we can rewrite equation (16.47) as

$$\delta \bar{c}^{\alpha} = \tfrac{1}{2} \varepsilon f^{\alpha}_{\beta\gamma} \bar{c}^{\beta} \bar{c}^{\gamma} . \tag{16.49}$$

In this form, we recognize the transformation (15.49). Moreover, the transformation (16.46) applied to the matrix $\mathbf{R}$ of the representation (15.6) leads to the transformation (15.45) for the field $\varphi(x)$. We have thus found a geometric interpretation to the transformations used in the derivation of WT identities for homogeneous spaces in an arbitrary system of coordinates.

16.5 Stochastic Equations

We now assume that equation (16.1) depends on a set of stochastic variables ν_a, the "noise", with normalized probability distribution $\mathrm{d}\rho(\nu)$:

$$E_{\alpha}(\varphi, \nu) = 0 . \tag{16.50}$$

The solution φ^{α} of the equation becomes a stochastic variable. Quantities of interest are now expectation values of functions of φ:

$$\langle F(\varphi) \rangle = \int \mathrm{d}\rho(\nu) \prod_{\alpha} \mathrm{d}\varphi^{\alpha}\, \delta\left[E_{\alpha}(\varphi, \nu)\right] \det \mathbf{E}\, F(\varphi). \tag{16.51}$$

After the set of transformations described in Section 16.2, this representation becomes

$$\langle F(\varphi) \rangle \propto \int \mathrm{d}\rho(\nu) \prod_{\alpha} (\mathrm{d}\varphi^{\alpha} \mathrm{d}\bar{\varphi}^{\alpha} \mathrm{d}\bar{c}^{\alpha} \mathrm{d}c^{\alpha})\, F(\varphi) \exp\left[-S(\varphi, \bar{\varphi}, c, \bar{c}, \nu)\right] \tag{16.52}$$

with S given by equation (16.19):

$$S = \bar{\varphi}^{\alpha} E_{\alpha}(\varphi, \nu) - c^{\alpha} E_{\alpha\beta}(\varphi, \nu) \bar{c}^{\beta} . \tag{16.53}$$

Let us introduce the function $\Sigma(\varphi, \bar{\varphi}, c, \bar{c})$ obtained after noise averaging:

$$\langle F(\varphi) \rangle \propto \int \prod_{\alpha} (\mathrm{d}\varphi^{\alpha} \mathrm{d}\bar{\varphi}^{\alpha} \mathrm{d}\bar{c}^{\alpha} \mathrm{d}c^{\alpha})\, F(\varphi) \exp\left[-\Sigma(\varphi, \bar{\varphi}, c, \bar{c})\right] \tag{16.54}$$

with

$$\exp\left[-\Sigma(\varphi, \bar{\varphi}, c, \bar{c})\right] = \int \mathrm{d}\rho(\nu) \exp\left[-S(\varphi, \bar{\varphi}, c, \bar{c}, \nu)\right] . \tag{16.55}$$

We have shown that S has a BRS symmetry. Applying the BRS operator (16.21) on both sides of equation (16.55) we conclude that $\Sigma(\varphi, \bar{\varphi}, c, \bar{c})$ is still BRS symmetric, although it no longer has the simple form (16.53), that is, a function linear in $\bar{\varphi}$ and $c\bar{c}$

$$\mathcal{D}\Sigma = 0 . \tag{16.56}$$

Moreover, because S is BRS exact the function Σ is also BRS exact, as simple algebraic manipulations based on the identity

$$f(\mathcal{D}X) = f(0) + \mathcal{D}\left[X g(\mathcal{D}X)\right] \quad \text{with} \quad g(x) = \frac{f(x) - f(0)}{x} ,$$

show.

Remark. The function S involves only a Grassmann combination of the form $c\bar{c}$. Therefore, if we multiply c by a phase $\mathrm{e}^{i\nu}$ and $\bar{c}$ by the complex conjugated phase $\mathrm{e}^{-i\nu}$, 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 (1.77) also shows.

16.5.1 Stochastic equations linear in the noise

A simple example. In the coming chapters stochastic equations of the simple algebraic form

$$E_\alpha(\nu,\varphi) \equiv E_\alpha(\varphi) - \nu_\alpha\,, \tag{16.57}$$

will be encountered. Introducing the Laplace transform of the measure $\mathrm{d}\rho(\nu)$,

$$\mathrm{e}^{w(\bar\varphi)} = \int \mathrm{d}\rho(\nu)\,\mathrm{e}^{\bar\varphi^\alpha \nu_\alpha}\,, \tag{16.58}$$

we obtain for the function Σ defined by (16.55)

$$\Sigma(\varphi,\bar\varphi,c,\bar c) = -w(\bar\varphi) + \bar\varphi^\alpha E_\alpha(\varphi) - c^\alpha E_{\alpha\beta}\bar c^\beta \tag{16.59a}$$

$$= \mathcal{D}\tilde\Sigma\,, \quad \tilde\Sigma = c^\alpha\left[E_\alpha(\varphi) - \frac{\partial}{\partial\bar\varphi^\alpha}\int_0^1 \mathrm{d}s\,w(s\bar\varphi)\right]. \tag{16.59b}$$

This is a minimal modification of expression (16.19).

Remarks.

(i) After integration over the noise, the expression of the function Σ in the notation of Grassmann coordinates is, in general, rather complicated. However, in the case of equation (16.57) with gaussian noise the additional term $w(\bar\varphi) = \frac{1}{2}w_{\alpha\beta}\bar\varphi^\alpha\bar\varphi^\beta$ is represented in the notation (16.35) by

$$\frac{1}{2}w_{\alpha\beta}\bar\varphi^\alpha\bar\varphi^\beta = \int \mathrm{d}\bar\theta\mathrm{d}\theta\,\frac{1}{2}w_{\alpha\beta}\frac{\partial\phi^\alpha}{\partial\bar\theta}\frac{\partial\phi^\beta}{\partial\theta}.$$

(ii) In the latter case, it is also possible to integrate explicitly over the $\bar\varphi$ variables. The resulting integrand corresponds to

$$\Sigma(\varphi,c,\bar c) = \tfrac{1}{2}E_\alpha(\varphi)[w^{-1}]^{\alpha\beta}E_\beta(\varphi) - c^\alpha E_{\alpha\beta}(\varphi)\bar c^\beta\,.$$

The BRS transformation of c is now non-linear:

$$\delta_{\mathrm{BRS}}c^\alpha = \varepsilon[w^{-1}]^{\alpha\beta}E_\beta(\varphi)\,.$$

We note that in this form the BRS transformation has a vanishing square only when φ is a solution of the equation $E(\varphi) = 0$. We conclude that the property $\mathcal{D}^2 = 0$ of BRS transformations is not true in all formulations and may be satisfied only after the introduction of auxiliary variables.

The general linear case. A slightly more general form will also be encountered:

$$E_\alpha(\varphi,\nu) = E_\alpha(\varphi) - e^a_\alpha(\varphi)\nu_a\,. \tag{16.60}$$

Then,

$$\partial_\beta E_\alpha(\varphi,\nu) = \partial_\beta E_\alpha(\varphi) - \partial_\beta e^a_\alpha(\varphi)\nu_a\,.$$

Here, we only consider the special example of gaussian stochastic variables ν with probability distribution,

$$\mathrm{d}\rho(\nu) = \Big(\prod_a \mathrm{d}\nu_a\Big)\exp(-\tfrac{1}{2}\nu_a[\Omega^{-1}]^{ab}\nu_b). \tag{16.61}$$

After integration over ν and some algebra we find that Σ can be written as

$$\Sigma(\varphi,\bar\varphi,c,\bar c) = \mathcal{D}\widetilde\Sigma(\varphi,\bar\varphi,c,\bar c), \tag{16.62a}$$

$$\widetilde\Sigma(\varphi,\bar\varphi,c,\bar c) = c^\alpha E_\alpha - \tfrac{1}{2}c^\alpha\left(w_{\alpha\beta}(\varphi)\bar\varphi^\beta - w_{\alpha\beta,\gamma}(\varphi)c^\beta\bar c^\gamma\right), \tag{16.62b}$$

with the definitions

$$w_{\alpha\beta} = e^a_\alpha\Omega_{ab}e^b_\beta\,, \qquad w_{\alpha\beta,\gamma} = e^a_\alpha\Omega_{ab}\partial_\gamma e^b_\beta\,.$$

Σ differs from S by the addition of a function quadratic in both $\bar\varphi$ and $c\bar c$.

16.5.2 BRS cohomology

We have seen in Section 16.5.1 that for a general stochastic equation linear in a gaussian noise, the weight function Σ obtained after noise averaging (equations (16.62)) is BRS exact and quadratic in $\{\bar{\varphi}, c\bar{c}\}$. Conversely, one may ask the question: what is the most general form of a function Σ quadratic in $\{\bar{\varphi}, c\bar{c}\}$ and BRS symmetric:

$$\mathcal{D}\Sigma(\varphi, \bar{\varphi}, c, \bar{c}) = 0\,. \tag{16.63}$$

We expand Σ in powers of $c\bar{c}$:

$$\Sigma(\varphi, \bar{\varphi}, c, \bar{c}) = \Sigma_0(\varphi, \bar{\varphi}) + \Sigma_1(\varphi, \bar{\varphi}, c, \bar{c}) + \Sigma_2(\varphi, c, \bar{c}), \tag{16.64}$$

where Σ_0 is quadratic in $\bar{\varphi}$, Σ_1 linear in $c\bar{c}$ and of first degree in $\bar{\varphi}$ and finally Σ_2 quadratic in $c\bar{c}$ (and independent of $\bar{\varphi}$).

We now write the BRS operator $\mathcal{D}$ as a sum $\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_-$ with the definitions

$$\mathcal{D}_+ \equiv \bar{c}^\alpha \partial_\alpha\,, \quad \mathcal{D}_- \equiv \bar{\varphi}^\alpha \frac{\partial}{\partial c^\alpha}\,. \tag{16.65}$$

The operator $\mathcal{D}_+$ corresponds to form differentiation. One verifies

$$\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_-\,, \quad \mathcal{D}_+^2 = \mathcal{D}_-^2 = 0\,, \quad \mathcal{D}_+\mathcal{D}_- + \mathcal{D}_-\mathcal{D}_+ = 0\,. \tag{16.66}$$

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

$$\mathcal{D}_-\Sigma_1 = -\mathcal{D}_+\Sigma_0\,, \tag{16.67a}$$
$$\mathcal{D}_-\Sigma_2 = -\mathcal{D}_+\Sigma_1\,, \tag{16.67b}$$
$$\mathcal{D}_+\Sigma_2 = 0\,. \tag{16.67c}$$

We now parametrize Σ_0 as

$$\Sigma_0 = A(\varphi) + \bar{\varphi}^\alpha E_\alpha(\varphi, \bar{\varphi}) = A(\varphi) + \mathcal{D}_-\left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right].$$

Setting $\bar{\varphi} = 0$ in equation (16.67a) we first find

$$\partial_\alpha A(\varphi) = 0 \;\Rightarrow A(\varphi) = \text{const.}.$$

Then,

$$\mathcal{D}_-\Sigma_1 = -\mathcal{D}_+\mathcal{D}_-\left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] = \mathcal{D}_-\mathcal{D}_+\left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right],$$

which has a solution

$$\Sigma_1 = \mathcal{D}_+\left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] + \mathcal{D}_-\tilde{\Sigma}_1\,,$$

up to possible terms which are not $\mathcal{D}_-$ exact. We now examine this possibility taking into account that Σ_1 has the general form

$$\Sigma_1 = w_{\alpha\beta} c^\alpha \bar{c}^\beta + w_{\alpha,\beta\gamma} \bar{\varphi}^\alpha c^\beta \bar{c}^\gamma\,.$$

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

$$\tilde{\Sigma}_1 = \tfrac{1}{2} w_{\alpha\beta,\gamma} c^\alpha c^\beta \bar{c}^\gamma.$$

16.5.2 BRS cohomology

We have seen in Section 16.5.1 that for a general stochastic equation linear in a gaussian noise, the weight function Σ obtained after noise averaging (equations (16.62)) is BRS exact and quadratic in $\{\bar{\varphi}, c\bar{c}\}$. Conversely, one may ask the question: what is the most general form of a function Σ quadratic in $\{\bar{\varphi}, c\bar{c}\}$ and BRS symmetric:

$$\mathcal{D}\Sigma(\varphi, \bar{\varphi}, c, \bar{c}) = 0 . \tag{16.63}$$

We expand Σ in powers of $c\bar{c}$:

$$\Sigma(\varphi, \bar{\varphi}, c, \bar{c}) = \Sigma_0(\varphi, \bar{\varphi}) + \Sigma_1(\varphi, \bar{\varphi}, c, \bar{c}) + \Sigma_2(\varphi, c, \bar{c}), \tag{16.64}$$

where Σ_0 is quadratic in $\bar{\varphi}$, Σ_1 linear in $c\bar{c}$ and of first degree in $\bar{\varphi}$ and finally Σ_2 quadratic in $c\bar{c}$ (and independent of $\bar{\varphi}$).

We now write the BRS operator $\mathcal{D}$ as a sum $\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_-$ with the definitions

$$\mathcal{D}_+ \equiv \bar{c}^\alpha \partial_\alpha , \quad \mathcal{D}_- \equiv \bar{\varphi}^\alpha \frac{\partial}{\partial c^\alpha} . \tag{16.65}$$

The operator $\mathcal{D}_+$ corresponds to form differentiation. One verifies

$$\mathcal{D} = \mathcal{D}_+ + \mathcal{D}_- , \quad \mathcal{D}_+^2 = \mathcal{D}_-^2 = 0 , \quad \mathcal{D}_+\mathcal{D}_- + \mathcal{D}_-\mathcal{D}_+ = 0 . \tag{16.66}$$

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

$$\mathcal{D}_-\Sigma_1 = -\mathcal{D}_+\Sigma_0 , \tag{16.67a}$$

$$\mathcal{D}_-\Sigma_2 = -\mathcal{D}_+\Sigma_1 , \tag{16.67b}$$

$$\mathcal{D}_+\Sigma_2 = 0 . \tag{16.67c}$$

We now parametrize Σ_0 as

$$\Sigma_0 = A(\varphi) + \bar{\varphi}^\alpha E_\alpha(\varphi, \bar{\varphi}) = A(\varphi) + \mathcal{D}_- \left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] .$$

Setting $\bar{\varphi} = 0$ in equation (16.67a) we first find

$$\partial_\alpha A(\varphi) = 0 \ \Rightarrow A(\varphi) = \text{const.} .$$

Then,

$$\mathcal{D}_-\Sigma_1 = -\mathcal{D}_+\mathcal{D}_- \left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] = \mathcal{D}_-\mathcal{D}_+ \left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] ,$$

which has a solution

$$\Sigma_1 = \mathcal{D}_+ \left[c^\alpha E_\alpha(\varphi, \bar{\varphi})\right] + \mathcal{D}_-\tilde{\Sigma}_1 ,$$

up to possible terms which are not $\mathcal{D}_-$ exact. We now examine this possibility taking into account that Σ_1 has the general form

$$\Sigma_1 = w_{\alpha\beta} c^\alpha \bar{c}^\beta + w_{\alpha,\beta\gamma} \bar{\varphi}^\alpha c^\beta \bar{c}^\gamma .$$

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

$$\tilde{\Sigma}_1 = \tfrac{1}{2} w_{\alpha\beta,\gamma} c^\alpha c^\beta \bar{c}^\gamma .$$

Then equation (16.67b) 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 $\bar{\varphi}$ has the simple 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 BRS exact, up to a constant:

$$\Sigma = \mathcal{D}\left[c^\alpha E_\alpha(\varphi,\bar{\varphi}) + \tfrac{1}{2}w_{\alpha\beta,\gamma}c^\alpha c^\beta \bar{c}^\gamma\right] + \text{const.}. \tag{16.68}$$

A more general analysis. One can look for general solutions Σ of equation (16.63), polynomial in $\bar{c}c$:

$$\Sigma = \sum_{p=0}^{n} \Sigma_p , \tag{16.69}$$

where Σ_p has the form

$$\Sigma_p = c^{\alpha_1}\bar{c}^{\beta_1}\ldots c^{\alpha_p}\bar{c}^{\beta_p}[\Sigma_p(\varphi,\bar{\varphi})]_{\alpha_1,\ldots,\alpha_p;\beta_1,\ldots,\beta_p} .$$

In particular, Σ_p being proportional to p factors $\bar{c}$ can be considered a p-form.

The equation $\mathcal{D}\Sigma = 0$ decomposes into

$$\mathcal{D}_+\Sigma_p = -\mathcal{D}_-\Sigma_{p+1} \quad \text{for } p \le n-1 \quad \text{and} \quad \mathcal{D}_+\Sigma_n = 0 . \tag{16.70}$$

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 has as a solution

$$\Sigma_0 = \mathcal{D}_-\tilde{\Sigma}_0 + \text{const.}.$$

We conclude that Σ can be written as

$$\Sigma = \mathcal{D}\widetilde{\Sigma} + \text{const.} \tag{16.71}$$

with

$$\tilde{\Sigma} = \sum_{p=0}^{n-1} \tilde{\Sigma}_p \,.$$

From this analysis we conclude that in the case of simply connected manifolds, any BRS symmetric function is BRS exact, up to a constant:

$$\mathcal{D}\Sigma = 0 \;\Rightarrow\; \Sigma = \mathcal{D}\widetilde{\Sigma} + \text{const.}\,. \tag{16.72}$$

An important consequence. To illustrate the importance of the property that the weight function Σ is not only BRS invariant but also BRS exact, we consider the following integral Z:

$$Z = \langle 1 \rangle = \int \prod_\alpha (\mathrm{d}\varphi^\alpha \mathrm{d}\bar{\varphi}^\alpha \mathrm{d}\bar{c}^\alpha \mathrm{d}c^\alpha) \exp\left[-\Sigma(\varphi, \bar{\varphi}, c, \bar{c})\right], \tag{16.73}$$

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 (16.50).

To this variation corresponds a variation $\delta\Sigma$ of Σ which, according to our previous analysis, can be written as

$$\delta\Sigma = \mathcal{D}\left[\delta\widetilde{\Sigma}(\varphi, \bar{\varphi}, c, \bar{c})\right]. \tag{16.74}$$

Therefore, δZ has the form

$$\delta Z = \int \prod_\alpha (\mathrm{d}\varphi^\alpha \mathrm{d}\bar{\varphi}^\alpha \mathrm{d}c^\alpha \mathrm{d}\bar{c}^\alpha) \mathcal{D}[\delta\widetilde{\Sigma}] \exp(-\Sigma). \tag{16.75}$$

Since $\mathcal{D}$ is a differential operator we can integrate by parts:

$$\delta Z = \int \prod_\alpha (\mathrm{d}\varphi^\alpha \mathrm{d}\bar{\varphi}^\alpha \mathrm{d}c^\alpha \mathrm{d}\bar{c}^\alpha) \delta\widetilde{\Sigma}\, \mathcal{D}\Sigma \exp(-\Sigma). \tag{16.76}$$

Then, using the BRS symmetry condition $\mathcal{D}\Sigma = 0$, we conclude that

$$\delta Z = 0\,. \tag{16.77}$$

Of course, this result is not surprising since by construction Z is a constant independent of $E_\alpha(\varphi)$. However, it tells us that equation (16.72) implies by itself this independence without any further assumption on the explicit form of Σ.

This result has important implications in quantized gauge theories.

16.6 Application: Stochastic Field Equations

We now consider a stochastic field equation which can be formally written as

$$E_\alpha(\varphi(x)) = \nu_\alpha(x), \tag{16.78}$$

in which α corresponds to the various components of the field $\varphi^\alpha(x)$ and $\nu_\alpha(x)$ is a field, hereafter called the noise, for which a probability distribution is provided. In most examples the noise $\nu_\alpha(x)$ will have a gaussian distribution.

Moreover, we will be interested in local stochastic field equations and local noise correlations, and find out in the example of the Langevin equation that the integration over the noise leads to large momentum divergences. It is, therefore, necessary to understand their properties from the point of view of renormalization and renormalization group. It is then convenient to associate with these equations a functional integral and a local action because renormalization of actions is much better understood than renormalization of equations. Moreover, in such a representation explicit integration over the noise becomes possible. This naturally leads to the geometric structure we have studied up to now in this chapter, and in particular to actions with BRS symmetries.

Starting from an equation as general as equation (16.78), it is possible to derive results which apply to several problems which have the same formal structure; for example, $E_\alpha(\varphi)$ can correspond to a classical field equation (of the form (16.34)) for a spin φ and ν_α to a random magnetic field (Section A17.2), or to a Langevin equation, as will be discussed in Section 16.7, and x then has to be understood as the collection of space and time variables. In addition, as we will show in Chapters 19,21, a similar structure emerges in the quantization of gauge theories.

16.6.1 The associated action

We write the noise probability distribution $[\mathrm{d}\rho(\nu)]$ as

$$[\mathrm{d}\rho(\nu)] = [\mathrm{d}\nu]\exp\left[-\sigma(\nu)\right]. \tag{16.79}$$

The most useful example corresponds to $\sigma(\nu)$ quadratic in $\nu_\alpha(x)$:

$$\sigma(\nu) = \frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\,\nu_\alpha(x)[\Omega^{-1}]^{\alpha\beta}(x,y)\nu_\beta(y), \tag{16.80}$$

in which the $\Omega_{\alpha\beta}(x,y)$ is a local symmetric operator.

The generating functional $\mathcal{Z}(J)$ of φ-field correlation functions is a noise expectation value of the form

$$\mathcal{Z}(J) = \left\langle \exp\left[\int \mathrm{d}x\,J_\alpha(x)\varphi^\alpha(x)\right]\right\rangle_\nu . \tag{16.81}$$

According to the discussion of Section 16.5 we can write $\mathcal{Z}(J)$ as (equation (17.36))

$$\mathcal{Z}(J) = \int [\mathrm{d}\rho(\nu)][\mathrm{d}\varphi\,\mathrm{d}\bar\varphi\,\mathrm{d}\bar c\,\mathrm{d}c]\exp\left[-\mathcal{S}(\varphi,\bar\varphi,c,\bar c,\nu) + \int \mathrm{d}x\,J_\alpha(x)\varphi^\alpha(x)\right], \tag{16.82}$$

where $\mathcal{S}(\varphi,c,c,\bar\varphi,\nu)$ is the action:

$$\mathcal{S}(\varphi,\bar\varphi,c,\bar c,\nu) = \int \mathrm{d}x\,\bar\varphi^\alpha(x)\left[E_\alpha(\varphi)-\nu(x)\right] - \int \mathrm{d}x\,\mathrm{d}y\,c^\alpha(x)E_{\alpha\beta}(x,y)\bar c^\beta(y), \tag{16.83}$$

and

$$E_{\alpha\beta}(x,y)=\frac{\delta E_\alpha(\varphi(x))}{\delta\varphi^\beta(y)}. \tag{16.84}$$

We recall that the identity (16.82) is true only if equation (16.78) has a unique solution. The verification of this condition in field theory may not be easy when the field equations are not linear. The Langevin equation we consider in Section 16.7 is a first order differential equation in time and thus this condition is indeed satisfied.

We now integrate over the noise, setting

$$\exp\left[w(\bar\varphi)\right]=\int[\mathrm{d}\rho(\nu)]\exp\left[\int\mathrm{d}x\,\bar\varphi^\alpha(x)\nu_\alpha(x)\right], \tag{16.85}$$

where $w(\bar\varphi)$ is the generating functional of connected ν-field correlation functions. In the gaussian example (16.80)

$$w(\bar\varphi)=\frac{1}{2}\int\mathrm{d}x\,\mathrm{d}y\,\bar\varphi(x)\Omega(x,y)\bar\varphi(y). \tag{16.86}$$

This yields the final expression

$$\mathcal{Z}(J)=\int[\mathrm{d}\varphi\,\mathrm{d}\bar\varphi\,\mathrm{d}\bar c\,\mathrm{d}c]\exp\left[-\mathcal{S}(\varphi,\bar\varphi,c,\bar c)+\int\mathrm{d}x\,J_\alpha(x)\varphi^\alpha(x)\right] \tag{16.87}$$

with

$$\mathcal{S}(\varphi,\bar\varphi,c,\bar c)=-w(\bar\varphi)+\int\mathrm{d}x\,\bar\varphi^\alpha(x)E_\alpha\bigl(\varphi(x)\bigr)-\int\mathrm{d}x\,\mathrm{d}y\,c^\alpha(x)E_{\alpha\beta}(x,y)\bar c^\beta(y). \tag{16.88}$$

This expression is, in more explicit notations adapted to field theory, the expression (16.59a). It, therefore, has a simple BRS symmetry. The transformations of the fields read

$$\delta\varphi^\alpha(x)=\varepsilon\bar c^\alpha(x),\quad \delta\bar c^\alpha(x)=0\,,\qquad \delta c^\alpha(x)=\varepsilon\bar\varphi^\alpha(x),\quad \delta\bar\varphi^\alpha(x)=0\,, \tag{16.89}$$

where ε is a Grassmann constant.

The action (16.88) can also be expressed in terms of the fields:

$$\phi^\alpha(x,\theta)=\varphi^\alpha(x)+\theta\bar c^\alpha(x),\qquad C^\alpha(x,\theta)=c^\alpha(x)+\theta\bar\varphi^\alpha(x),$$

as has been done in Section 16.3, to render the BRS obvious. Moreover, in some cases it is useful to combine all fields in one superfield:

$$\phi^\alpha(x,\theta,\bar\theta)=\phi^\alpha(x,\theta)+C^\alpha(x,\theta)\bar\theta=\varphi^\alpha(x)+\theta\bar c^\alpha(x)+c^\alpha(x)\bar\theta+\theta\bar\theta\bar\varphi^\alpha(x)\,. \tag{16.90}$$

In the case of the gaussian noise (16.80) the action $\mathcal{S}(\phi)$ then reads

$$\mathcal{S}(\phi)=\int\mathrm{d}\bar\theta\,\mathrm{d}\theta\left\{\int\mathrm{d}x\,\theta\frac{\partial\phi^\alpha}{\partial\bar\theta}E_\alpha\left[\phi(x,\theta,\bar\theta)\right]+\int\mathrm{d}x\,\mathrm{d}y\,\frac{\partial\phi^\alpha}{\partial\bar\theta}\Omega_{\alpha\beta}\frac{\partial\phi^\beta}{\partial\theta}\right\}. \tag{16.91}$$

16.6.2 BRS symmetry and WT identities

To the BRS symmetry corresponds a set of WT identities which can be derived by introducing sources for all fields,

$$\mathcal{Z}(J,\bar{J},\eta,\bar{\eta}) = \int [\mathrm{d}\varphi][\mathrm{d}\bar{\varphi}][\mathrm{d}\bar{c}][\mathrm{d}c] \exp\Big[-\mathcal{S}(\varphi,\bar{\varphi},c,\bar{c}) \\ + \int \mathrm{d}x (J_\alpha\varphi^\alpha + \bar{J}_\alpha\bar{\varphi}^\alpha + \eta_\alpha\bar{c}^\alpha + c^\alpha\bar{\eta}_\alpha)\Big], \tag{16.92}$$

and performing in the functional integral a change of variables of the form (16.89). Since the transformations are linear in the fields, the consequences are simple, the generating functional of connected correlation functions satisfies

$$\int \mathrm{d}x \left(J_\alpha(x)\frac{\delta}{\delta\eta_\alpha(x)} - \bar{\eta}_\alpha(x)\frac{\delta}{\delta\bar{J}_\alpha(x)}\right) W(J,\bar{J},\eta,\bar{\eta}) = 0\,, \tag{16.93}$$

and the 1PI generating functional $\Gamma(\varphi,\bar{\varphi},c,\bar{c})$ is BRS symmetric, that is, is also invariant under the transformation (16.89). This property is expressed by the equation

$$\mathcal{D}\,\Gamma(\varphi,\bar{\varphi},c,\bar{c}) \equiv \int \mathrm{d}x \left[\bar{c}^\alpha(x)\frac{\delta\Gamma}{\delta\varphi^\alpha(x)} + \bar{\varphi}^\alpha(x)\frac{\delta\Gamma}{\delta c^\alpha(x)}\right] = 0\,. \tag{16.94}$$

If the action $\mathcal{S}(\varphi,\bar{\varphi},c,\bar{c})$ is local and renormalizable by power counting, and if a BRS invariant regularization can be found, then equation (16.94) implies that the counter-terms needed to render the theory finite are BRS symmetric. Equation (16.94) is stable under renormalization and the renormalized action $\mathcal{S}_\mathrm{r}$ is BRS symmetric, that is, also satisfies

$$\mathcal{D}\mathcal{S}_\mathrm{r}(\varphi,\bar{\varphi},c,\bar{c}) = 0\,. \tag{16.95}$$

Renormalized action. We assume that as a consequence of power counting arguments and fermion number conservation $\mathcal{S}_\mathrm{r}$ has the general form

$$\mathcal{S}_\mathrm{r} = -\int \mathrm{d}x\,\mathrm{d}y\, E^\mathrm{r}_{\alpha\beta}(\varphi)c^\alpha\bar{c}^\beta - \tfrac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\,\bar{\varphi}(x)\bar{\varphi}(y)\Omega^\mathrm{r}(x,y) + \int \mathrm{d}x\,\bar{\varphi}^\alpha(x)E^\mathrm{r}_\alpha(\varphi) + E^\mathrm{r}(\varphi). \tag{16.96}$$

It then follows from the analysis of Section 16.5.2 that $E^\mathrm{r}(\varphi)$ is an irrelevant constant and

$$E^\mathrm{r}_{\alpha\beta} = \frac{\partial E^\mathrm{r}_\alpha(x)}{\partial\varphi^\beta(y)}. \tag{16.97}$$

The form (16.88) is thus preserved by the renormalization. In the parametrization (16.90) these properties have a simple interpretation, they correspond to invariance under the translation of the coordinate θ.

16.7 Langevin and Fokker–Planck Equations

An important example of stochastic field equations is provided by Langevin equations governing field time evolutions. To keep the notation as simple as possible we discuss the properties of the Langevin equation in the example of a one-component scalar field $\varphi(x,t)$, where x is the position in the d-dimensional euclidean space and t the time. The Langevin equation is a first order in time stochastic differential equation, straightforward generalization of the equation introduced in Chapter 4. We consider here equations of the generic form

$$\dot{\varphi}(x,t) = -\tfrac{1}{2}\Omega L\left[\varphi(x,t)\right] + \nu(x,t), \tag{16.98}$$

where $L\left[\varphi(x)\right]$ is a local functional of $\varphi(x)$ and Ω^{-1} a constant introduced for convenience, to provide a time scale. The noise field $\nu(x,t)$ is a stochastic field which we assume to have a gaussian local distribution $[\mathrm{d}\rho(\nu)]$:

$$[\mathrm{d}\rho(\nu)] = [\mathrm{d}\nu] \exp\left[-\int \mathrm{d}^d x \, \mathrm{d}t \, \nu^2(x,t)/2\Omega\right], \tag{16.99}$$

which can also be characterized by its one- and two-point functions,

$$\langle \nu(x,t)\rangle = 0\,, \qquad \langle \nu(x,t)\nu(x',t')\rangle = \Omega\,\delta(x-x')\delta(t-t'). \tag{16.100}$$

We discuss the problem of renormalization on this simple example though many results apply to more general equations and gaussian distributions of noise (see for, example, Sections 17.3,17.4).

The Fokker–Planck equation. Given the noise (16.99), and some initial conditions for the field $\varphi(x,t)$, the Langevin equation (16.98) generates a time-dependent field distribution $P(t,\varphi(x))$:

$$P\big(t,\varphi(x)\big) = \left\langle \delta\left(\varphi(x,t) - \varphi(x)\right)\right\rangle_\nu . \tag{16.101}$$

The derivation of Section 4.3 can be immediately generalized and yields a Fokker–Planck equation, the field theory form of equation (4.23):

$$\dot{P}(\varphi,t) = -\Omega \mathcal{H}_{\mathrm{FP}} P(\varphi,t)\,, \tag{16.102}$$

where $\mathcal{H}_{\mathrm{FP}}$, the Fokker–Planck hamiltonian, is given by

$$\mathcal{H}_{\mathrm{FP}}\left(\varphi,\frac{\delta}{\delta\varphi}\right) = -\frac{1}{2}\int \mathrm{d}^d x \frac{\delta}{\delta\varphi(x)}\left[\frac{\delta}{\delta\varphi(x)} + L\left(\varphi(x)\right)\right]. \tag{16.103}$$

The dissipative Langevin equation. In Section 17.1 a special case is discussed more thoroughly—the purely dissipative Langevin equation which corresponds to the choice

$$L\left[\varphi(x)\right] = \frac{\delta\mathcal{A}}{\delta\varphi(x)}, \tag{16.104}$$

and generalizes the example studied in Section 4.5. The functional $\mathcal{A}(\varphi)$ is a static (time-independent) euclidean action for the scalar field φ, for example,

$$\mathcal{A}(\varphi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\left(\partial_\mu\varphi(x)\right)^2 + V\left(\varphi(x)\right)\right].$$

In Chapter 4 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 theories. Here, we discuss new problems which arise because, as we show, the Langevin equation in general requires renormalizations. In the case (16.104) 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 Langevin equation finite. To discuss this problem, we have to set up a formalism more directly amenable to the ordinary methods of quantum field theory. This can be done by constructing a functional integral representation of the time-dependent φ-field correlation functions in terms of an associated local action, which, in this framework, it is natural to call dynamic action.

16.8 Time-Dependent Correlation Functions and Equilibrium

We now construct a functional integral representation for the generating functional $\mathcal{Z}(J)$ of dynamic correlation functions of the field $\varphi(x,t)$ solution of equation (16.98), and then discuss in a special example the relation between ST symmetry and equilibrium properties.

16.8.1 Dynamic action

The generating functional $\mathcal{Z}(J)$ of dynamic correlation functions of the field $\varphi(x,t)$ solution of equation (16.98), is given by the noise expectation value:

$$\begin{aligned}\mathcal{Z}(J) &= \left\langle \exp\left[\int \mathrm{d}^d x\, \mathrm{d}t\, J(x,t)\varphi(x,t)\right]\right\rangle_{\nu}, \\ &= \int [\mathrm{d}\nu] \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}t \left(\frac{1}{2\Omega}\nu^2(x,t) - J(x,t)\varphi(x,t)\right)\right]. \end{aligned} \tag{16.105}$$

To impose equation (16.98), following the method explained in Section 16.6, we insert the identity

$$\int [\mathrm{d}\varphi] \det M \delta\left[\dot{\varphi} + (\Omega/2)L(\varphi) - \nu\right] = 1\,, \tag{16.106}$$

where M is the differential operator,

$$M = \frac{\partial}{\partial t} + \frac{\Omega}{2}\frac{\delta L}{\delta\varphi(x,t)}, \tag{16.107}$$

into expression (16.105). We then find

$$\mathcal{Z}(J) = \int [\mathrm{d}\nu]\,[\mathrm{d}\varphi] \det M\, \delta\left[\dot{\varphi} + \tfrac{1}{2}\Omega L(\varphi) - \nu\right] \exp\left[\int \mathrm{d}^d x\, \mathrm{d}t (J\varphi - \nu^2/2\Omega)\right]. \tag{16.108}$$

The δ-function can be used to integrate over the noise ν:

$$\mathcal{Z}(J) = \int [\mathrm{d}\varphi] \det M \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}t \left(\frac{1}{2\Omega}\left(\dot{\varphi} + \tfrac{1}{2}\Omega L(\varphi)\right)^2 - J\varphi\right)\right]. \tag{16.109}$$

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

$$\det M \propto \exp \operatorname{tr} \ln \left[1 + \left(\frac{\partial}{\partial t}\right)^{-1} \frac{\Omega}{2}\frac{\delta L}{\delta\varphi}\right]. \tag{16.110}$$

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 Ω all terms thus vanish when one takes the trace, except the first one which yields

$$\det M \propto \exp\left\{\theta(0)\frac{\Omega}{2}\int \mathrm{d}t\,\mathrm{d}^d x\,\frac{\delta L[\varphi(x,t)]}{\delta\varphi(x',t)}\bigg|_{x'=x}\right\}. \tag{16.111}$$

For the ill-defined quantity $\theta(0)$ we choose $\theta(0)=1/2$, a choice symmetric in time, for reasons we have already explained in Section 4.6. The final expression then formally reads

$$\mathcal{Z}(J)=\int[\mathrm{d}\varphi]\exp\left[-\mathcal{S}(\varphi)+\int \mathrm{d}^d x\,\mathrm{d}t\,J(x,t)\varphi(x,t)\right], \tag{16.112}$$

$$\mathcal{S}(\varphi)=\frac{1}{2\Omega}\int \mathrm{d}^d x\,\mathrm{d}t\,\left(\dot\varphi+\tfrac{1}{2}\Omega L(\varphi)\right)^2-\frac{\Omega}{4}\int \mathrm{d}t\,\mathrm{d}^d x\,\frac{\delta L[\varphi(x,t)]}{\delta\varphi(x',t)}\bigg|_{x'=x}. \tag{16.113}$$

When the force in the Langevin equation derives from a static action (equation (16.104)), the term linear in $\dot\varphi$ in the action $\mathcal{S}$ is a total time derivative and contributes only to boundary terms.

Note that the dynamic action (16.113) is the generalization to field theory of the action (4.49), and can thus be directly obtained by writing the solution $P(\varphi,t)$ of the Fokker–Planck equation (16.102) as a functional integral.

16.8.2 ST symmetry and equilibrium correlation functions

Langevin equations belong to the general class of stochastic equations discussed in Section 16.6. We exploit this remark more systematically in the next section. However, we already note that the ST symmetry discussed in Section 16.1 implies identities for correlation functions.

We consider, here, only the example (16.104). Then the substitution in equation (16.102),

$$P=\mathrm{e}^{-\mathcal{A}/2}\,\widetilde{P},$$

transforms the hamiltonian (16.103) into a hermitian positive hamiltonian. This property allows to prove the convergence of P at time $+\infty$ towards the static distribution $\mathrm{e}^{-\mathcal{A}}$ (provided the latter is normalizable). It is at the basis of the idea of stochastic quantization, the Langevin equation being there only a device to generate the static distribution $\mathrm{e}^{-\mathcal{A}}$, as in the example of Section 18.13.

Another proof relies on ST symmetry. It follows from the identity (16.106) that the determinant of $\mathbf{M}$,

$$M(x,t;x',t')=\langle x,t|\left(\frac{\partial}{\partial t}+\frac{\Omega}{2}\frac{\delta^2\mathcal{A}}{\delta\varphi\delta\varphi}\right)|x',t'\rangle, \tag{16.114}$$

generates an invariant measure for a set of non-linear transformations which translate $\nu(x,t)$ by a function $\mu(x,t)$ independent of φ. For $\mu(x,t)$ infinitesimal the variation of φ is

$$\delta\varphi(x,t)=\int \mathrm{d}t'\,\mathrm{d}x'\,M^{-1}(x,t;x',t';\varphi)\mu(x',t'). \tag{16.115}$$

In an infinitesimal change of variables of the form of a transformation (16.115) the variations of the action and the source term in the functional integral (16.109) are

$$\delta\left\{\frac{1}{2\Omega}\int \mathrm{d}x\,\mathrm{d}t\left(\dot\varphi+\frac{\Omega}{2}\frac{\delta\mathcal{A}}{\delta\varphi}\right)^2\right\}=\frac{1}{\Omega}\int \mathrm{d}x\,\mathrm{d}t\left[\dot\varphi(x,t)+\frac{\Omega}{2}\frac{\delta\mathcal{A}}{\delta\varphi(x,t)}\right]\mu(x,t),$$

$$\delta\left[\int \mathrm{d}x\,\mathrm{d}t\,J\varphi\right]=\int \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}x'\,\mathrm{d}t'\,J(x,t)M^{-1}(x',t';x,t)\mu(x',t').$$

Expressing as usual the invariance of the functional integral under a change of variables, we obtain the identity

$$\left[\frac{1}{\Omega}\frac{\partial}{\partial t}\frac{\delta}{\delta J(x,t)}+\frac{1}{2}\frac{\delta\mathcal{A}}{\delta\varphi(x,t)}\left(\frac{\delta}{\delta J}\right)\right]\mathcal{Z}(J)=\int \mathrm{d}x'\,\mathrm{d}t'\,J(x',t')\mathcal{Z}_M(x',t';x,t) \qquad (16.116)$$

with the definition

$$\mathcal{Z}_M(x,t;x',t';J)=M^{-1}(x',t';x,t;\delta/\delta J)\mathcal{Z}(J), \qquad (16.117)$$

or, equivalently,

$$\left\{\frac{\partial}{\partial t}+\frac{\Omega}{2}\frac{\delta^2\mathcal{A}}{[\delta\varphi(x,t)]^2}\left(\frac{\delta}{\delta J}\right)\right\}\mathcal{Z}_M(x,t;x',t';J)=\delta(x-x')\delta(t-t')\mathcal{Z}(J). \qquad (16.118)$$

Equation (16.116), 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 take a special source

$$J(x,t')=J(x)\delta(t'-\tau), \qquad (16.119)$$

and set $t=\tau$ in equation (16.116). Then the equation involves only $\mathcal{Z}_M(x',\tau;x,\tau;J)$. If we expand $\mathcal{Z}_M$ in powers of Ω, again due to the $\theta(t)$ function, only the first term survives and, with $\theta(0)=\frac{1}{2}$, $\mathcal{Z}_M$ becomes

$$\mathcal{Z}_M(x',\tau;x,\tau;J)=\tfrac{1}{2}\,\delta(x-x')\mathcal{Z}(J). \qquad (16.120)$$

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 (16.104) the term

$$\left.\frac{\partial}{\partial t}\frac{\delta}{\delta J(x,t)}\mathcal{Z}(J)\right|_{t=\tau}$$

vanishes at large time, so that the limiting functional $\mathcal{Z}(J)$ satisfies as a consequence of equations (16.116) and (16.120)

$$\frac{\delta\mathcal{A}}{\delta\varphi(x)}\left(\frac{\delta}{\delta J}\right)\mathcal{Z}(J)=J(x)\mathcal{Z}(J), \qquad (16.121)$$

which implies

$$\mathcal{Z}(J)=\int[\mathrm{d}\varphi(x)]\exp\left[-\mathcal{A}(\varphi)+\int \mathrm{d}x\,J(x)\varphi(x)\right]. \qquad (16.122)$$

From these considerations we conclude that some form of equations (16.116) and (16.117) 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 the equal-time correlation functions have in the large time limit, the correlation functions of the renormalized static theory corresponding to $\mathcal{A}(\varphi)$.

16.9 Renormalization and BRS Symmetry

Divergences and the problem of renormalization. In dimension $d > 0$, the expression (16.113) is ill-defined when $L(\varphi)$ is a local functional because the contribution of the determinant is formally proportional to $\delta^d(0)$:

$$\ln\det M \propto \int \mathrm{d}^d x \left.\frac{\delta L[\varphi(x,t)]}{\delta\varphi(x',t)}\right|_{x'=x} \propto \delta^d(0).$$

We have already encountered a similar situation in the discussion of homogeneous spaces in Chapters 14, 15 where a divergent measure term appeared too. The determinant has thus to be regularized. Again we have two choices:

(i) With dimensional regularization, terms like $\delta^d(0)$ vanish and, therefore, the determinant can be completely omitted. With this convention the expression (16.112) can be used in practical perturbative calculations.

(ii) It is, however, interesting to keep this divergent term in some regularized form, as we shall do in the discussion of the renormalization of the dynamic theory, in order to understand its geometric meaning. This can be achieved with lattice regularization.

Let us show, however, that even with dimensional regularization a non-trivial renormalization problem arises. We take the example of a dissipative Langevin equation (equation (16.104)) with the four-dimensional action $\mathcal{A}(\varphi)$:

$$\mathcal{A}(\varphi) = \int \mathrm{d}^4 x \left[\tfrac{1}{2}\left(\partial_\mu\varphi\right)^2 + \tfrac{1}{2}m^2\varphi^2 + (g/4!)\varphi^4\right]. \tag{16.123}$$

From the dynamic action (16.113) we calculate the propagator Δ of the $\varphi(x,t)$ field. After Fourier transformation, as a function of $\mathbf{k}$ and ω the variables associated with space and time respectively, the propagator reads

$$\Delta(\mathbf{k},\omega) = \frac{\Omega}{\omega^2 + \Omega^2(k^2+m^2)^2/4}. \tag{16.124}$$

In the standard power counting analysis, as presented in Chapter 9, it is generally assumed that the propagator is, in momentum representation and when all arguments become large, a homogeneous function. This is not the case here. However, the propagator has a homogeneous asymptotic behaviour if we scale the frequency ω 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:

$$[\varphi] = 1\,, \tag{16.125}$$

and the interactions coming from $(\delta\mathcal{A}/\delta\varphi)^2$ have dimensions 4 and 6:

$$\frac{\delta\mathcal{A}}{\delta\varphi} = -\nabla^2\varphi + m^2\varphi + g\varphi^3/6\,. \tag{16.126}$$

General renormalization theory tells us only that the renormalized action is the most general local functional of canonical dimension 6, that is, containing all vertices of non-positive dimensions. Such a functional in general depends on more parameters than the bare action (16.113); 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 (16.105–16.113). To prove that the Langevin equation can be renormalized, we have, therefore, to find identities satisfied by correlation functions, which imply relations between counter-terms and ensure that the structure of the action (16.113) is preserved by renormalization.

We now consider a particular example of a Langevin equation (16.98):

$$\dot{\varphi}(x,t) + \tfrac{1}{2}\Omega\left[-\nabla^2\varphi + L_{\text{int.}}(\varphi)\right] = \nu(x,t), \tag{16.127}$$

in which $L_{\text{int.}}(\varphi)$ is a polynomial in φ, and $\nu(x,t)$ the gaussian noise with distribution (16.99). 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. The dynamic action corresponding to equation (16.127) is then,

$$\mathcal{S} = \int \mathrm{d}t\,\mathrm{d}^d x \left\{-\tfrac{1}{2}\Omega\bar{\varphi}^2 + \bar{\varphi}\left[\dot{\varphi} + \tfrac{1}{2}\Omega\left(-\nabla^2\varphi + L_{\text{int.}}(\varphi)\right)\right] - cM\bar{c}\right\}, \tag{16.128}$$

where M is the operator (16.107),

$$M = \frac{\partial}{\partial t} - \tfrac{1}{2}\Omega\nabla^2 + \tfrac{1}{2}\Omega\frac{\partial L_{\text{int.}}[\varphi(x,t)]}{\partial\varphi}. \tag{16.129}$$

In this form the associated dynamic action (16.128) is BRS symmetric. 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 Fokker–Planck equation. However, to exhibit more clearly the general algebraic structure, it is useful to keep it in the form of a fermion integral.

Power counting and renormalization. Note that, below, the canonical dimension of a quantity Q is denoted by $[Q]$. As discussed in Section 16.8.1, from the point of view of power counting, 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 (16.128) then has dimension $d+2$ since the integration measure here is $\mathrm{d}t\,\mathrm{d}^d x$. The dimensions $[\bar{\varphi}]$, $[\varphi]$, $[\bar{c}]$ and $[c]$ of the fields follow:

$$[\bar{\varphi}] = \tfrac{1}{2}(d+2), \quad [\varphi] = \tfrac{1}{2}(d-2), \quad [\bar{c}] + [c] = d\,. \tag{16.130}$$

The theory is thus renormalizable by power counting in dimension d if the dimension of $L_{\text{int.}}(\varphi)$ is

$$[L_{\text{int.}}(\varphi)] = \tfrac{1}{2}(d+2),$$

The dimension of M is then $[M] = [L_{\text{int.}}(\varphi)] - \tfrac{1}{2}(d-2) = 2$. 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 $\mathcal{S}_{\mathrm{r}}$ 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

$$[\bar{c}] + [c] + [\bar{\varphi}] = \tfrac{3}{2}d + 1 > d + 2\,,$$

and, therefore, M_{r}, the renormalized coefficient of cc, cannot depend on $\bar{\varphi}$.

longer be derived from a Langevin equation by the algebraic transformations (16.105–16.113). To prove that the Langevin equation can be renormalized, we have, therefore, to find identities satisfied by correlation functions, which imply relations between counter-terms and ensure that the structure of the action (16.113) is preserved by renormalization.

We now consider a particular example of a Langevin equation (16.98):

$$\dot{\varphi}(x,t)+\tfrac{1}{2}\Omega\left[-\nabla^2\varphi+L_{\text{int.}}(\varphi)\right]=\nu(x,t), \tag{16.127}$$

in which $L_{\text{int.}}(\varphi)$ is a polynomial in φ, and $\nu(x,t)$ the gaussian noise with distribution (16.99). 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. The dynamic action corresponding to equation (16.127) is then,

$$\mathcal{S}=\int \mathrm{d}t\,\mathrm{d}^d x\left\{-\tfrac{1}{2}\Omega\bar{\varphi}^2+\bar{\varphi}\left[\dot{\varphi}+\tfrac{1}{2}\Omega\left(-\nabla^2\varphi+L_{\text{int.}}(\varphi)\right)\right]-cM\bar{c}\right\}, \tag{16.128}$$

where M is the operator (16.107),

$$M=\frac{\partial}{\partial t}-\tfrac{1}{2}\Omega\nabla^2+\tfrac{1}{2}\Omega\frac{\partial L_{\text{int.}}[\varphi(x,t)]}{\partial\varphi}. \tag{16.129}$$

In this form the associated dynamic action (16.128) is BRS symmetric. 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 Fokker–Planck equation. However, to exhibit more clearly the general algebraic structure, it is useful to keep it in the form of a fermion integral.

Power counting and renormalization. Note that, below, the canonical dimension of a quantity Q is denoted by $[Q]$. As discussed in Section 16.8.1, from the point of view of power counting, 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 (16.128) then has dimension $d+2$ since the integration measure here is $\mathrm{d}t\,\mathrm{d}^d x$. The dimensions $[\bar{\varphi}]$, $[\varphi]$, $[\bar{c}]$ and $[c]$ of the fields follow:

$$[\bar{\varphi}]=\tfrac{1}{2}(d+2),\quad [\varphi]=\tfrac{1}{2}(d-2),\quad [\bar{c}]+[c]=d\,. \tag{16.130}$$

The theory is thus renormalizable by power counting in dimension d if the dimension of $L_{\text{int.}}(\varphi)$ is

$$[L_{\text{int.}}(\varphi)]=\tfrac{1}{2}(d+2),$$

The dimension of M is then $[M]=[L_{\text{int.}}(\varphi)]-\frac{1}{2}(d-2)=2$. 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 $\mathcal{S}_{\mathrm{r}}$ 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

$$[\bar{c}]+[c]+[\bar{\varphi}]=\tfrac{3}{2}d+1>d+2\,,$$

and, therefore, M_{r}, the renormalized coefficient of cc, cannot depend on $\bar{\varphi}$.

The results of Section 16.6 (which follow from BRS symmetry) then apply and the structure of the dynamic action (16.128) is 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 $L_{\text{int.}}(\varphi)$ satisfies $[L_{\text{int.}}] \le \frac{1}{2}(d+2)$ and the dimension d is larger than 2, the Langevin equation (16.127) with the gaussian noise (16.99) can be renormalized. Moreover, since the renormalized action remains quadratic in $\bar{\varphi}$, the renormalized noise is still gaussian. Note that 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 particular case (16.104) more precise results can be obtained which we describe in Section 17.1 (for a one-component field equation the drift force in equation (16.127) can of course always be written as the variation of an action as in (16.104)).

Bibliographical Notes

The ST has been introduced in the context of gauge theories by

A.A. Slavnov, *Theor. Math. Phys.* 10 (1972) 99; J.C. Taylor, *Nucl. Phys.* B33 (1971) 436;

as well as the BRS symmetry,

C. Becchi, A. Rouet and R. Stora, *Commun. Math. Phys.* 42 (1975) 127; *Ann. Phys.* 98 (1976) 28.

One generally quotes also

I.V. Tyutin, Lebedev preprint FIAN 39 (1975), unpublished.

The discussion of Section 16.1 is inspired by

B.W. Lee and J. Zinn-Justin, *Phys. Rev.* D7 (1973) 1049 (Appendix).

The discussion of the BRS symmetry in this different context and its application to the Langevin equation follows

J. Zinn-Justin, *Nucl. Phys.* B275 [FS18] (1986) 135.

17 FROM LANGEVIN EQUATION TO SUPERSYMMETRY

In Chapter 16 we have already explained how to associate to the Langevin or Fokker–Planck equations a dynamic action (see in particular Sections 16.6 and 16.7). We have shown that the corresponding formalism is well suited to the discussion of a number of properties of these stochastic dynamic equations, like renormalization or equilibrium distributions.

We have discovered quite generally that the dynamic action has a BRS 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 is thus conservative (the corresponding stochastic processes satisfy the detailed balance condition). We show that the dynamic action then has a second Grassmann symmetry which, combined with the first one, provides the simplest example of supersymmetry: quantum mechanics supersymmetry. We discuss a few consequences of the supersymmetry like WT identities, renormalization properties and equilibrium distribution. We describe some subtleties of the formalism in the case of two-dimensional models defined on various manifolds.

In the two last sections, we present, for completeness, a very brief discussion of supersymmetric quantum field theories, since supersymmetry has been proposed as a principle to solve the so-called *hierarchy 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 20.1.4). Exact supersymmetry would also solve the problem of vanishing or at least unnatural small value of the cosmological constant (see Section 22.7.2). However, experimentally the, as yet undiscovered, supersymmetric partners of known particles have to be much heavier. Supersymmetry is necessarily broken and the problem of the cosmological constant only slightly improved. Finally, gauging of supersymmetry naturally leads to unification with gravity, because the commutators of supersymmetry currents involve the energy–momentum tensor.

In Section 17.5, we construct simple examples of supersymmetric theories for scalar superfields. The new technical feature is the combination of spin and supersymmetry. In Section 17.6 we briefly describe supersymmetry in four dimensions, a structure potentially relevant for four-dimensional particle physics.

The appendix contains some general remarks on supersymmetry and a short discussion of the random field Ising model, and the consequences of its perturbative supersymmetry.

17 FROM LANGEVIN EQUATION TO SUPERSYMMETRY

In Chapter 16 we have already explained how to associate to the Langevin or Fokker–Planck equations a dynamic action (see in particular Sections 16.6 and 16.7). We have shown that the corresponding formalism is well suited to the discussion of a number of properties of these stochastic dynamic equations, like renormalization or equilibrium distributions.

We have discovered quite generally that the dynamic action has a BRS 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 is thus conservative (the corresponding stochastic processes satisfy the detailed balance condition). We show that the dynamic action then has a second Grassmann symmetry which, combined with the first one, provides the simplest example of supersymmetry: quantum mechanics supersymmetry. We discuss a few consequences of the supersymmetry like WT identities, renormalization properties and equilibrium distribution. We describe some subtleties of the formalism in the case of two-dimensional models defined on various manifolds.

In the two last sections, we present, for completeness, a very brief discussion of supersymmetric quantum field theories, since supersymmetry has been proposed as a principle to solve the so-called *hierarchy 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 20.1.4). Exact supersymmetry would also solve the problem of vanishing or at least unnatural small value of the cosmological constant (see Section 22.7.2). However, experimentally the, as yet undiscovered, supersymmetric partners of known particles have to be much heavier. Supersymmetry is necessarily broken and the problem of the cosmological constant only slightly improved. Finally, gauging of supersymmetry naturally leads to unification with gravity, because the commutators of supersymmetry currents involve the energy–momentum tensor.

In Section 17.5, we construct simple examples of supersymmetric theories for scalar superfields. The new technical feature is the combination of spin and supersymmetry. In Section 17.6 we briefly describe supersymmetry in four dimensions, a structure potentially relevant for four-dimensional particle physics.

The appendix contains some general remarks on supersymmetry and a short discussion of the random field Ising model, and the consequences of its perturbative supersymmetry.

17.1 The Purely Dissipative Langevin Equation

We have shown in Section 16.3 that when a stochastic equation has, in symbolic notations, the form $\delta\mathcal{A}/\delta\varphi = \nu$, the associated action $\mathcal{S}(\phi)$ has two independent BRS symmetries which, in the superfield notation (16.90), correspond to translations of θ and $\bar{\theta}$. The Langevin equation in the special case (16.104):

$$\dot{\varphi}(x,t) = -\frac{\Omega}{2}\frac{\delta\mathcal{A}}{\delta\varphi(x,t)} + \nu(x,t), \tag{17.1}$$

almost shares this property. The force term is a gradient but not the time derivative $\dot{\varphi}$. Somewhat surprisingly, however, it is possible to find a slightly modified BRS transformation which leaves the dynamic action (16.128) invariant and which combines non-trivially with the first one to yield the simplest example of supersymmetry.

17.1.1 Supersymmetry

In the superfield notation (16.90),

$$\phi^{\alpha}(x,t;\bar{\theta},\theta) = \varphi^{\alpha}(x,t) + \theta\bar{c}^{\alpha}(x,t) + c^{\alpha}(x,t)\bar{\theta} + \theta\bar{\theta}\bar{\varphi}^{\alpha}(x,t)\,,$$

the action (16.128) corresponding to the Langevin equation (17.1) with the noise gaussian distribution (16.99)

$$[\mathrm{d}\rho(\nu)] = [\mathrm{d}\nu]\exp\left[-\int \mathrm{d}^d x\,\mathrm{d}t\,\nu^2(x,t)/2\Omega\right],$$

can be rewritten as

$$\mathcal{S}(\phi) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t\left[\frac{2}{\Omega}\int \mathrm{d}^d x\,\bar{\mathrm{D}}\phi\mathrm{D}\phi + \mathcal{A}(\phi)\right], \tag{17.2}$$

with the definitions:

$$\bar{\mathrm{D}} = \frac{\partial}{\partial\bar{\theta}}, \quad \mathrm{D} = \frac{\partial}{\partial\theta} - \bar{\theta}\frac{\partial}{\partial t}. \tag{17.3}$$

For convenience we have rescaled equation (17.1) by a factor $2/\Omega$.

$\bar{\mathrm{D}}$ and D satisfy the anticommutation relations:

$$\mathrm{D}^2 = \bar{\mathrm{D}}^2 = 0\,, \qquad \mathrm{D}\bar{\mathrm{D}} + \bar{\mathrm{D}}\mathrm{D} = -\frac{\partial}{\partial t}\,. \tag{17.4}$$

We then introduce two other Grassmann-type differential operators Q, $\bar{\mathrm{Q}}$:

$$\mathrm{Q} = \frac{\partial}{\partial\theta}, \qquad \bar{\mathrm{Q}} = \frac{\partial}{\partial\bar{\theta}} + \theta\frac{\partial}{\partial t}, \tag{17.5}$$

Both *anticommute* with D and $\bar{\mathrm{D}}$ and satisfy

$$\mathrm{Q}^2 = \bar{\mathrm{Q}}^2 = 0\,, \qquad \mathrm{Q}\bar{\mathrm{Q}} + \bar{\mathrm{Q}}\mathrm{Q} = \frac{\partial}{\partial t}\,. \tag{17.6}$$

The two pairs $\mathrm{D}, \bar{\mathrm{D}}$ and $\mathrm{Q}, \bar{\mathrm{Q}}$ provide the simplest example of generators of supersymmetry (see for example Section A17.1). We already know that Q (a translation of θ) is the generator of a BRS symmetry of the action (17.2). Let us verify that $\bar{\mathrm{Q}}$ is the generator of an additional symmetry. If we perform a variation of ϕ of the form

$$\delta\phi(t,\theta,\bar{\theta}) = \bar{\varepsilon}\bar{\mathrm{Q}}\phi\,, \tag{17.7}$$

which in component form reads

$$\begin{array}{ll} \delta\varphi = c\bar{\varepsilon}\,, & \delta c = 0\,, \\ \delta\bar{c} = (\bar{\varphi} - \dot{\varphi})\,\bar{\varepsilon}\,, & \delta\bar{\varphi} = \dot{c}\bar{\varepsilon}\,, \end{array} \tag{17.8}$$

we observe that the variation of the action density is a total derivative. This is obvious for $\mathcal{A}$ because it does not explicitly depend on t and $\bar{\theta}$. For the remaining term the additional property that $\bar{\mathrm{Q}}$ anticommutes with D and $\bar{\mathrm{D}}$ has to be used:

$$\delta\left[\bar{\mathrm{D}}\phi\mathrm{D}\phi\right] = \bar{\mathrm{D}}\left[\bar{\varepsilon}\bar{\mathrm{Q}}\phi\right]\mathrm{D}\phi + \bar{\mathrm{D}}\phi\mathrm{D}\left[\bar{\varepsilon}\bar{\mathrm{Q}}\phi\right] = \bar{\varepsilon}\bar{\mathrm{Q}}\left[\bar{\mathrm{D}}\phi\mathrm{D}\phi\right].$$

The action is thus supersymmetric. The operators D and $\bar{\mathrm{D}}$ are *covariant derivatives* from the point of view of supersymmetry.

This supersymmetry is directly related with the property that the corresponding Fokker–Planck hamiltonian (16.102) is then equivalent to a positive hamiltonian of the form (4.41).

Remarks.

(i) The anticommutator (17.6) of $\bar{\mathrm{Q}}$ and Q generates time translations.

(ii) It is possible to emphasize the symmetric role played by $\bar{\theta}$ and θ by performing the substitution $t \mapsto t + \frac{1}{2}\theta\bar{\theta}$. We find it more convenient to remain with the original variables but this is mainly a matter of taste.

(iii) Considering the fermions as real dynamic variables, we can write the hamiltonian associated with the supersymmetric action in boson–fermion space. We find that the functional integral describes both the Langevin equation and its time-reversed form.

17.1.2 Ward–Takahashi (WT) identities

The consequence of BRS symmetry is simple, correlation functions are invariant under a translation of the coordinate θ. The transformation (17.7) has a slightly more complicated form. It leads for the generating functional $\mathcal{W}(J)$ to the equation:

$$\int \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\bar{\mathrm{Q}}J(x,t,\theta,\bar{\theta})\frac{\delta\mathcal{W}}{\delta J(x,t,\theta,\bar{\theta})} = 0\,.$$

Connected correlation functions $W^{(n)}(x_i,t_i,\theta_i,\bar{\theta}_i)$ and proper vertices $\Gamma^{(n)}(p_i,\omega_i,\theta_i,\bar{\theta}_i)$ thus satisfy the WT identities:

$$\bar{\mathrm{Q}}\,W^{(n)}(x_i,t_i,\theta_i,\bar{\theta}_i) = 0\,, \qquad \bar{\mathrm{Q}}\,\Gamma^{(n)}(x_i,t_i,\theta_i,\bar{\theta}_i) = 0 \tag{17.9}$$

with

$$\bar{\mathrm{Q}} \equiv \sum_{k=1}^{n}\left(\frac{\partial}{\partial\bar{\theta}_k} + \theta_k\frac{\partial}{\partial t_k}\right).$$

After Fourier transformation over space and time the operator $\bar{\mathrm{Q}}$ takes the form

$$\bar{\mathrm{Q}} = \sum_{k=1}^{n}\left(\frac{\partial}{\partial\bar{\theta}_k} - i\omega_k\theta_k\right). \tag{17.10}$$

Example: two-point function. Let us explore the implications of WT identities for a two-point function. As the relations (17.5,17.6) show, supersymmetry implies invariance by translation in time and in θ. We can, therefore, write any two-point function $W^{(2)}$ as

$$W^{(2)} = A(t_1 - t_2) + (\theta_1 - \theta_2)\left[(\bar\theta_1 + \bar\theta_2)B(t_1 - t_2) + (\bar\theta_1 - \bar\theta_2)C(t_1 - t_2)\right]. \qquad (17.11)$$

The WT identity (17.9) then implies

$$2B(t) = \frac{\partial A}{\partial t}. \qquad (17.12)$$

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 $t_1 = t_2$. We find

$$2C(t) = -\epsilon(t)\frac{\partial A}{\partial t}, \qquad (17.13)$$

where $\epsilon(t)$ is the sign of t. This leads to the remarkable form:

$$W^{(2)} = \left\{1 + \tfrac{1}{2}(\theta_1 - \theta_2)\left[\bar\theta_1 + \bar\theta_2 - (\bar\theta_1 - \bar\theta_2)\epsilon(t_1 - t_2)\right]\frac{\partial}{\partial t_1}\right\} A(t_1 - t_2). \qquad (17.14)$$

17.1.3 Renormalization of the dissipative Langevin equation

In the special case of the supersymmetric dynamic action (17.2) a comparison between the two explicit quadratic terms in ϕ of the action yields the relation between dimensions:

$$[t] - [\theta] - [\bar\theta] = 0 \ \Rightarrow [\mathrm{d}t] + [\mathrm{d}\theta] + [\mathrm{d}\bar\theta] = 0\,. \qquad (17.15)$$

(We recall that since integration and differentiation over anticommuting variables are equivalent operations, the dimension of $\mathrm{d}\theta$ is $-[\theta]$.)

Therefore, the term proportional to $\mathcal{A}(\phi)$ in the action has the same canonical dimension as in the static case: 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 equation (17.15) also implies

$$2[\phi] = d + [t],$$

equation which relates the dimensions of field and time.

We then write in superfield notation the most general form of the renormalized action $\mathcal{S}_\mathrm{r}$ consistent with the results derived in Section 16.9:

$$\mathcal{S}_\mathrm{r} = \int \mathrm{d}^d x\,\mathrm{d}t\,\mathrm{d}\bar\theta\,\mathrm{d}\theta \left[\frac{2}{\Omega}\frac{\partial\phi}{\partial\bar\theta}\left(Z'\frac{\partial\phi}{\partial\theta} - Z\bar\theta\frac{\partial\phi}{\partial t}\right) + \bar\theta\frac{\partial\phi}{\partial\bar\theta}L(\phi)\right], \qquad (17.16)$$

and impose the constraints coming from the supersymmetry of the bare action. The transformation (17.7) is linearly represented on the fields ϕ and, therefore, the renormalized action remains symmetric. Performing the transformation (17.7) and expressing the invariance of the action we obtain two equations. First identifying the coefficient of $(\partial\phi/\partial\bar\theta)(\partial\phi/\partial t)$ we find

$$\mathcal{Z}' = \mathcal{Z}\,. \tag{17.17}$$

The second equation comes from the variation of the last term of expression (17.16):

$$\int \mathrm{d}\bar\theta \frac{\partial\phi}{\partial\bar\theta} L(\phi) = 0\,. \tag{17.18}$$

Thus $L(\phi)$ is a total derivative and can be written as

$$L(\phi) = \frac{\delta\,(\mathcal{A}_\mathrm{r})}{\delta\phi}. \tag{17.19}$$

An integration by parts over $\bar\theta$ of the last term in equation (17.16) finally yields the supersymmetric form of the renormalized action:

$$\mathcal{S}_\mathrm{r}(\phi) = \int \mathrm{d}\bar\theta\,\mathrm{d}\theta\,\mathrm{d}t \left[\frac{2Z_t}{\Omega}\int \mathrm{d}^d x\,\bar{\mathrm{D}}\phi\mathrm{D}\phi + \mathcal{A}_\mathrm{r}(\phi)\right]. \tag{17.20}$$

After renormalization the drift force in the Langevin equation is thus still of the form of the variation of an action.

Remark. Because the Fokker–Planck equation has static (time-independent) solutions which are not of the form (16.104), 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.

17.2 Supersymmetry and Equilibrium Correlation Functions

The supersymmetry of action (17.2) 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 $\mathcal{A}(\varphi)$. To simplify the notation we consider the action (17.2), but the generalization to actions of the type examined in Sections 17.3,17.4 is straightforward. We assume that the initial conditions in the Langevin equation are given at time t' and we calculate the equal-time correlation functions at time t''. The source J associated with the field ϕ then has the special form:

$$J(x,t,\theta,\bar\theta) = J(x)\delta(t-t'')\delta(\theta)\delta(\bar\theta). \tag{17.21}$$

We consider the s-dependent dynamic action:

$$\mathcal{S}(\phi,s) = (1-s)\int_{t'}^{t''} \mathrm{d}\bar\theta\,\mathrm{d}\theta\,\mathrm{d}t \left[\frac{2}{\Omega}\int \mathrm{d}^d x\,\bar{\mathrm{D}}\phi\mathrm{D}\phi + \mathcal{A}(\phi)\right] + s\mathcal{A}\left[\phi(t'',\theta=\bar\theta=0)\right]. \tag{17.22}$$

For $s=0$ we recover the dynamic action for time-dependent correlation functions relevant when all times are in the interval $[t',t'']$ and for $s=1$, because the source has the form (17.21), 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:

$$\frac{\partial}{\partial s}\mathcal{S}(\phi,s) = -\int_{t'}^{t''} \mathrm{d}\bar\theta\mathrm{d}\theta\mathrm{d}t \left[\frac{2}{\Omega}\int \mathrm{d}^d x\,\bar{\mathrm{D}}\phi\mathrm{D}\phi + \mathcal{A}(\phi)\right] + \mathcal{A}\left[\phi(t'',\theta=\bar\theta=0)\right]. \tag{17.23}$$

We first neglect the breaking of supersymmetry due to the boundary condition at t'. Our discussion is thus strictly correct only in the case the system already is at equilibrium, that is, $t' = -\infty$. In this case the insertion of the operator $\partial S/\partial s$ in an equal-time correlation function generates a two time correlation function. In Section 17.1.2 we have determined the most general two time connected correlation functions satisfying the requirements of supersymmetry and causality. Using $t'' > t$ and $\theta'' = \bar{\theta}'' = 0$, we can rewrite equation (17.14) as

$$\langle \mathcal{O}(t)\mathcal{O}(t'')\rangle = \left(1 + \theta\bar{\theta}\frac{\partial}{\partial t}\right) A(t - t'') = A(t + \theta\bar{\theta} - t''). \tag{17.24}$$

We can, therefore, replace $\phi(t, \theta, \bar{\theta})$ by $\varphi(t + \theta\bar{\theta})$ in the operator insertion. We then note:

$$\left(\frac{\partial}{\partial\theta} - \bar{\theta}\frac{\partial}{\partial t}\right) f(t + \theta\bar{\theta}) = 0\,,$$

and, therefore, the insertion of the first term in the r.h.s. of equation (17.23) immediately vanishes. We remain with the insertion of the operator $R(\phi)$:

$$R(\phi) = -\int_{t'}^{t''} \mathrm{d}t\, \mathrm{d}\bar{\theta}\, \mathrm{d}\theta\, \mathcal{A}\left[\varphi(t + \theta\bar{\theta})\right] + \mathcal{A}\left[\varphi(t'')\right]. \tag{17.25}$$

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 r.h.s. and we remain with:

$$R(\phi) = \mathcal{A}\left[\varphi(t')\right]. \tag{17.26}$$

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. We have, therefore, 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 (17.8) 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 fields in (17.8) by some 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 fields taken at $t = t'$. Due to cluster properties, connected correlation functions involving the insertion of such operators vanish in the large time separation limit, $|t'' - t'| \to \infty$. This implies that equal-time correlation functions converge at large time towards static equilibrium correlation functions. In explicit calculations we shall 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.

17.3 Stochastic Quantization of Two-Dimensional Chiral Models

It follows from the analysis of Section 16.9 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.

We have shown in Chapter 15 that two-dimensional models which 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 15.4–15.6. We here describe Langevin equations for $G\times G/G$ chiral models, G being a simple compact group. Note that some expressions explicitly refer to unitary groups but the generalization to other groups is simple.

A Langevin equation for chiral fields. In chiral models the field $g(x)$ varies in some representation of a Lie group G. We have shown in Section 15.4 that a static action $\mathcal{A}$ with only two derivatives has the form:

$$\mathcal{A}=-\frac{\beta}{2}\int \mathrm{d}^d x\ \operatorname{tr} j_\mu^2(x), \tag{17.27}$$

where $j_\mu(x)$ is the G-current given by

$$j_\mu(x)=g^{-1}(x)\partial_\mu g(x) \tag{17.28}$$

(it belongs to the Lie algebra 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 (equation (15.71)) expresses the current conservation $\partial_\mu j_\mu(x)=0$.

It is possible to write a group covariant Langevin equation which leads to an equilibrium distribution corresponding to the static action (17.27):

$$j_0(x,t)=(\Omega\beta/2)\partial_\mu j_\mu(x,t)+g^{-1}(x,t)\nu(x,t)-\nu^\dagger(x,t)g(x,t), \tag{17.29}$$

where the current $j_0(x,t)$ is defined by

$$j_0=g^{-1}\partial_t g \tag{17.30}$$

(below the index 0 always refers to time). The noise $\nu(x,t)$ is a general complex matrix with gaussian probability distribution $[\mathrm{d}\rho(\nu)]$,

$$[\mathrm{d}\rho(\nu)]=[\mathrm{d}\nu]\exp\left[-\frac{1}{2\Omega}\int \mathrm{d}^d x\,\mathrm{d}t\ \operatorname{tr}\left(\nu^\dagger(x,t)\nu(x,t)\right)\right]. \tag{17.31}$$

Note that $\nu(x,t)$ belongs to the linear representation of G and not only to the Lie algebra: it has more degrees of freedom than the field.

The dynamic action. To construct the dynamic action we follow the same steps as in preceding sections. However, since we have to integrate over $g(x,t)$ with the group invariant measure, the expressions are slightly modified as we expect from the analysis of Section 16.4. Similar expressions will again 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 19.1):

$$\nabla_\mu = \partial_\mu + [j_\mu, \cdot] \qquad \nabla_0 = \partial_t + [j_0, \cdot]\,. \tag{17.32}$$

In an infinitesimal space–time-dependent group transformation,

$$g(x,t) \mapsto g(x,t)\,(1 + \ell(x,t))\,, \tag{17.33}$$

the currents transform like

$$j_\mu \mapsto j_\mu + \nabla_\mu \ell, \qquad j_0 \mapsto j_0 + \nabla_0 \ell\,. \tag{17.34}$$

Therefore, the analogue of the operator $\mathbf{M}$ introduced in equation (16.129) is now defined, when acting on a field ℓ, as

$$\mathbf{M}\ell = \nabla_0 \ell - \tfrac{1}{2}\beta\Omega\partial_\mu\nabla_\mu \ell + \ell g^{-1}\nu + \nu^\dagger g\ell\,. \tag{17.35}$$

Introducing fermion fields $\bar c$ and c as well as a Lagrange multiplier $\bar\varphi$, all belonging to the Lie algebra of G, we can write, in the normalization of equation (17.2), the dynamic action $\mathcal{S}$ resulting from the integration over the noise as

$$\mathcal{S} = \int \mathrm{d}^d x\,\mathrm{d}t\,\frac{2}{\Omega}\,\mathrm{tr}\left[(\bar\varphi + c\bar c)\,(\bar\varphi + \bar c c) + \left(j_0 - \tfrac{1}{2}\Omega\beta\partial_\mu j_\mu\right)\bar\varphi - c\left(\nabla_0 - \tfrac{1}{2}\Omega\beta\partial_\mu\nabla_\mu\right)\bar c\right]. \tag{17.36}$$

We have used in this expression the identity $(\bar c c)^\dagger = -c\bar c$.

Note the appearance of a quartic fermion term induced by the dependence of the noise term on the field $g(x,t)$ in equation (17.29). As we expect from the analysis of Section 16.4, the set of BRS transformations which leave the action invariant is now

$$\delta g = \varepsilon g \bar c\,, \qquad \delta \bar c = -\varepsilon \bar c^2, \tag{17.37}$$

$$\delta c = \varepsilon \bar\varphi\,, \qquad \delta\bar\varphi = 0\,. \tag{17.38}$$

The BRS transformation of the current j_μ induced by equation (17.37) can then be written as

$$\delta j_\mu = \varepsilon \nabla_\mu \bar c\,. \tag{17.39}$$

The BRS transformations (17.37–17.39) are exactly those that will again appear in the quantization of non-abelian gauge theories in Chapter 19. The gauge field associated with the group G transforms as the current in (17.39).

We now show that the Langevin equation (17.29) is the natural generalization of equation (17.1). We introduce the superfield $\mathcal{G}$, element of the group G,

$$\mathcal{G} = g(1 + \theta\bar c)(1 + c\bar\theta)(1 + \theta\bar\theta\bar\varphi)\,. \tag{17.40}$$

The set of equations (17.37,17.38) is then equivalent to

$$\delta\mathcal{G} = \varepsilon\frac{\partial}{\partial\theta}\mathcal{G}\,. \tag{17.41}$$

From $\mathcal{G}$, we derive a set of associated currents J_μ,

$$J_\mu = \mathcal{G}^{-1}\partial_\mu\mathcal{G}\,, \tag{17.42}$$

and $J_{\bar\theta}$, J_θ:

$$J_{\bar\theta} = \mathcal{G}^{-1}\bar{\mathrm{D}}\mathcal{G}\,, \quad J_\theta = \mathcal{G}^{-1}\mathrm{D}\mathcal{G} \tag{17.43}$$

where $\bar{\mathrm{D}}, \mathrm{D}$ are the covariant derivatives (17.3).

The action (17.36) can then be written in a form analogous to (17.2) as

$$\mathcal{S} = -\int \mathrm{d}^d x\, \mathrm{d}t\, \mathrm{d}\bar\theta\, \mathrm{d}\theta\, \mathrm{tr}\left(\frac{2}{\Omega} J_{\bar\theta} J_\theta + \frac{\beta}{2} J_\mu^2\right). \tag{17.44}$$

In this form it is clear that $\mathcal{S}$ is supersymmetric. Supersymmetry implies that equal-time $g(x,t)$-field correlation functions converge at time $+\infty$ towards the correlation functions of the static action (17.27), justifying the choice of the Langevin equation (17.29). We show in the next section that the equation (17.29) 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 (17.7)): 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.

17.4 Langevin Equation and Riemannian Manifolds

We have already seen in Section 4.8 that it is impossible to write a Langevin equation which only depends on the geometry of the manifold. We, therefore, construct a Langevin equation in the following way: we consider smooth manifolds embedded in an euclidean space and defined by a set of equations constraining the euclidean coordinates χ_α:

$$E^s(\chi_\alpha) = 0\,. \tag{17.45}$$

To simplify calculations, we assume that we have solved locally these equations and expressed some components σ^s of χ_α as functions of a set of independent components φ^i:

$$\chi_\alpha \equiv \left\{\sigma^s(\varphi) \quad \varphi^i\right\}. \tag{17.46}$$

To construct a Langevin equation on the manifold, we start from a Langevin equation of type (17.1) in the embedding euclidean space,

$$\dot\chi_\alpha(x,t) = -\frac{\Omega}{2}\frac{\delta\mathcal{A}}{\delta\chi_\alpha(x,t)} + \nu_\alpha(x,t), \tag{17.47}$$

in which $\nu_\alpha(x,t)$ is a gaussian noise:

$$\langle \nu_\alpha(x,t)\nu_\beta(x',t')\rangle = \Omega\delta(t-t')\delta(x-x')\delta_{\alpha\beta}\,. \tag{17.48}$$

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 now project $\dot\chi_\alpha(x,t)$ onto the space tangent to the manifold at point $\chi_\alpha(x,t)$. For this purpose we introduce some notation. On the manifold the variations $\delta\chi_\alpha$ of the field χ_α are constrained by

$$\frac{\partial E^s}{\partial\chi_\alpha}\delta\chi_\alpha = 0\,. \tag{17.49}$$

We introduce an orthogonal basis e_a^α for the variations of χ_α:

$$\frac{\partial E^s}{\partial \chi_\alpha} e_a^\alpha = 0\,, \tag{17.50}$$

with the orthogonality conditions

$$e_a^\alpha e_b^\alpha = \delta_{ab}\,. \tag{17.51}$$

In terms of the coordinates φ^i, equation (17.50) can be written as

$$e_a^s = \partial_i \sigma^s e_a^i\,, \tag{17.52}$$

and, therefore, equation (17.51) becomes, after elimination of the components e_a^s,

$$e_a^i g_{ij} e_b^j = \delta_{ab}\,, \tag{17.53}$$

where we have introduced the metric tensor g_{ij} on the manifold (see Appendix A15.3.2):

$$g_{ij} = \delta_{ij} + \partial_i \sigma^s \partial_j \sigma^s. \tag{17.54}$$

Equation (17.53) shows that the matrix e_a^i is the inverse of the vielbein (Section 22.6).

We can now construct from equation (17.47) a Langevin equation on the manifold:

$$\dot\chi_\alpha(x,t) = e_a^\alpha e_a^\beta \left(-\frac{\Omega}{2}\frac{\delta \mathcal{A}}{\delta \chi_\beta} + \nu_\beta\right). \tag{17.55}$$

In terms of the independent components φ^i the equation reads

$$\dot\varphi^i = -\frac{\Omega}{2} e_a^i \left(e_a^j \frac{\delta \mathcal{A}}{\delta \varphi^j} + \partial_j \sigma^s e_a^j \frac{\delta \mathcal{A}}{\delta \sigma^s}\right) + e_a^i \left(e_a^j \nu_j + e_a^j \partial_j \sigma^s \nu_s\right). \tag{17.56}$$

We observe that

$$\frac{\delta \mathcal{A}}{\delta \varphi^j}(\varphi,\sigma) + \partial_j \sigma^s(\varphi) \frac{\delta \mathcal{A}}{\delta \sigma^s} = \partial_j \mathcal{A}\big(\varphi, \sigma(\varphi)\big), \tag{17.57}$$

in which in the r.h.s. $\partial_j \mathcal{A}$ means total derivative of $\mathcal{A}$ with respect to φ^j.

As a consequence of equation (17.53) the inverse metric tensor g^{ij} is given by

$$g^{ij} = e_a^i e_a^j\,. \tag{17.58}$$

Equation (17.56) can thus be written as

$$\dot\varphi^i = -\tfrac{1}{2}\Omega g^{ij} \partial_j \mathcal{A} + g^{ij} t_j^\alpha \nu_\alpha\,, \tag{17.59}$$

with the definition

$$t_i^\alpha \equiv \begin{cases} t_i^j = \delta_{ij}\,, \\ t_i^s = \partial_i \sigma^s. \end{cases} \tag{17.60}$$

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 4.8.

The quantities t_i^α satisfy

$$t_i^\alpha t_j^\alpha = g_{ij}\,, \tag{17.61}$$

and

$$\partial_i t^\alpha_j t^\alpha_k = g_{kl}\Gamma^l_{ij}\,, \tag{17.62}$$

in which Γ^l_{ij} is the usual Christoffel symbol (equation (22.41)). This equation can be rewritten in covariant form as

$$\nabla_i t^\alpha_j t^\alpha_k = 0\,, \tag{17.63}$$

in which ∇_i is the covariant derivative on the manifold.

Dynamic action. It is convenient to rewrite equation (17.59) as

$$g_{ij}\dot{\varphi}^j + \tfrac{1}{2}\Omega\partial_i\mathcal{A} - t^\alpha_i\nu_\alpha = 0\,. \tag{17.64}$$

Introducing a Lagrange multiplier $\bar{\varphi}^i$ and fermion fields c^i and c^i, we can write the corresponding dynamic action $\mathcal{S}$, before integration over the noise, as the sum of two contributions $\mathcal{S}_0$ and $\mathcal{S}_1$:

$$\mathcal{S} = \mathcal{S}_0 + \mathcal{S}_1 \tag{17.65}$$

with

$$\mathcal{S}_0 = \int \mathrm{d}x\,\mathrm{d}t\left(\frac{1}{2\Omega}\nu_\alpha\nu_\alpha - \frac{2}{\Omega}t^\alpha_i\nu_\alpha\bar{\varphi}^i + \frac{2}{\Omega}c^i\partial_j t^\alpha_i\bar{c}^j\nu_\alpha\right), \tag{17.66}$$

$$\mathcal{S}_1 = \int \mathrm{d}x\,\mathrm{d}t\frac{2}{\Omega}\left[\bar{\varphi}^i\left(g_{ij}\dot{\varphi}^j + \tfrac{1}{2}\Omega\partial_i\mathcal{A}\right) + c^i g_{ij}\dot{\bar{c}}^j - c^i\partial_k g_{ij}\dot{\varphi}^j\bar{c}^k - \tfrac{1}{2}\Omega c^i\partial_i\partial_j\mathcal{A}\,\bar{c}^j\right]. \tag{17.67}$$

Integration over the noise after a short calculation leads to

$$\mathcal{S} = \mathcal{S}_1 + \int \mathrm{d}x\,\mathrm{d}t\frac{2}{\Omega}\left(-\bar{\varphi}^i g_{ij}\bar{\varphi}^j + 2\bar{\varphi}^i g_{il}\Gamma^l_{jk}c^j\bar{c}^k + c^i\bar{c}^j c^k\bar{c}^l\partial_i\partial_j g_{kl}\right). \tag{17.68}$$

Introducing now the superfield ϕ^i:

$$\phi^i = \varphi^i + \theta\bar{c}^i + c^i\bar{\theta} + \theta\bar{\theta}\bar{\varphi}^i, \tag{17.69}$$

and the covariant derivatives (17.3) we can rewrite the action as

$$\mathcal{S}(\phi) = \int \mathrm{d}t\,\mathrm{d}\bar{\theta}\,\mathrm{d}\theta\left[\int \mathrm{d}x\,\frac{2}{\Omega}\bar{\mathrm{D}}\phi^i g_{ij}(\phi)\mathrm{D}\phi^j + \mathcal{A}(\phi)\right]. \tag{17.70}$$

This supersymmetric form then implies that at equilibrium the equal-time field configurations are weighted by the measure $[\mathrm{d}\rho(\varphi)]$:

$$[\mathrm{d}\rho(\varphi)] = \prod_i [\mathrm{d}\varphi^i]\sqrt{g(\varphi)}\exp\left[-\mathcal{A}(\varphi,\sigma(\varphi))\right]. \tag{17.71}$$

This was the field distribution we had in mind when we wrote equation (17.59). Note that in contrast to the Langevin equation (17.59), the dynamic action (17.70) and, therefore, the correlation functions at equilibrium depend only on the geometry of the manifold. In the case of homogeneous spaces the original action $\mathcal{A}(\chi)$ is simply:

$$\mathcal{A}(\chi) = \frac{1}{2}\int (\partial_\mu\chi_\alpha)^2\,\mathrm{d}x\,, \tag{17.72}$$

which, as a functional of φ, becomes on the manifold:

$$\mathcal{A}(\varphi, \sigma(\varphi)) = \frac{1}{2}\int g_{ij}(\varphi)\partial_\mu\varphi^i\partial_\mu\varphi^j \mathrm{d}x \,. \tag{17.73}$$

One can verify that the expression (17.70) contains the expression (17.44) as a special case.

Note, finally, that the functional integral representation for the generating functional $\mathcal{Z}(J)$ can also be rewritten, introducing the superfield X in euclidean coordinates:

$$X_\alpha = \chi_\alpha + \theta\bar{c}_\alpha + c_\alpha\bar{\theta} + \bar{\varphi}_\alpha\theta\bar{\theta}\,, \tag{17.74}$$

in the form

$$\begin{aligned}\mathcal{Z}(J) &= \int [\mathrm{d}\phi]\exp\left[-\mathcal{S}(\phi) + \int \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\bar{\theta}\,\mathrm{d}\theta\, J_i\phi^i\right]\\ &= \int [\mathrm{d}X\,]\,\delta\left[E^s(X)\right]\exp\left(-\Sigma(X) + \int \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\bar{\theta}\,\mathrm{d}\theta\, J_i\phi^i\right)\end{aligned} \tag{17.75}$$

with, for $\Sigma(X)$, the expression

$$\Sigma(X) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t\left[\int \mathrm{d}x\,\frac{2}{\Omega}\bar{\mathrm{D}}X_\alpha\mathrm{D}X_\alpha + \mathcal{A}(X)\right]. \tag{17.76}$$

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 15 and supersymmetry.

17.5 Scalar Supersymmetric Fields Below Four Dimensions

In the two coming sections we present, for completeness, a very brief introduction to supersymmetric quantum field theories. The new feature of supersymmetry in higher dimensions, compared to the quantum mechanics case, is the combination of supersymmetry with spin since fermions have spins, and this leads to modifications we want to discuss here. We begin with models involving only a scalar superfield in dimensions two and three, where fermion spin is associated with Pauli matrices. As it will become clear shortly, such models constitute the simplest generalization of supersymmetric quantum mechanics as it naturally appears for example in the study of stochastic differential equations of Langevin type (see Section 17.1).

We first consider a simple supersymmetric model, with one scalar superfield, in three euclidean dimensions. We then examine the example of the supersymmetric non-linear σ model, very much as we have done in the non supersymmetric examples. To discuss supersymmetric models we use the superfield notation already introduced in supersymmetric quantum mechanics.

17.5.1 The supersymmetric scalar field in three dimensions

Supersymmetry and Majorana spinors in $d = 3$. We first explain our notation, summarize a few properties of Majorana spinors in three dimensions and construct a scalar superfield.

In three euclidean dimensions the spin group is $SU(2)$ (see Appendix A8). A spinor transforms like

$$\psi_U = U\psi\,, \quad U \in SU(2) \ .$$

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

$$\sigma_2\sigma_\mu\sigma_2 = -{}^T\sigma_\mu \ \Rightarrow \ U^* = \sigma_2 U \sigma_2 \ .$$

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 Section 5.3.

It is convenient to define a conjugated spinor,

$$\bar\psi = {}^T\psi\sigma_2 \ \Rightarrow \ \bar\psi_\alpha = i\epsilon_{\alpha\beta}\psi_\beta\,, \tag{17.77}$$

($\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha}$, $\epsilon_{12} = 1$) which thus transforms like

$$[{}^T\psi\sigma_2]_U = [{}^T\psi\sigma_2]U^\dagger\,.$$

Corresponding to fermion spinors we need a spinor θ_α of real Grassmann coordinates. We also define

$$\bar\theta = {}^T\theta\sigma_2 \ \Rightarrow \ \bar\theta_\alpha = i\epsilon_{\alpha\beta}\theta_\beta\,.$$

Since the only non-vanishing product is $\theta_1\theta_2$, we have

$$\bar\theta_\alpha\theta_\beta = \tfrac{1}{2}\delta_{\alpha\beta}\bar\theta\cdot\theta\,.$$

The scalar product $\bar\theta\cdot\theta$ is

$$\bar\theta\cdot\theta = -2i\theta_1\theta_2 \ \Rightarrow \theta_\alpha\bar\theta_\beta = i\delta_{\alpha\beta}\theta_1\theta_2\,.$$

If $\{\bar\theta', \theta'\}$ is another pair of coordinates, because $\sigma_2\sigma_\mu$ is symmetric, one finds

$$\bar\theta\sigma_\mu\theta' = {}^T\theta\sigma_2\sigma_\mu\theta' = -\bar\theta'\sigma_\mu\theta\,, \tag{17.78}$$

and for the same reason

$$\bar\theta\psi = \bar\psi\theta\,.$$

Other useful identities are

$$(\bar\theta\psi)^2 = -\tfrac{1}{2}(\bar\theta\theta)(\bar\psi\psi), \quad (\bar\theta\not{p}\psi)^2 = \tfrac{1}{2}p^2(\bar\psi\psi)(\bar\theta\theta).$$

Finally, we integrate over θ_1, θ_2 with the measure

$$\mathrm{d}^2\theta \equiv \tfrac{1}{2}i\mathrm{d}\theta_2\mathrm{d}\theta_1\,,$$

which is invariant under complex conjugation. Then,

$$\int \mathrm{d}^2\theta\, \bar\theta_\alpha \theta_\beta = \tfrac{1}{2}\delta_{\alpha\beta}\,, \quad \int \mathrm{d}^2\theta\, \bar\theta \cdot \theta = 1\,.$$

With this convention, the δ-function $\delta^2(\theta'-\theta)$ in θ space is

$$\delta^2(\theta'-\theta) = (\bar\theta' - \bar\theta)\cdot(\theta'-\theta). \tag{17.79}$$

Superfields and covariant derivatives. A scalar superfield $\Phi(\theta)$ has the expansion

$$\Phi(\theta, x) = \varphi(x) + \bar\theta\psi(x) + \tfrac{1}{2}\bar\theta\theta F(x)\,. \tag{17.80}$$

Again, although only two θ variables are independent, it is convenient to define the covariant derivatives D_α and $\bar{\mathrm{D}}_\alpha$ ($\bar{\mathrm{D}} = \sigma_2 \mathrm{D}$)

$$\bar{\mathrm{D}}_\alpha \equiv \frac{\partial}{\partial \bar\theta_\alpha} - (\not{\partial}\theta)_\alpha\,, \quad \mathrm{D}_\alpha \equiv \frac{\partial}{\partial \theta_\alpha} - (\bar\theta\not{\partial})_\alpha\,. \tag{17.81}$$

Then the anticommutation relation is

$$\{\bar{\mathrm{D}}_\alpha, \mathrm{D}_\beta\} = -2[\not{\partial}]_{\alpha\beta}\,.$$

Also,

$$\mathrm{D}_\alpha \bar{\mathrm{D}}_\alpha = \frac{\partial}{\partial\theta_\alpha}\frac{\partial}{\partial\bar\theta_\alpha} - (\bar\theta\not{\partial})_\alpha \frac{\partial}{\partial\bar\theta_\alpha} - \frac{\partial}{\partial\theta_\alpha}(\not{\partial}\theta)_\alpha + \bar\theta_\alpha\theta_\alpha\partial^2\,.$$

Since the σ_μ are traceless, using the identity (17.78) one verifies that $\mathrm{D}\bar{\mathrm{D}}$ can also be written as

$$\mathrm{D}_\alpha \bar{\mathrm{D}}_\alpha = \frac{\partial}{\partial\theta_\alpha}\frac{\partial}{\partial\bar\theta_\alpha} - 2(\bar\theta\not{\partial})_\alpha \frac{\partial}{\partial\bar\theta_\alpha} + \bar\theta_\alpha\theta_\alpha\partial^2\,,$$

and, therefore, in component form,

$$\mathrm{D}_\alpha \bar{\mathrm{D}}_\alpha \Phi = 2F - 2\bar\theta\not{\partial}\psi + \bar\theta\theta\partial^2\varphi\,.$$

Supersymmetry generators and WT identities. Supersymmetry is generated by the operators Q_α (or $\bar{\mathrm{Q}} = \sigma_2\mathrm{Q}$) which anticommute with $\bar{\mathrm{D}}_\alpha$ (and thus D_α):

$$\mathrm{Q}_\alpha = \frac{\partial}{\partial\theta_\alpha} + (\bar\theta\not{\partial})_\alpha\,.$$

Then,

$$\{\bar{\mathrm{Q}}_\alpha, \mathrm{Q}_\beta\} = 2[\not{\partial}]_{\alpha\beta}\,.$$

The anticommutator of $\bar{\mathrm{Q}}$ and Q now is the generator of space translations: supersymmetry implies space translation invariance.

Supersymmetry also implies WT identities for correlation functions. The n-point function $W^{(n)}(p_k, \theta_k)$ in momentum space satisfies

$$\mathrm{Q}_\alpha W^{(n)} \equiv \left[\sum_k \frac{\partial}{\partial\theta^k_\alpha} + i(\bar\theta^k \not{p}_k)_\alpha\right] W^{(n)}(p,\theta) = 0\,.$$

To solve this equation, we set

$$W^{(n)}(p,\theta) = F^n(p,\theta)\exp\left[\frac{i}{2n}\sum_{jk}\bar\theta^j(\not p_j - \not p_k)\theta^k\right], \tag{17.82}$$

where $F^{(n)}$ is a symmetric function in the exchange $\{p_i,\theta_i\}\leftrightarrow\{p_j,\theta_j\}$. It then satisfies

$$\sum_k \frac{\partial}{\partial\theta^k_\alpha}F^{(n)}(p,\theta) = 0\,,$$

that is, it is translation invariant in θ space.

In the case of the two-point function this leads to the general form

$$\begin{aligned} W^{(2)}(p,\theta',\theta) &= A(p^2)\left[1 + C(p^2)\delta^2(\theta'-\theta)\right]\mathrm{e}^{i\bar\theta\not p\theta'}\,. \\ &= A(p^2)\left[1 + C(p^2)(\bar\theta'-\bar\theta)(\theta'-\theta) + i\bar\theta\not p\theta' - \tfrac14 p^2\bar\theta\theta\bar\theta'\theta'\right]. \end{aligned} \tag{17.83}$$

The same identities apply to the 1PI correlations $\Gamma^{(n)}$.

17.5.2 General $O(N)$ symmetric action

We now consider a general $O(N)$ invariant supersymmetric action $\mathcal{S}(\Phi)$

$$\mathcal{S}(\Phi) = \int \mathrm{d}^3x\,\mathrm{d}^2\theta\left[\tfrac12 \mathrm{D}\Phi\cdot\bar{\mathrm{D}}\Phi + U(\Phi^2)\right], \tag{17.84}$$

where Φ is an N-component vector.

In component notation,

$$\int \mathrm{d}^2\theta\,\mathrm{D}\Phi\bar{\mathrm{D}}\Phi = -\int \mathrm{d}^2\theta\,\Phi\mathrm{D}_\alpha\bar{\mathrm{D}}_\alpha\Phi = -\bar\psi\not\partial\psi + (\partial_\mu\varphi)^2 - F^2. \tag{17.85}$$

Then, since

$$\Phi^2 = \varphi^2 + 2\varphi\bar\theta\psi - \tfrac12(\bar\theta\theta)(\bar\psi\psi) + F\varphi\bar\theta\theta\,, \tag{17.86}$$

we find,

$$\int \mathrm{d}^2\theta\,U(\Phi^2) = U'(\varphi^2)\left(-\tfrac12\bar\psi\psi + F\varphi\right) - U''(\varphi^2)(\bar\psi\varphi)(\varphi\psi).$$

In the case of the free theory $U(\rho)\equiv\mu\rho$, the action in component form is

$$\mathcal{S} = \int \mathrm{d}^3x\,\left[-\tfrac12\bar\psi\not\partial\psi + \tfrac12(\partial_\mu\varphi)^2 - \tfrac12F^2 + \mu\left(-\tfrac12\bar\psi\psi + F\varphi\right)\right].$$

After integration over the auxiliary F field the action becomes

$$\mathcal{S} = \int \mathrm{d}^3x\,\left[-\tfrac12\bar\psi\not\partial\psi - \tfrac12\mu\bar\psi\psi + \tfrac12(\partial_\mu\varphi)^2 + \tfrac12\mu^2\varphi^2\right].$$

For a generic super-potential $U(\rho)$ one finds

$$\mathcal{S} = \int \mathrm{d}^3x\,\left[-\tfrac12\bar\psi\not\partial\psi + \tfrac12(\partial_\mu\varphi)^2 - \tfrac12F^2 + U'(\varphi^2)\left(-\tfrac12\bar\psi\psi + F\varphi\right) - U''(\varphi^2)(\bar\psi\varphi)(\varphi\psi)\right], \tag{17.87}$$

and thus again after integration over the auxiliary F,

$$\mathcal{S} = \int \mathrm{d}^3x \left[-\tfrac{1}{2}\bar{\psi}\not{\partial}\psi + \tfrac{1}{2}(\partial_\mu\varphi)^2 - \tfrac{1}{2}U'(\varphi^2)\bar{\psi}\psi - U''(\varphi^2)(\bar{\psi}\varphi)(\varphi\psi) + \tfrac{1}{2}\varphi^2 U'^2(\varphi^2)\right]. \tag{17.88}$$

Note that the theory violates fermion number conservation (Majorana spinors) and parity symmetry. Actually a space reflection (see definition in Section A8.1.6) is equivalent to the change $U \mapsto -U$. Therefore, theories with $\pm U$ have the same physical properties.

Moreover, under the form (17.87) we notice 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 that transforms the action into a free action. This transformation implies that the partition function is one at zero-temperature.

Power counting. In three dimensions the superfield has canonical dimension 1/2 and $\theta, \bar{\theta}$ dimensions $-1/2$

$$[\theta] = -\tfrac{1}{2}, \quad [\Phi] = \tfrac{1}{2}.$$

We infer that the $(\Phi^2)^2$ is the only theory renormalizable in three dimensions. Note that we have not considered here the Φ^3 theory for a one-component field, which is super-renormalizable (and thus has a renormalizable generalization in four dimensions). Prior to a more refined analysis one expects coupling constant and field renormalizations (with logarithmic divergences) and a mass renormalization with linear divergences. Using the solution (17.83) for the two-point function $\Gamma^{(2)}$, 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 the mass renormalization.

We, therefore, now consider the potential

$$U(\rho) = \mu\rho + \tfrac{1}{2}g\rho^2. \tag{17.89}$$

The super-propagator. To generate perturbation theory we need the field propagator. We notice that in momentum space and in terms of the field components,

$$[-\mathrm{D}\bar{\mathrm{D}} + 2\mu]\delta^2(\theta' - \theta) = 4\left[-1 - i\bar{\theta}\not{k}\theta' + \tfrac{1}{2}\mu(\bar{\theta}' - \bar{\theta})(\theta' - \theta) + \tfrac{1}{4}k^2\bar{\theta}\theta\bar{\theta}'\theta'\right]. \tag{17.90}$$

The super-propagator Δ of the ϕ-field

$$\left(-\mathrm{D}\bar{\mathrm{D}} + 2\mu\right)\Delta(k, \theta, \theta') = \delta^2(\theta' - \theta),$$

thus can be obtained by solving the equation

$$2\mu\varphi - 2F + 2\bar{\theta}(i\not{k} + \mu)\psi + \bar{\theta}\theta[(k^2 + \mu] = J(\theta).$$

It is given by

$$\begin{aligned}\Delta(k, \theta, \theta') &= \frac{1 + \frac{1}{2}\mu\delta^2(\theta - \theta')}{k^2 + \mu^2}\, \mathrm{e}^{i\bar{\theta}\not{k}\theta'} \\ &= \frac{1}{k^2 + \mu^2}\left[1 + \tfrac{1}{2}\mu(\bar{\theta}\theta + \bar{\theta}'\theta') + \bar{\theta}[i\not{k} - \mu]\theta' - \tfrac{1}{4}k^2\bar{\theta}\theta\bar{\theta}'\theta'\right].\end{aligned} \tag{17.91}$$

Clearly, one reads in equation (17.91) the $\langle\varphi(k)\varphi(-k)\rangle$ propagator $(k^2+\mu^2)^{-1}$ and the $\langle\bar{\psi}(k)\psi(-k)\rangle$ propagator $(i\not{k}-\mu)/(k^2+\mu^2)$. The coefficients of $(\bar{\theta}\theta+\bar{\theta}'\theta')$ and of $(\bar{\theta}\theta\bar{\theta}'\theta')$ are the $\langle\varphi(k)F(-k)\rangle$ and $\langle F(k)F(-k)\rangle$ propagators, respectively.

One-loop divergences. The one-loop diagram has a simple form

$$2\delta^2(\theta-\theta')(N+2)g\,\mathrm{tr}\,\Delta(k,\theta,\theta)=2\delta^2(\theta-\theta')\frac{N+2}{(2\pi)^3}g\int\frac{\mathrm{d}^3k}{k^2+\mu^2}\,.$$

It exhibits indeed the expected linear mass divergence and requires a mass renormalization $\delta\mu$, for example,

$$\delta\mu=-(N+2)g\frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{k^2}\,,$$

which vanishes in dimensional regularization.

The contribution to the four-point function is proportional to the usual bubble diagram

$$g^2\int\frac{\mathrm{d}^3k}{(2\pi)^3}\Delta(k,\theta,\theta')\Delta(p-k,\theta,\theta').$$

Then,

$$\Delta(k,\theta,\theta')\Delta(p-k,\theta,\theta')=\frac{[1+\mu\delta^2(\theta'-\theta)]\,\mathrm{e}^{i\bar{\theta}\not{p}\theta'}}{(k^2+\mu^2)[(p+k)^2+\mu^2]}\,.$$

Notice the cancellation of the factor $\mathrm{e}^{i\bar{\theta}\not{k}\theta'}$ which renders the integral more convergent that one would naively expect. The integral over k then yields the finite three-dimensional scalar bubble diagram

$$B(p)=\frac{1}{(2\pi)^3}\int\frac{\mathrm{d}^3k}{(k^2+\mu^2)[(p+k)^2+\mu^2]}=\frac{1}{4\pi p}\mathrm{Arctan}(p/2|\mu|)\,,$$

and no coupling constant renormalization is required, like in the non-supersymmetric, super-renormalizable $(\phi^2)^2$ field theory.

17.5.3 The $O(N)$ supersymmetric non-linear σ model

We now consider the supersymmetric non-linear σ-model in 2 dimensions. The action

$$\mathcal{S}(\Phi)=\frac{1}{2\kappa}\int\mathrm{d}^2x\,\mathrm{d}^2\theta\,\mathrm{D}\Phi\cdot\bar{\mathrm{D}}\Phi \tag{17.92}$$

involves a scalar superfield Φ, an N-component vector, which satisfies

$$\Phi\cdot\Phi=1\,.$$

Using the identity (17.86) one can write the constraint in component form:

$$\varphi^2=1\,,\quad\varphi\cdot\psi=0\,,\quad F\cdot\varphi=\tfrac{1}{2}\bar{\psi}\cdot\psi\,.$$

Using then the identity (17.85) and integrating over the auxiliary field F one obtains the action in component form:

$$\mathcal{S}=\frac{1}{2\kappa}\int\mathrm{d}^2x\,\left[(\partial_\mu\varphi)^2-\bar{\psi}\not{\partial}\psi-\tfrac{1}{4}(\bar{\psi}\psi)^2\right],$$

The constraint can also be implemented by introducing a superfield

$$\mathbf{L}(x,\theta) = L(x) + \bar\theta\ell(x) + \tfrac{1}{2}\bar\theta\theta\lambda(x),$$

where $L(x)$, $\lambda(x)$ and $\ell(x)$ are the Lagrange multiplier fields, and adding to the action

$$\mathcal{S}_L = \frac{1}{\kappa}\int \mathrm{d}^2x\,\mathrm{d}^2\theta\,\mathbf{L}(x,\theta)\left[\Phi^2(x,\theta)-1\right]. \tag{17.93}$$

The partition function is given by ($\mathcal{S}(\Phi,\mathbf{L}) = \mathcal{S} + \mathcal{S}_L$):

$$\mathcal{Z} = \int [d\Phi][\mathrm{d}L]\,\mathrm{e}^{-\mathcal{S}(\Phi,\mathbf{L})}.$$

In terms of component fields, after integration over the auxiliary field F, the total action reads

$$\mathcal{S} = \frac{1}{2\kappa}\int \mathrm{d}^2x\{\varphi(-\nabla^2+L^2)\varphi - \bar\psi(\not\partial + L)\psi + \lambda(\varphi^2-1) - 2\bar\ell(\bar\psi\cdot\varphi)\}. \tag{17.94}$$

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 IR divergences it is also necessary here to introduce a mass term by breaking the $O(N)$ symmetry explicitly, which can be obtained by adding a term linear in Φ in the action.

Much is known about the model in two-dimensions including the exact S-matrix. The calculation of the RG β-function at one-loop in the superfield formalism is analogous to the calculation in the non-supersymmetric model; one simply changes the propagator. Actually the β-function has been calculated up to four loops. In the minimal subtraction scheme,

$$\beta(\kappa) = -(N-2)\tilde\kappa - \tfrac{3}{2}\zeta(3)(N-2)(N-3)\tilde\kappa^5 + O(\tilde\kappa^6),$$

with $\tilde\kappa = \kappa/2\pi$.

17.6 Supersymmetry in Four Dimensions

In this section we present, for completeness, a very brief discussion of supersymmetry in four dimensions, because four dimensions is the dimension relevant for particle physics, and the implementation of supersymmetry has features that depend on dimensions. Moreover, we have not included gauge fields yet, something obviously important for applications to particle physics.

The algebra of generators of supersymmetry is briefly described in Section A17.1. We again use the superfield formalism of Section 17.5 and for instance combine scalar boson and spin 1/2 fermion fields into scalar superfields.

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. In any space dimension, the dimension $[\bar\theta]$ of the coordinates $\bar\theta$ is $[\theta] = [A] - [\psi] = -1/2$. If we give to $\bar\theta$ Dirac spinor indices the corresponding Grassmann algebra has 8 different generators $\theta_\alpha, \bar\theta_\alpha$. Therefore, the volume elements $\prod_\alpha \mathrm{d}\theta_\alpha \mathrm{d}\bar\theta_\alpha$ has dimension 4 and thus only a constant action is allowed. To circumvent this difficulty, one uses the specific properties of the spinor group Spin(4) (see Section

A8.2) which is $SU(2) \times SU(2)$. First the spinorial representation can be reduced into left-handed and right-handed components (Weyl spinors, see Appendix A8.2). Moreover, the representations of $SU(2)$ are self-conjugated. It is thus sufficient to associate Grassmann coordinates with the two Weyl spinors. We only need four Grassmann coordinates $\theta_\alpha, \bar\theta_\alpha$. In this form the Grassmann volume element has only dimension 2 and an action bilinear in the scalar superfield can be written: such an action term 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 $\mathrm{d}\theta_1\mathrm{d}\theta_2$ ($\mathrm{d}\bar\theta_1\mathrm{d}\bar\theta_2$ respectively) which have dimension 1. The corresponding contribution is called an F-term.

To write the generators of supersymmetry (see Section 17.5) it is convenient to introduce a notation (different from (17.77)) for the two-spinor transforming under the conjugated representation of $SU(2)$

$$\theta^\alpha = (\sigma_2)_{\alpha\beta}\theta_\beta \equiv -i\epsilon_{\alpha\beta}\theta_\beta\,, \tag{17.95}$$

where $\epsilon_{\alpha\beta}$ is the antisymmetric tensor. We also define the four matrices σ_μ, where for $\mu = 1,2,3$, σ_μ are just the usual Pauli matrices and $\sigma_4 = -i\mathbf{1}$. Two relations are useful in what follows:

$$\sigma_2\sigma_\mu\sigma_2 = -\sigma^*_\mu, \qquad \sigma^\dagger_\mu\sigma_\nu + \sigma^\dagger_\nu\sigma_\mu = 2\delta_{\mu\nu}\,.$$

Then, under a $SU(2) \times SU(2)$ transformation,

$$\theta'_\alpha = U^\beta_\alpha\theta_\beta\,, \quad \bar\theta^{\alpha\prime} = V^{*\alpha}_\beta\bar\theta^\beta\,, \quad \sigma_\mu x'_\mu = V\sigma_\mu x_\mu U^\dagger, \quad U,V \in SU(2).$$

$SU(2)$ invariant quantities are then, for example, $\theta^\alpha\theta_\alpha$, $\bar\theta^\alpha\bar\theta_\alpha$ or $\bar\theta^\alpha\sigma_{\mu\alpha}^{\ \beta}\theta_\beta$.

The supersymmetry generators,

$$\bar{\mathrm{Q}}_\alpha = \frac{\partial}{\partial\bar\theta^\alpha} + \frac{1}{2}\sigma_{\mu\alpha}^{\ \beta}\theta_\beta\frac{\partial}{\partial x_\mu}, \tag{17.96a}$$

$$\mathrm{Q}^\alpha = \frac{\partial}{\partial\theta_\alpha} + \frac{1}{2}\bar\theta^\beta\sigma_{\mu\beta}^{\ \alpha}\frac{\partial}{\partial x_\mu}, \tag{17.96b}$$

belong to the representations $(0,1/2)$ and $(1/2,0)$ of $SU(2) \times SU(2)$, respectively. They have the anticommutation relations,

$$\{\bar{\mathrm{Q}}_\alpha, \mathrm{Q}^\beta\} = \sigma_{\mu\alpha}^{\ \beta}\frac{\partial}{\partial x_\mu}, \qquad \{\mathrm{Q},\mathrm{Q}\} = \{\bar{\mathrm{Q}},\bar{\mathrm{Q}}\} = 0\,. \tag{17.97}$$

Note, therefore, that the supersymmetry current has spin 3/2. This implies in particular 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. Although massless spin 1/2 fermions do exist (neutrinos) they do not obey the consequences of the corresponding WT identities. It is actually a problem of the supersymmetry assumption that none of the partners of known particles in supermultiplets have been found yet. Note finally that since the commutator of supersymmetries is the generator of translations any attempt to gauge supersymmetry (which would provide a solution to the unwanted Goldstone fermion) results in a theory which contains 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.

To $\mathrm{Q}, \bar{\mathrm{Q}}$ correspond two operators $\mathrm{D}, \bar{\mathrm{D}}$ which also satisfy the anticommutation relations of supersymmetry generators, anticommute with them and thus play the role of covariant derivatives (see for instance Section A17.1):

$$\bar{\mathrm{D}}_\alpha = \frac{\partial}{\partial\bar{\theta}^\alpha} - \frac{1}{2}\sigma_{\mu\alpha}^{\ \ \beta}\theta_\beta\frac{\partial}{\partial x_\mu}, \qquad \mathrm{D}^\alpha = \frac{\partial}{\partial\theta_\alpha} - \frac{1}{2}\bar{\theta}^\beta\sigma_{\mu\beta}^{\ \ \alpha}\frac{\partial}{\partial x_\mu}, \tag{17.98}$$

$$\{\bar{\mathrm{D}}_\alpha, \mathrm{D}^\beta\} = -\sigma_{\mu\beta}^{\ \ \alpha}\frac{\partial}{\partial x_\mu}, \qquad \{\mathrm{D}, \mathrm{D}\} = \{\bar{\mathrm{D}}, \bar{\mathrm{D}}\} = 0\,. \tag{17.99}$$

17.6.1 Scalar chiral superfields

Because $\bar{\mathrm{D}}$ anticommutes with $\bar{\mathrm{Q}}, \mathrm{Q}$ the fields which satisfy

$$\bar{\mathrm{D}}_\alpha\phi(x,\theta,\bar{\theta}) = 0\,, \tag{17.100}$$

form a space of representation for the supersymmetry generators. The general solution of equation (17.100) can be written as

$$\phi(x,\theta,\bar{\theta}) = \phi(y,\theta),$$

where the new space coordinate y is

$$y_\mu = x_\mu + \tfrac{1}{2}\bar{\theta}^\alpha\sigma_{\mu\alpha}^{\ \ \beta}\theta_\beta\,. \tag{17.101}$$

Note that y also satisfies the equation

$$\mathrm{Q}^\alpha y_\mu(x,\theta,\bar{\theta}) = 0\,. \tag{17.102}$$

Acting on functions of y, θ, the generators $\mathrm{Q}, \bar{\mathrm{Q}}$ then take the form

$$\bar{\mathrm{Q}}_\alpha = \frac{\partial}{\partial\bar{\theta}^\alpha} + \sigma_{\mu\alpha}^{\ \ \beta}\theta_\beta\frac{\partial}{\partial y_\mu}, \quad \mathrm{Q}^\alpha = \frac{\partial}{\partial\theta_\alpha}.$$

Note that the variable y plays a role similar to time in the 1-D example. 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

$$\phi(y,\theta) = A(y) + \psi^\alpha(y)\theta_\alpha + \tfrac{1}{2}E(y)\theta^\alpha\theta_\alpha\,. \tag{17.103}$$

A and E are two complex scalar fields; A as well as ϕ itself have dimension one and thus E has dimension two. A renormalizable action can be at most quadratic in E and, as in the 1-D case, E does not propagate and can be eliminated from the action by using the corresponding equation of motion.

The action of the supersymmetry generator $\bar{\eta}^\alpha\bar{\mathrm{Q}}_\alpha - \eta_\alpha\mathrm{Q}^\alpha$ in component form is then

$$\delta A = \eta_\alpha\psi^\alpha\,, \tag{17.104a}$$

$$\delta\psi^\beta = \bar{\eta}^\alpha\sigma_{\mu\alpha}^{\ \ \beta}\partial_\mu A - \eta^\beta E\,, \tag{17.104b}$$

$$\delta E = -\bar{\eta}^\alpha\sigma_{\mu\alpha}^{\ \ \gamma}\partial_\mu\psi_\gamma\,. \tag{17.104c}$$

For later purposes it is useful to also expand ϕ in terms of the space coordinate x:

$$\begin{aligned}\phi = A(x) + \psi^{\alpha}(x)\theta_{\alpha} + \tfrac{1}{2}\partial_{\mu}A(x)\bar{\theta}^{\alpha}\sigma_{\mu\alpha}^{\ \ \beta}\theta_{\beta} + \tfrac{1}{2}E(x)\theta^{\alpha}\theta_{\alpha} \\ - \tfrac{1}{4}\bar{\theta}^{\alpha}\sigma_{\mu\alpha}^{\ \ \beta}\partial_{\mu}\psi_{\beta}(x)\theta^{\gamma}\theta_{\gamma} + \tfrac{1}{16}\partial^{2}A(x)\theta^{\alpha}\theta_{\alpha}\bar{\theta}^{\beta}\bar{\theta}_{\beta}\,.\end{aligned} \tag{17.105}$$

In the same way left-handed chiral superfields can be defined which satisfy $\mathrm{D}^{\alpha}\bar{\phi} = 0$ and thus depend on $\bar{\theta}$ and a space variable $\bar{y}$,

$$\bar{y}_{\mu} = x_{\mu} - \tfrac{1}{2}\bar{\theta}^{\alpha}\sigma_{\mu\alpha}^{\ \ \beta}\theta_{\beta}, \qquad \bar{\mathrm{Q}}_{\alpha}\bar{y}_{\mu} = 0\,. \tag{17.106}$$

The corresponding chiral field can be written as

$$\bar{\phi}(\bar{y},\bar{\theta}) = A^{*}(\bar{y}) + \bar{\psi}^{\alpha}(\bar{y})\bar{\theta}_{\alpha} + \tfrac{1}{2}E^{*}(\bar{y})\bar{\theta}^{\alpha}\bar{\theta}_{\alpha}\,. \tag{17.107}$$

The action of the supersymmetry generators is then

$$\delta A^{*} = \bar{\eta}^{\alpha}\bar{\psi}_{\alpha}\,, \tag{17.108a}$$
$$\delta\bar{\psi}_{\beta} = \sigma_{\mu\beta}^{\ \ \alpha}\eta_{\alpha}\partial_{\mu}A^{*} + \bar{\eta}_{\beta}E^{*}\,, \tag{17.108b}$$
$$\delta E^{*} = -\partial_{\mu}\bar{\psi}^{\gamma}\sigma_{\mu\gamma}^{\ \ \alpha}\eta_{\alpha}\,. \tag{17.108c}$$

The expansion of $\bar{\phi}$ in terms of the space variable x reads as

$$\begin{aligned}\bar{\phi} = A^{*}(x) + \bar{\psi}^{\alpha}(x)\bar{\theta}_{\alpha} - \tfrac{1}{2}\partial_{\mu}A^{*}(x)\bar{\theta}^{\alpha}\sigma_{\mu\alpha}^{\ \ \beta}\theta_{\beta} + \tfrac{1}{2}E^{*}(x)\bar{\theta}^{\alpha}\bar{\theta}_{\alpha} \\ + \tfrac{1}{4}\bar{\partial}_{\mu}\psi^{\alpha}(x)\sigma_{\mu\alpha}^{\ \ \beta}\theta_{\beta}\bar{\theta}^{\gamma}\bar{\theta}_{\gamma} + \tfrac{1}{16}\partial^{2}A(x)\theta^{\alpha}\theta_{\alpha}\bar{\theta}^{\beta}\bar{\theta}_{\beta}\,.\end{aligned} \tag{17.109}$$

The free action. We can now construct a free action

$$\begin{aligned}\mathcal{L}_{D} &= \phi(\bar{y},\bar{\theta})\phi(y,\theta), \\ \int \mathrm{d}^{4}x \prod_{\alpha} \mathrm{d}\bar{\theta}_{\alpha}\mathrm{d}\theta_{\alpha}\,\mathcal{L}_{D} &= \int \mathrm{d}^{4}x\left(\partial_{\mu}A^{*}\partial_{\mu}A - \bar{\psi}^{\alpha}\sigma_{\mu\alpha}^{\ \ \beta}\partial_{\mu}\psi_{\beta} - E^{*}E\right).\end{aligned} \tag{17.110}$$

The supersymmetry of the lagrangian $\mathcal{L}_{D}$ is obvious here since it does not depend explicitly on $\theta, \bar{\theta}$ and x_{μ}. If we have several superfields we simply add the corresponding contributions.

Interaction terms. We here 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:

$$\mathcal{L}_{F}(\phi) = \bar{c}_{i}\phi_{i} + \tfrac{1}{2}M_{ij}\phi_{i}\phi_{j} + \tfrac{1}{3!}g_{ijk}\phi_{i}\phi_{j}\phi_{k}\,. \tag{17.111}$$

Integrating over θ_1, θ_2, we obtain

$$\int \mathrm{i}\mathrm{d}\theta_{2}\mathrm{d}\theta_{1}\,\mathcal{L}_{F} = \bar{c}_{i}E_{i} + M_{ij}\left(A_{i}E_{j} - \tfrac{1}{2}\psi_{i}^{\alpha}\psi_{\alpha j}\right) + \tfrac{1}{2}g_{ijk}\left(A_{i}A_{j}E_{k} - \psi_{i}^{\alpha}\psi_{\alpha j}A_{k}\right). \tag{17.112}$$

Adding the kinetic term (17.110), (17.112) and its conjugated left-handed contribution

$$\begin{aligned}\bar{\mathcal{L}}_{F}(\bar{\phi}) &= c_{i}^{*}\bar{\phi}_{i} + \tfrac{1}{2}M_{ij}^{*}\bar{\phi}_{i}\bar{\phi}_{j} + \tfrac{1}{3!}g_{ijk}^{*}\bar{\phi}_{i}\bar{\phi}_{j}\bar{\phi}_{k}\,, \\ \int \mathrm{i}\mathrm{d}\bar{\theta}_{2}\mathrm{d}\bar{\theta}_{1}\,\bar{\mathcal{L}}_{F} &= c_{i}^{*}E_{i}^{*} + M_{ij}^{*}\left(A_{i}^{*}E_{j}^{*} - \tfrac{1}{2}\bar{\psi}_{i}^{\alpha}\bar{\psi}_{\alpha j}\right) + \tfrac{1}{2}g_{ijk}^{*}\left(A_{i}^{*}A_{j}^{*}E_{k}^{*} - \bar{\psi}_{i}^{\alpha}\bar{\psi}_{\alpha j}A_{k}^{*}\right).\end{aligned} \tag{17.113}$$

we obtain a physically sensible supersymmetric lagrangian. A useful exercise is to verify, using the explicit expressions (17.104) and (17.108), 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:

$$\frac{\partial \mathcal{L}}{\partial E_i} = E_i^* + \bar{c}_i + M_{ij} A_j + \tfrac{1}{2} g_{ijk} A_j A_k = 0\,, \tag{17.114a}$$

$$\frac{\partial \mathcal{L}}{\partial E_i^*} = E_i + c_i^* + M_{ij}^* A_j^* + \tfrac{1}{2}{}^* g_{ijk} A_j^* A_k^* = 0\,. \tag{17.114b}$$

The lagrangian can then be cast into the form:

$$\begin{aligned}\mathcal{L} = \partial_\mu A^* \partial_\mu A - \bar{\psi}^\alpha \sigma_{\mu\dot{\alpha}}^{\ \beta} \partial_\mu \psi_\beta + \frac{\partial \mathcal{L}_F}{\partial \phi_i}(A) \frac{\partial \bar{\mathcal{L}}_F}{\partial \phi_i}(A^*) \\ - \frac{1}{2} \frac{\partial^2 \mathcal{L}_F}{\partial \phi_i \partial \phi_j}(A) \psi_i^\alpha \psi_{\alpha j} - \frac{1}{2} \frac{\partial^2 \bar{\mathcal{L}}_F}{\partial \bar{\phi}_i \bar{\partial} \phi_j}(\bar{A}) \bar{\psi}_i^\alpha \bar{\psi}_{\alpha j}\,,\end{aligned}$$

a form reminiscent of expressions (16.113) and (17.88).

17.6.2 Vector superfields

Up to now we have explained how to construct a supersymmetric action containing only scalar and spin 1/2 fermion fields. For a realistic theory of particles, vector fields are also required. Therefore, from now on, we again assume some of the concepts which will be introduced only in Chapters 18,19.

We consider a general real superfield, which we call a vector superfield. We can parametrize it in the form:

$$\begin{aligned}V = B + \chi^\alpha \theta_\alpha + \bar{\chi}^\alpha \bar{\theta}_\alpha + \tfrac{1}{2} C \theta^\alpha \theta_\alpha + \tfrac{1}{2} C^* \bar{\theta}^\alpha \bar{\theta}_\alpha + i 2^{-1/2} V_\mu \bar{\theta}^\alpha \sigma_{\mu\dot{\alpha}}^{\ \beta} \theta_\beta \\ + \tfrac{1}{2} \bar{\theta}^\alpha \left(\bar{\lambda}_\alpha - \tfrac{1}{2} \sigma_{\mu\dot{\alpha}}^{\ \beta} \partial_\mu \chi_\beta \right) \theta^\gamma \theta_\gamma + \tfrac{1}{2} \left(\lambda^\beta + \tfrac{1}{2} \partial_\mu \bar{\chi}^\alpha(x) \sigma_{\mu\dot{\alpha}}^{\ \beta} \right) \theta_\beta \bar{\theta}^\gamma \bar{\theta}_\gamma \\ + \tfrac{1}{4} \left(K + \tfrac{1}{4} \partial^2 B(x) \right) \theta^\alpha \theta_\alpha \bar{\theta}^\beta \bar{\theta}_\beta\,,\end{aligned} \tag{17.115}$$

in which B, K are real fields. The reason for such a parametrization will become clearer below. We first note that in this form the vector superfield contains four real scalar fields, four Weyl spinors and one vector field. However, supersymmetry naturally leads to a generalized form of gauge symmetry. We consider two scalar superfields $\bar{\phi}, \phi$ of the form (17.103,17.106). A generalized abelian supersymmetric gauge transformation is then defined by the translation

$$V \mapsto V + \phi + \bar{\phi}\,. \tag{17.116}$$

In component form, this leads to

$$\begin{aligned} &B \mapsto B + A + A^*, && \chi \mapsto \chi + \psi\,, \\ &C \mapsto C + E\,, && V_\mu \mapsto V_\mu - 2^{-1/2} i \partial_\mu (A - A^*)\,, \\ &\lambda \mapsto \lambda\,, && K \mapsto K\,. \end{aligned}$$

With our parametrization λ and K are gauge invariant, while V_μ transforms like a usual abelian gauge field.

In a theory which has this kind of gauge invariance, the vector superfield can be simplified, the fields B, C, χ being eliminated by a gauge transformation. This characterizes the non-supersymmetric *Wess–Zumino (WZ) gauge.* The vector field then takes the form,

$$V = i2^{-1/2} V_\mu \bar{\theta}^\alpha \sigma_{\mu\alpha}^{\ \ \beta} \theta_\beta + \tfrac{1}{2}\bar{\theta}^\alpha \bar{\lambda}_\alpha \theta^\beta \theta_\beta + \tfrac{1}{2}\lambda^\beta \theta_\beta \bar{\theta}^\alpha \bar{\theta}_\alpha + \tfrac{1}{4} K \theta^\alpha \theta_\alpha \bar{\theta}^\beta \bar{\theta}_\beta \,. \tag{17.117}$$

The physical degrees of freedom, as we shall see, reduce to a massless vector field V_μ and a massless spin $1/2$ fermion λ, since the field K does not propagate.

It is important to note that, in this gauge, powers of the vector field have the property

$$V^2 = -\tfrac{1}{4} V_\mu V_\mu \theta^\alpha \theta_\alpha \bar{\theta}^\beta \bar{\theta}_\beta \,, \qquad V^n = 0 \text{ for } n \geq 3 \,. \tag{17.118}$$

Supersymmetric curvature tensor. The vector superfield is not chiral and, therefore, as in the 1-D case, the operators $\bar{\mathrm{D}}, \mathrm{D}$ can be used to construct other superfields. Note that quantities of the form $\bar{\mathrm{D}}^\beta \bar{\mathrm{D}}_\beta W$ and $\mathrm{D}^\beta \mathrm{D}_\beta W$, where W is an arbitrary superfield, are right- and left-handed chiral fields, respectively, because the product of three operators D or $\bar{\mathrm{D}}$ vanishes:

$$\bar{\mathrm{D}}_\alpha \bar{\mathrm{D}}^\beta \bar{\mathrm{D}}_\beta W = \mathrm{D}^\alpha \mathrm{D}^\beta \mathrm{D}_\beta W = 0 \quad \forall\, W \,.$$

In particular, the quantities

$$F^\alpha = \tfrac{1}{2} \bar{\mathrm{D}}^\beta \bar{\mathrm{D}}_\beta \mathrm{D}^\alpha V \,, \qquad \bar{F}_\alpha = \tfrac{1}{2} \mathrm{D}^\beta \mathrm{D}_\beta \bar{\mathrm{D}}_\alpha V \,, \tag{17.119}$$

are chiral and gauge invariant (in the sense (17.116)) since for example

$$\bar{\mathrm{D}}^\beta \bar{\mathrm{D}}_\beta \mathrm{D}^\alpha (\phi + \bar{\phi}) = \bar{\mathrm{D}}^\beta \bar{\mathrm{D}}_\beta \mathrm{D}^\alpha \phi = -\sigma_{\mu\beta}^{\ \ \alpha} \partial_\mu \bar{\mathrm{D}}^\beta \phi = 0 \,,$$

(we have used the anticommutation relation (17.99) 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 $\bar{y}$:

$$\begin{aligned} V &= i2^{-1/2} V_\mu(y) \bar{\theta}^\alpha \sigma_{\mu\alpha}^{\ \ \beta} \theta_\beta + \tfrac{1}{2}\bar{\theta}^\alpha \bar{\lambda}_\alpha \theta^\beta \theta_\beta + \tfrac{1}{2}\lambda^\beta(y) \theta_\beta \bar{\theta}^\alpha \bar{\theta}_\alpha \\ &\quad + \tfrac{1}{4}\left[K(y) - i\partial_\mu V_\mu(y)\right] \theta^\alpha \theta_\alpha \bar{\theta}^\beta \bar{\theta}_\beta \,, \\ &= i2^{-1/2} V_\mu(\bar{y}) \bar{\theta}^\alpha \sigma_{\mu\alpha}^{\ \ \beta} \theta_\beta + \tfrac{1}{2}\bar{\theta}^\alpha \bar{\lambda}_\alpha \theta^\beta \theta_\beta + \tfrac{1}{2}\lambda^\beta(\bar{y}) \theta_\beta \bar{\theta}^\alpha \bar{\theta}_\alpha \\ &\quad + \tfrac{1}{4}\left[K(\bar{y}) + i\partial_\mu V_\mu(\bar{y})\right] \theta^\alpha \theta_\alpha \bar{\theta}^\beta \bar{\theta}_\beta \,. \end{aligned}$$

We then find

$$\begin{aligned} F^\alpha &= -\lambda^\alpha(y) - i2^{-3/2} \theta^\beta \left(\sigma_\nu^\dagger \sigma_\mu\right)_\beta^{\ \alpha} F_{\mu\nu}(y) - K(y)\theta^\alpha + \tfrac{1}{2}\partial_\mu \bar{\lambda}^\beta(y) \sigma_{\mu\beta}^{\ \ \alpha} \theta^\gamma \theta_\gamma \,, \\ \bar{F}_\alpha &= \bar{\lambda}_\alpha(\bar{y}) - i2^{-3/2} \left(\sigma_\mu \sigma_\nu^\dagger\right)_\alpha^{\ \beta} F_{\mu\nu}(\bar{y}) \bar{\theta}_\beta + K(\bar{y})\bar{\theta}_\alpha + \tfrac{1}{2}\sigma_{\mu\alpha}^{\ \ \beta} \partial_\mu \lambda_\beta(\bar{y}) \bar{\theta}^\gamma \bar{\theta}_\gamma \end{aligned}$$

with the usual notation $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. We observe that indeed F and $\bar{F}$ depend only on the gauge invariant combinations $F_{\mu\nu}, \lambda, K$. A check of these expressions is provided by the simple relation (derived from the definitions (17.119) after a few commutations)

$$\mathrm{D}_\alpha F^\alpha + \bar{\mathrm{D}}^\alpha \bar{F}_\alpha = 0 \,.$$

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,

$$\int \mathrm{d}\theta_1 \mathrm{d}\theta^1\, F^\alpha F_\alpha = 2\bar{\lambda}^\alpha \sigma_{\mu\alpha}{}^{\beta}\partial_\mu \lambda_\beta + K^2 - \tfrac{1}{2}F_{\mu\nu}F_{\mu\nu} + \tfrac{1}{4}\epsilon_{\lambda\mu\nu\rho}F_{\lambda\mu}F_{\nu\rho}\,, \tag{17.120}$$

where we have used the identity

$$\tfrac{1}{2}\operatorname{tr}\sigma^\dagger_\lambda \sigma_\mu \sigma^\dagger_\nu \sigma_\rho = \epsilon_{\lambda\mu\nu\rho} + (\delta_{\lambda\mu}\delta_{\nu\rho} + \delta_{\mu\nu}\delta_{\rho\lambda} - \delta_{\lambda\nu}\delta_{\mu\rho})\,.$$

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

$$\mathcal{S} = \int \mathrm{d}^4 x\,\left[\tfrac{1}{4}F^2_{\mu\nu} - \bar{\lambda}^\alpha \sigma_{\mu\alpha}{}^{\beta}\partial_\mu\lambda_\beta - K^2\right]. \tag{17.121}$$

This action could also have been obtained by noting that the vector superfield has dimension zero. A term of the form $V\bar{\mathrm{D}}\bar{\mathrm{D}}\mathrm{D}\mathrm{D}V$ has dimension 2 and is thus candidate to be a D term.

Note finally that it is also possible to give a mass to a vector superfield by adding the D contribution of V^2.

Gauge invariant interactions. A charged scalar superfield transforms under global $U(1)$ as $\phi \mapsto \mathrm{e}^{ie\Lambda}\,\phi$. However, if we want to introduce space-dependent $U(1)$ transformations consistent with supersymmetry we cannot simply replace Λ by $\Lambda(x)$. We have to introduce the constant Λ by a scalar chiral superfield:

$$\phi(y,\theta) \mapsto \mathrm{e}^{ie\Lambda(y,\theta)}\,\phi(y,\theta) \qquad \text{with} \quad \bar{\mathrm{D}}_\alpha \Lambda = 0\,, \tag{17.122a}$$

$$\bar{\phi}(\bar{y},\bar{\theta}) \mapsto \mathrm{e}^{-ie\bar{\Lambda}(\bar{y},\bar{\theta})}\,\bar{\phi}(\bar{y},\bar{\theta}) \qquad \text{with} \quad \mathrm{D}^\alpha \bar{\Lambda} = 0\,. \tag{17.122b}$$

We see immediately that if the charges and couplings are such that the interaction term (17.111) is invariant under global $U(1)$ transformations, it is also invariant under the transformations (17.122). However, the free term (17.110) is not invariant since

$$\bar{\phi}\phi \mapsto \bar{\phi}\phi\,\mathrm{e}^{ie(\Lambda-\bar{\Lambda})}\,.$$

To render the kinetic term gauge invariant we replace it by

$$\mathcal{L}_D = \bar{\phi}\,\mathrm{e}^{eV}\,\phi\,,$$

where the vector superfield V must transform like

$$V \mapsto V - i(\Lambda - \bar{\Lambda}),$$

transformation indeed consistent with (17.116).

At first sight it would seem that such a theory cannot be renormalizable, but at least in the WZ gauge, the expansion of the exponential reduces to three first terms (property (17.118)) and contains no term of dimension larger than four.

We have thus presented all the ingredients necessary to construct a supersymmetric version of QED.

Non-abelian gauge theories. The extension to non-abelian gauge theories is not difficult. 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

$$\phi \mapsto \mathrm{e}^{\Lambda}\,\phi\,, \quad \Lambda \equiv \Lambda_a t_a\,, \qquad \bar{\phi} \mapsto \bar{\phi}\,\mathrm{e}^{\bar{\Lambda}}, \quad \bar{\Lambda} \equiv \bar{\Lambda}_a t_a\,.$$

A gauge invariant lagrangian is

$$\mathcal{L}_D = \bar{\phi}\,\mathrm{e}^{V}\,\phi\,, \quad \text{with } V = V_a t_a\,,$$

where the gauge field transforms like

$$\mathrm{e}^{V} \mapsto \mathrm{e}^{-\bar{\Lambda}}\,\mathrm{e}^{V}\,\mathrm{e}^{-\Lambda}\,.$$

The generalized form of the curvature F^{α} is

$$F^{\alpha} = \bar{\mathrm{D}}^{\beta}\bar{\mathrm{D}}_{\beta}\,\mathrm{e}^{-V}\,\mathrm{D}^{\alpha}\,\mathrm{e}^{V}\,, \tag{17.123}$$

which is chiral and indeed transforms like

$$F^{\alpha} \mapsto \mathrm{e}^{\Lambda}\,F^{\alpha}\,\mathrm{e}^{-\Lambda}\,. \tag{17.124}$$

Finally, the contribution to the action takes the form $\operatorname{tr} F^{\alpha}F_{\alpha}$.

Bibliographical Notes

In the discussion of the Langevin equation we follow

J. Zinn-Justin, *Nucl. Phys.* B275 [FS18] (1986) 135.

The relation between supersymmetry and dissipative Langevin or Fokker–Planck equations have been shown in

E. Witten, *Nucl. Phys.* B188 (1981) 513; M.V. Feigel'man and A.M. Tsvelik, *Sov. Phys.-JETP* 56 (1982) 823; H. Nakazato, M. Namiki, I. Ohba and K. Okano, *Prog. Theor. Phys.* 70 (1983) 298; E. Egorian and S. Kalitzin, *Phys. Lett.* 129B (1983) 320.

Our construction of the dynamic action provides a first example of the so-called Nicolai map, which appears in supersymmetric theories

H. Nicolai, *Phys. Lett.* 89B (1980) 341; *Nucl. Phys.* B176 (1980) 419.

The characterization of the equilibrium distribution by supersymmetry arguments is adapted from

R. Kirschner, *Phys. Lett.* 139B (1984) 180; E. Gozzi, *Phys. Lett.* 143B (1984) 183.

The non-linear σ-model has been considered in

R. Bausch, H.K. Janssen and Y. Yamazaki, *Z. Phys.* B37 (1980) 163.

Supersymmetry has been introduced in Field Theory context by

J. Wess and B. Zumino, *Nucl. Phys.* B70 (1974) 39.

Anticommuting coordinates were used in

A. Salam and J. Strathdee, *Phys. Lett.* 51B (1974) 353; S. Ferrara, B. Zumino and J. Wess, *Phys. Lett.* 51B (1974) 239.

In Section 17.6 we have followed

J. Wess and J. Bagger, *Supersymmetry and Supergravity* (Princeton University Press, Princeton 1983).

Numerous other textbooks have been devoted to supersymmetry see, for example,

S. Ferrara, *Supersymmetry and Supergravity (Lecture Notes in Physics 208)*, (Springer-Verlag, Berlin 1984); B.S. DeWitt, *Supermanifolds* (Cambridge University Press, Cambridge 1984).

Many articles are reprinted in

Supersymmetry, S. Ferrara ed. (North-Holland and World Scientific 1987).

The supersymmetric $O(N)$ non linear σ model has been described in

E. Witten, *Phys. Rev.* D16 (1977) 2991.

The study of general two-dimensional supersymmetric non-linear σ models has been initiated in

D.Z. Freedman, P.K. Townsend, *Nucl. Phys.* B177 (1981) 282,

and UV properties first discussed in

L. Alvarez-Gaume, D.Z. Freedman, *Commun. Math. Phys.* 80 (1981) 443.

The general four-loop β-function is given in

Marcus T. Grisaru, A.E.M. van de Ven and D. Zanon, *Nucl. Phys.* B277 (1986) 409, *Phys. Lett.* B173 (1986) 423.

RG-β functions in supersymmetric theories are related to instanton calculus in

V.A. Novikov, M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *Nucl. Phys.* B229 (1983) 381; A.Yu. Morozov, A.M. Perelomov and M.A. Shifman, *Nucl. Phys.* B248 (1984) 279.

The supersymmetry and the dimensional reduction in the random-field Ising model are discussed in

G. Parisi and N. Sourlas, *Phys. Rev. Lett.* 43 (1979) 744; J. Cardy, *Phys. Lett.* 125B (1983) 470; A. Klein and J. Fernando-Perez, *Phys. Lett.* 125B (1983) 473.

The consequences of dimensional reduction contradict physical intuition

Y. Imry and S.K. Ma, *Phys. Rev. Lett.* 35 (1975)

and are disproved in

J. Imbrie, *Commun. Math. Phys.* 98 (1985) 145; J. Bricmont and A. Kupiainen, *Phys. Rev. Lett.* 59 (1987) 1829.

APPENDIX A17
BRS AND SUPERSYMMETRY. THE RANDOM FIELD ISING MODEL

A17.1 Extension of BRS Symmetries: Supersymmetry

A trivial extension of BRS transformation can be constructed by considering a function of several Grassmann variables $f(\theta_1, \ldots, \theta_n)$ 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:

$$\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j} + \frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_i} = 0 . \tag{A17.1}$$

Translations on Grassmann variables form the equivalent of usual abelian Lie groups; non-trivial extensions, "non-abelian", will correspond to non-vanishing anticommutators.

First a few general remarks are in order:

(i) If Q is the unique generator of this generalized structure called "supergroup", then exponentiation implies

$$\mathrm{e}^{\varepsilon_1 \mathrm{Q}}\, \mathrm{e}^{\varepsilon_2 \mathrm{Q}} = \mathrm{e}^{(\varepsilon_1+\varepsilon_2)\mathrm{Q}} ,$$

in which ε_1 and ε_2 are Grassmann variables. Expanding both sides we obtain the condition

$$-\varepsilon_1\varepsilon_2 \mathrm{Q}^2 = 0 \Rightarrow \mathrm{Q}^2 = 0 , \tag{A17.2}$$

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 call Q_i, L_a the anticommuting and commuting elements respectively, the general structure of a Lie superalgebra is

$$\{\mathrm{Q}_i, \mathrm{Q}_j\} = c_{ija} L_a , \quad [\mathrm{Q}_i, L_a] = d_{iaj}\mathrm{Q}_j , \quad [L_a, L_b] = f_{abc} L_c .$$

(iii) Triplets of generators thus satisfy a mixed Jacobi identity of the form:

$$[\mathrm{Q}_i, (\mathrm{Q}_j\mathrm{Q}_k + \mathrm{Q}_k\mathrm{Q}_j)] + \text{cyclic permutations } (ijk) = 0 . \tag{A17.3}$$

Many aspects of the theory 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 shall construct only a few simple examples.

Supersymmetry and time translation. We want to realize our algebra under the form of differential operators. From the previous considerations we conclude that we need at least two Grassmann variables $\bar\theta$ and θ and a usual 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 ($A17.2$) we find two non-trivial combinations:

$$\mathrm{Q} = \frac{\partial}{\partial\theta} + a\bar{\theta}\frac{\partial}{\partial t}, \qquad \bar{\mathrm{Q}} = \frac{\partial}{\partial\bar{\theta}} + \bar{a}\theta\frac{\partial}{\partial t}, \tag{A17.4}$$

in which a and $\bar{a}$ are two ordinary arbitrary constants. It follows that

$$\mathrm{Q}\bar{\mathrm{Q}} + \bar{\mathrm{Q}}\mathrm{Q} = (\bar{a} + a)\frac{\partial}{\partial t}. \tag{A17.5}$$

The Lie subgroup is the group of translations on the variable t. This is a realization of quantum mechanics supersymmetry as discussed in Section 17.1.

This structure can immediately be generalized to d-dimensional space:

$$\mathrm{Q}_\alpha = \frac{\partial}{\partial\theta_\alpha} + a^\mu_{\alpha\beta}\bar{\theta}_\beta\frac{\partial}{\partial x_\mu}, \qquad \bar{\mathrm{Q}}_\alpha = \frac{\partial}{\partial\bar{\theta}_\alpha} + \theta_\beta\bar{a}^\mu_{\beta\alpha}\frac{\partial}{\partial x_\mu}, \tag{A17.6}$$

in which x_μ corresponds to the d commuting variables, $\bar{\theta}_\alpha$ and θ_α to anticommuting variables and $a^\mu_{\alpha\beta}$, $\bar{a}^\mu_{\alpha\beta}$ are constants. Then,

$$\begin{cases} \mathrm{Q}_\alpha\mathrm{Q}_\beta + \mathrm{Q}_\beta\mathrm{Q}_\alpha = 0\,, \\ \bar{\mathrm{Q}}_\alpha\bar{\mathrm{Q}}_\beta + \bar{\mathrm{Q}}_\beta\bar{\mathrm{Q}}_\alpha = 0\,, \\ \mathrm{Q}_\alpha\bar{\mathrm{Q}}_\beta + \bar{\mathrm{Q}}_\beta\mathrm{Q}_\alpha = (\bar{a}^\mu_{\alpha\beta} + a^\mu_{\alpha\beta})\dfrac{\partial}{\partial x_\mu}. \end{cases} \tag{A17.7}$$

Again the Lie subgroup is a product of translation groups. Note that to $\bar{\mathrm{Q}}, \mathrm{Q}$ correspond two other supersymmetry generators $\bar{\mathrm{D}}, \mathrm{D}$,

$$\mathrm{D}_\alpha = \frac{\partial}{\partial\theta_\alpha} - \bar{a}^\mu_{\alpha\beta}\bar{\theta}_\beta\frac{\partial}{\partial x_\mu}, \quad \bar{\mathrm{D}}_\alpha = \frac{\partial}{\partial\bar{\theta}_\alpha} - \bar{\theta}_\beta\bar{a}^\mu_{\beta\alpha}\frac{\partial}{\partial x_\mu}, \tag{A17.8}$$

which anticommute with $\bar{\mathrm{Q}}, \mathrm{Q}$ and thus can play the role of covariant derivatives.

A17.2 Supersymmetry: The Random Field Ising Model

It can be shown that the long distance properties of an Ising model in a random magnetic field can be described by a stochastic field equation of the form:

$$(-\nabla^2 + m^2)\varphi(x) + \frac{g}{3!}\varphi^3(x) = h(x), \tag{A17.9}$$

in which the field $\varphi(x)$ represents the Ising spin and $h(x)$, the magnetic field, has a gaussian distribution,

$$\langle h(x)h(x')\rangle = \alpha\delta(x - x'), \quad \langle h(x)\rangle = 0\,. \tag{A17.10}$$

Equation ($A17.9$) is an example of a gradient equation of the form (16.34) and, therefore, leads to two BRS symmetries. Introducing the superfield $\phi(x, \bar{\theta}, \theta)$, we can write the associated action $\mathcal{S}(\phi)$, after averaging over the magnetic field, as

$$\mathcal{S}(\phi) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta \left(\frac{\alpha}{2}\int \mathrm{d}^d x \frac{\partial\phi}{\partial\bar{\theta}}\frac{\partial\phi}{\partial\theta} + \mathcal{A}(\phi)\right) \tag{A17.11}$$

with

$$\mathcal{A}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{g}{4!}\phi^4\right]. \tag{A17.12}$$

The remarkable property of the action (A17.11) is that it has a larger symmetry group: translations of $\bar{\theta}$ and θ as already known but also "rotations" which leave the line element:

$$(\mathrm{d}s)^2 = \mathrm{d}x_\mu \mathrm{d}x_\mu + \alpha\, \mathrm{d}\bar{\theta}\, \mathrm{d}\theta \tag{A17.13}$$

invariant. In addition to transformations internal to the $\{\bar{\theta}, \theta\}$ and x_μ spaces one finds the two infinitesimal transformations,

$$\theta = \theta' + \alpha a_\mu x_\mu\,, \qquad x_\mu = x'_\mu - 2a_\mu \bar{\theta}', \tag{A17.14}$$

$$\bar{\theta} = \bar{\theta}' + \alpha \bar{a}_\mu x_\mu\,, \qquad x_\mu = x'_\mu - 2\theta' \bar{a}_\mu\,. \tag{A17.15}$$

The vectors a_μ and $\bar{a}_\mu$ are anticommuting elements of the Grassmann algebra. If we call Q_μ and $\bar{\mathrm{Q}}_\mu$ the corresponding generators, we find the anticommutation relations:

$$\mathrm{Q}_\mu \mathrm{Q}_\nu + \mathrm{Q}_\nu \mathrm{Q}_\mu = 4\alpha \delta_{\mu\nu}\, \theta \frac{\partial}{\partial \bar{\theta}}\,, \tag{A17.16}$$

$$\bar{\mathrm{Q}}_\mu \bar{\mathrm{Q}}_\nu + \bar{\mathrm{Q}}_\nu \bar{\mathrm{Q}}_\mu = -4\alpha \delta_{\mu\nu}\, \bar{\theta} \frac{\partial}{\partial \theta}\,, \tag{A17.17}$$

$$\bar{\mathrm{Q}}_\mu \mathrm{Q}_\nu + \mathrm{Q}_\nu \bar{\mathrm{Q}}_\mu = 2\alpha \left[\delta_{\mu\nu}(\theta \frac{\partial}{\partial \theta} - \bar{\theta}\frac{\partial}{\partial \bar{\theta}}) + x_\mu \partial_\nu - x_\mu \partial_\nu\right]. \tag{A17.18}$$

We recognize in the r.h.s. generators of internal transformations.

Using this symmetry it is possible to prove a property of dimensional reduction for action (A17.11) by a variant of the method of Section 17.2. One shows that the correlation functions calculated in d dimensions with the action (A17.11) 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. It has unfortunately consequences which contradict physical intuition. A reason can be found in the starting point: equation (A17.9) in the region of interest, m^2 small, has certainly for some fields a large number of solutions and, therefore, the whole method is not applicable without modifications. A similar problem, called Gribov's ambiguity, arises in gauge theories.

18 ABELIAN GAUGE THEORIES

With this chapter we begin the study of a new class of geometric models that have played a central role in the theory of fundamental interactions for more than a century: gauge theories. They are characterized by new physical properties and new technical difficulties. Gauge theories can be considered in some sense as geometric generalizations of the principal chiral model. In particular, the pure abelian gauge field leads to a free field theory, as the pure abelian chiral model.

We devote first a whole chapter to a simple and physically important example, the abelian gauge theory, whose physical realization is Quantum Electrodynamics (QED). However, since many excellent textbooks deal extensively with QED, we will mainly concentrate on the formal aspects of abelian gauge theories.

The set-up of this chapter is the following: We begin with elementary considerations about the massive vector field in perturbation theory. We show that coupling to 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 the massless vector limit can be defined. The corresponding field theories are gauge invariant. We then discuss the specific properties of gauge invariant theories and mention the IR problem of physical observables. We quantize gauge theories starting directly from first principles. The formal equivalence between different gauges is established.

In Section 18.5 regularization methods are presented which allow overcoming the new difficulties one encounters in gauge theories. The abelian gauge symmetry, broken by gauge fixing terms, then leads to a set of WT identities which are used to prove the renormalizability of the theory. The gauge dependence of correlation functions in a set of covariant gauges is determined. Renormalization group equations follow and we calculate the RG β-function at leading order.

As an introduction to the next chapter, we analyse the abelian Higgs mechanism. Finally, we comment about stochastic quantization of gauge theories.

The appendix contains the derivation of the Casimir effect, some more details about gauge dependence and the calculation of one-loop divergences with Schwinger's representation.

18.1 The Massive Vector Field

The quantization of the free massive vector field does not immediately follow from the quantization of the scalar field and thus requires a short discussion. In the first part of this section we work in real time with the metric $\{+,-,-,\cdots\}$ where the first component is the time component. Space-time is denoted by $\{t \equiv x_0 = -ix_d, x_i\}$, $x_i \in \mathbb{R}^{d-1}$, and the vector field $A_\mu \equiv \{A_0, A_i\}$. Summation over repeated lower and upper indices is implied. Finally, time derivative is occasionally indicated by

$$\frac{\partial A_i}{\partial t} \equiv \dot{A}_i \,.$$

The local $O(1,d-1)$ invariant classical action for a free massive vector field then can be written as

$$\mathcal{A}(A) = -\int \mathrm{d}t\,\mathrm{d}^{d-1}x\,\left[\tfrac{1}{4}F_{\mu\nu}(t,x)F^{\mu\nu}(t,x) - \tfrac{1}{2}m^2 A_\mu(t,x)A^\mu(t,x)\right] \tag{18.1}$$

with

$$F_{\mu\nu}(t,x) = \partial_\mu A_\nu(t,x) - \partial_\nu A_\mu(t,x). \tag{18.2}$$

One may wonder about the peculiar form of this derivative term, but it is straightforward to verify that the additional covariant term one could think of adding, $\partial_\mu A_\mu \partial^\nu A^\nu$, depending on its sign corresponds either to a A_0 field with a negative metric, or to an unbounded potential.

Quantization. The action (18.1) 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

$$\begin{aligned}\mathcal{A}(A) = \int \mathrm{d}t\,\mathrm{d}^{d-1}x\,&\left[\frac{1}{2}\dot{A}_i(t,x)\left(\delta_{ij} - \frac{\partial_i\partial_j}{\partial_\perp^2 - m^2}\right)\dot{A}_j(t,x) - \frac{1}{2}m^2 A_i^2(x) - \frac{1}{4}F_{ij}^2(t,x)\right.\\ &+\frac{1}{2}\left((-\partial_\perp^2 + m^2)A_0(t,x) + \partial_i\dot{A}_i(t,x)\right)(-\partial_\perp^2 + m^2)^{-1}\\ &\left.\times\left((-\partial_\perp^2 + m^2)A_0(t,x) + \partial_i\dot{A}_i(t,x)\right)\right],\end{aligned} \tag{18.3}$$

where roman indices denote space components, and $\partial_\perp^2$ is the space laplacian.

Therefore, A_0 is not a dynamical degree of freedom and the corresponding field equation

$$\frac{\delta\mathcal{A}}{\delta A_0(t,x)} = \left(-\partial_\perp^2 + m^2\right)A_0(t,x) + \partial_i\dot{A}_i(t,x) = 0\,, \tag{18.4}$$

is a constraint equation that 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. The reduced lagrangian density then takes the form:

$$\mathcal{L}(A_i) = \frac{1}{2}\dot{A}_i(t,x)\left(\delta_{ij} - \frac{\partial_i\partial_j}{\partial_\perp^2 - m^2}\right)\dot{A}_j(t,x) - \frac{1}{2}m^2 A_i^2(t,x) - \frac{1}{4}F_{ij}^2(t,x)\,. \tag{18.5}$$

We denote by E_i (because it becomes the electric field in the massless limit) the momentum conjugated to A_i,

$$E_i = \frac{\partial\mathcal{L}}{\partial\dot{A}_i} = \dot{A}_i - \partial_i\left(\partial_\perp^2 - m^2\right)^{-1}\partial_j\dot{A}_j\,.$$

After Legendre transformation we obtain the corresponding hamiltonian density,

$$\mathcal{H}(E_i, A_i) = \frac{1}{2}E_i(x)\left(\delta_{ij} - \frac{1}{m^2}\partial_i\partial_j\right)E_j(x) + \frac{1}{2}m^2 A_i^2(x) + \frac{1}{4}F_{ij}^2(x)\,. \tag{18.6}$$

The differential operator $-\partial_i\partial_j$ being non-negative, the hamiltonian is positive. The quantization procedure from now on is standard. It leads to an euclidean functional integral in which appears the euclidean reduced lagrangian. This lagrangian has the unpleasant properties that it is non-local and not $O(d)$ space–time symmetric. We note, however, that the dependence in A_0 of the action (18.3) is quadratic. Therefore, we can proceed in the following way: we substitute in the functional integral representation of the partition function the initial euclidean 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 field independent (see Appendix A19 for the more interesting non-abelian case). This shows that if we had ignored all these subtleties about quantization and immediately used the euclidean action,

$$\mathcal{S}(A_\mu) = \int \mathrm{d}^d x \left[\tfrac{1}{4}F_{\mu\nu}^2(x) + \tfrac{1}{2}m^2 A_\mu^2(x)\right] \tag{18.7}$$

with

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x), \tag{18.8}$$

we would have obtained the correct result (note that in the continuation to imaginary time we have set $A_0 = -iA_d$).

Correlation functions. The generating functional $\mathcal{Z}(J)$ of A_μ-field correlation functions is given by

$$\mathcal{Z}(J) = \int [\mathrm{d}A_\mu] \exp\left[-\mathcal{S}(A) + \int \mathrm{d}^d x\, J_\mu(x) A_\mu(x)\right]. \tag{18.9}$$

Integrating over all fields we obtain the generating functional $\mathcal{W}(J)$ of connected functions,

$$\mathcal{W}(J) = \ln \mathcal{Z}(J) = \frac{1}{2}\int \mathrm{d}^d k\, \tilde{J}_\mu(k) \Delta_{\mu\nu}(k) \tilde{J}_\nu(-k), \tag{18.10}$$

in which $\Delta_{\mu\nu}(k)$ is the vector field propagator in the Fourier representation,

$$\Delta_{\mu\nu}(k) = \frac{\delta_{\mu\nu} + k_\mu k_\nu / m^2}{k^2 + m^2}. \tag{18.11}$$

We verify that at the pole $k^2 = -m^2$ (the mass-shell) the numerator is a projector transverse to the vector $\mathbf{k}$, propagating $(d-1)$ components belonging to the vector representation of the $O(d-1)$ subgroup of $O(d)$ which leaves the vector $\mathbf{k}$ invariant.

As we have explained in Section 9.3, field theories involving massive vector fields coupled to matter are renormalizable only in dimensions $d \le 2$, which is bad for four-dimensional physics. A directly related disease (because the propagator is a homogeneous function of $\mathbf{k}$ and m) is the impossibility to pass continuously to the massless limit. However, we note that if the source $J_\mu(x)$ has the form of a conserved current,

$$\partial_\mu J_\mu(x) = 0 \Leftrightarrow k_\mu \tilde{J}_\mu(k) = 0\,, \tag{18.12}$$

then expression (18.10) reduces to

$$\mathcal{W}(J) = \frac{1}{2}\int \mathrm{d}^d k\, \tilde{J}_\mu(k) \frac{1}{k^2 + m^2} \tilde{J}_\mu(-k). \tag{18.13}$$

This means that the propagator can be replaced by $\delta_{\mu\nu}/(k^2+m^2)$ which behaves like the propagator of a scalar particle. In this case both problems of large momentum behaviour and massless limit are solved. One may now wonder why we have not used at once such a propagator: the reason is that it propagates, in addition to a vector field, a scalar particle with negative metric (like the regulator fields of Section 9.5). This is better illustrated by a short calculation.

More general propagators, interpretation. Let us add to the action (18.7), a term of the form of a regulator field action:

$$\mathcal{S}(A_\mu,\chi)=\mathcal{S}(A_\mu)-\frac{1}{2}\int \mathrm{d}^d x\left[(\partial_\mu\chi)^2+\mu^2\chi^2\right]. \tag{18.14}$$

In the absence of a source for χ, the generating functional

$$\mathcal{Z}(J)=\int[\mathrm{d}A_\mu][\mathrm{d}\chi]\exp\left[-\mathcal{S}(A_\mu,\chi)+\int \mathrm{d}^d x\, J_\mu(x)A_\mu(x)\right], \tag{18.15}$$

is proportional to the functional (18.9).

We now change variables, $A_\mu\mapsto A'_\mu$, in the integral (18.15),

$$A_\mu(x)=A'_\mu(x)+\frac{1}{m}\partial_\mu\chi(x). \tag{18.16}$$

This change leaves $F_{\mu\nu}$ invariant. If the source satisfies the conservation equation (18.12), the source term is not modified. Only the vector field mass term is affected:

$$\frac{1}{2}m^2A_\mu^2=\frac{1}{2}m^2A'^2_\mu+mA'_\mu\partial_\mu\chi+\frac{1}{2}\left(\partial_\mu\chi\right)^2, \tag{18.17}$$

and, therefore,

$$\mathcal{S}(A'_\mu,\chi)=\mathcal{S}(A'_\mu)+\int \mathrm{d}^d x\left(mA'_\mu\partial_\mu\chi-\frac{1}{2}\mu^2\chi^2\right). \tag{18.18}$$

We now integrate over $\chi(x)$ and find a new action,

$$\mathcal{S}_\xi(A'_\mu)=\mathcal{S}(A'_\mu)+\frac{1}{2\xi}\int\left(\partial_\mu A'_\mu\right)^2\mathrm{d}^d x, \tag{18.19}$$

in which we have set

$$\xi=\mu^2/m^2. \tag{18.20}$$

It is easy to calculate the corresponding propagator:

$$[\Delta_\xi]_{\mu\nu}(k)=\frac{\delta_{\mu\nu}}{k^2+m^2}+\frac{(\xi-1)k_\mu k_\nu}{(k^2+m^2)\left(k^2+\xi m^2\right)}. \tag{18.21}$$

For all finite values of ξ the propagator behaves at large momentum like a scalar field propagator. By varying ξ from 0 to $+\infty$ we describe a set of *gauges*: $\xi=0$ corresponds to Landau's gauge, $\xi=1$ is Feynman's gauge. For $\xi=\infty$ (the *unitary gauge*) we recover the original propagator of the vector field.

The massless limit. The propagator $[\Delta_\xi]_{\mu\nu}(k)$ has a zero mass limit,

$$[\Delta_\xi]_{\mu\nu}(k)=\frac{\delta_{\mu\nu}}{k^2}+(\xi-1)\frac{k_\mu k_\nu}{(k^2)^2}. \tag{18.22}$$

However, for values of $\xi\neq1$, the term proportional to $1/k^4$ may generate IR divergences in interacting theories for dimensions $d\leq4$.

propagator: the reason is that it propagates, in addition to a vector field, a scalar particle with negative metric (like the regulator fields of Section 9.5). This is better illustrated by a short calculation.

More general propagators, interpretation. Let us add to the action (18.7), a term of the form of a regulator field action:

$$\mathcal{S}(A_\mu, \chi) = \mathcal{S}(A_\mu) - \frac{1}{2}\int \mathrm{d}^d x \left[(\partial_\mu \chi)^2 + \mu^2 \chi^2\right]. \tag{18.14}$$

In the absence of a source for χ, the generating functional

$$\mathcal{Z}(J) = \int [\mathrm{d}A_\mu]\,[\mathrm{d}\chi] \exp\left[-\mathcal{S}(A_\mu, \chi) + \int \mathrm{d}^d x\, J_\mu(x) A_\mu(x)\right], \tag{18.15}$$

is proportional to the functional (18.9).

We now change variables, $A_\mu \mapsto A'_\mu$, in the integral (18.15),

$$A_\mu(x) = A'_\mu(x) + \frac{1}{m}\partial_\mu \chi(x). \tag{18.16}$$

This change leaves $F_{\mu\nu}$ invariant. If the source satisfies the conservation equation (18.12), the source term is not modified. Only the vector field mass term is affected:

$$\frac{1}{2}m^2 A_\mu^2 = \frac{1}{2}m^2 A_\mu'^2 + m A'_\mu \partial_\mu \chi + \frac{1}{2}(\partial_\mu \chi)^2, \tag{18.17}$$

and, therefore,

$$\mathcal{S}(A'_\mu, \chi) = \mathcal{S}(A'_\mu) + \int \mathrm{d}^d x \left(m A'_\mu \partial_\mu \chi - \frac{1}{2}\mu^2 \chi^2\right). \tag{18.18}$$

We now integrate over $\chi(x)$ and find a new action,

$$\mathcal{S}_\xi(A'_\mu) = \mathcal{S}(A'_\mu) + \frac{1}{2\xi}\int (\partial_\mu A'_\mu)^2\, \mathrm{d}^d x, \tag{18.19}$$

in which we have set

$$\xi = \mu^2/m^2. \tag{18.20}$$

It is easy to calculate the corresponding propagator:

$$[\Delta_\xi]_{\mu\nu}(k) = \frac{\delta_{\mu\nu}}{k^2 + m^2} + \frac{(\xi - 1) k_\mu k_\nu}{(k^2 + m^2)(k^2 + \xi m^2)}. \tag{18.21}$$

For all finite values of ξ the propagator behaves at large momentum like a scalar field propagator. By varying ξ from 0 to $+\infty$ we describe a set of *gauges*: $\xi = 0$ corresponds to Landau's gauge, $\xi = 1$ is Feynman's gauge. For $\xi = \infty$ (the *unitary gauge*) we recover the original propagator of the vector field.

The massless limit. The propagator $[\Delta_\xi]_{\mu\nu}(k)$ has a zero mass limit,

$$[\Delta_\xi]_{\mu\nu}(k) = \frac{\delta_{\mu\nu}}{k^2} + (\xi - 1)\frac{k_\mu k_\nu}{(k^2)^2}. \tag{18.22}$$

However, for values of $\xi \neq 1$, the term proportional to $1/k^4$ may generate IR divergences in interacting theories for dimensions $d \leq 4$.

18.2 Action with Fermion Matter

We conclude from the previous analysis that vector fields coupled to conserved currents, and thus associated with continuous symmetries, are candidates for the construction of theories renormalizable in four dimensions. We now construct the simplest example.

We start from a free action for a massive fermion:

$$\mathcal{S}_{\mathrm{F}}(\bar\psi,\psi) = -\int \mathrm{d}^d x\, \bar\psi(x)\,(\not\partial + M)\,\psi(x), \tag{18.23}$$

and want to add an $O(d)$ invariant coupling to a vector field.

As we have already noted in Section 8.2.3, the action (18.23) for a free Dirac fermion has a $U(1)$ symmetry associated with the conservation of the fermion number:

$$\psi(x) = \mathrm{e}^{i\Lambda}\,\psi'(x), \qquad \bar\psi(x) = \mathrm{e}^{-i\Lambda}\,\bar\psi'(x). \tag{18.24}$$

To this symmetry corresponds a current whose expression is obtained by calculating the variation of the action under a space-dependent group transformation (see Appendix A13.1). If Λ is space-dependent, the variation of the action is

$$\delta\mathcal{S}_{\mathrm{F}} = -i\int \mathrm{d}^d x\, \bar\psi(x)\not\partial\Lambda(x)\psi(x), \tag{18.25}$$

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 term which, from the point of view of power counting, has a chance to be renormalizable is proportional to

$$\int \mathrm{d}^d x\, \bar\psi(x)\not A\psi(x)\,.$$

This interaction term has exactly the form of a vector field linearly coupled to the conserved current $J_\mu(x)$. The action of a fermion interacting with a gauge field then takes the form

$$\mathcal{S}_{\mathrm{F}}(A_\mu,\bar\psi,\psi) = -\int \mathrm{d}^d x\, \bar\psi(x)\,(\not\partial + M + ie\not A)\,\psi(x). \tag{18.26}$$

The parameter e is the current–vector field coupling constant.

The transformations of Section 18.1, which have led to the propagator (18.21), rely on a change of variables in the functional integral of the form (18.16):

$$A_\mu(x) = -\frac{1}{e}\partial_\mu\Lambda(x) + A'_\mu(x). \tag{18.27}$$

We verify that the induced variation of $\mathcal{S}_{\mathrm{F}}(A_\mu,\bar\psi,\psi)$ can be cancelled by a change of the fermion variables of the form of the transformation (18.24) with a space-dependent function $\Lambda(x)$. Therefore, the algebraic transformations which allow to pass from a unitary non-renormalizable action to an action non-unitary but renormalizable by power counting remain justified.

This result is valid for the action and transverse A-field correlation functions. Eventually we shall have to discuss the problem of matter correlation functions. However, we first want to examine the special properties of the massless vector field theory.

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 A13.3). But $T_{\mu\nu}$ has at least dimension 4 for $d = 4$ and once coupled to a field of dimension at least one generates a non-renormalizable interaction. A similar argument applies to a spin 3/2 field coupled to the supersymmetry current.

18.3 Massless Vector Field: Abelian Gauge Symmetry

The total action, sum of contributions (18.26) and (18.7), has in the massless vector field limit $m = 0$ a remarkable property: it is invariant, at the classical level, under *local* (i.e. with a space-dependent parameter) $U(1)$ transformations. Such a symmetry is called a $U(1)$ *gauge symmetry* and the vector field is then called a *gauge field.* This symmetry has a geometric interpretation which we discuss now.

18.3.1 Gauge symmetry

The invariance of the fermion part of the action can be seen as a consequence of the replacement of the derivative ∂_μ of the free fermion theory by the *covariant derivative* $\partial_\mu + ieA_\mu$, which allows the transformation of the gauge field to cancel the term coming from the derivative. We are thus reminded of the concepts of covariant derivative, affine connection, curvature and parallel transport introduced for Riemannian manifolds (see Chapter 22). In particular, we note the similarity with the rotations of the local frame considered in Section 22.6. The difference is that the vectors which here are parallel-transported do not belong to the space tangential to the manifold (which is flat), but are vectors for Lie group transformations.

Let us now indicate the correspondence:

(i) $\psi(x)$ and $\bar\psi(x)$ are vectors for $U(1)$ transformations.

(ii) $A_\mu(x)$ is the affine connection. The connection is expected to have three indices, one, here called 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 here.

(iii) The covariant derivative is

$$\mathrm{D}_\mu = \partial_\mu + ieA_\mu \,. \tag{18.28}$$

(iv) It follows that the curvature tensor is

$$ieF_{\mu\nu} = [\mathrm{D}_\mu, \mathrm{D}_\nu] = ie\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right). \tag{18.29}$$

Finally, since the group is abelian, it is easy to write down explicitly the parallel transporter $U(C)$, which is an element of the $U(1)$ group, in terms of a line integral:

$$U(C) = \exp\left[-ie\oint_C A_\mu(s)\mathrm{d}s_\mu\right]. \tag{18.30}$$

This result is valid for the action and transverse A-field correlation functions. Eventually we shall have to discuss the problem of matter correlation functions. However, we first want to examine the special properties of the massless vector field theory.

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 A13.3). But $T_{\mu\nu}$ has at least dimension 4 for $d = 4$ and once coupled to a field of dimension at least one generates a non-renormalizable interaction. A similar argument applies to a spin 3/2 field coupled to the supersymmetry current.

18.3 Massless Vector Field: Abelian Gauge Symmetry

The total action, sum of contributions (18.26) and (18.7), has in the massless vector field limit $m = 0$ a remarkable property: it is invariant, at the classical level, under *local* (i.e. with a space-dependent parameter) $U(1)$ transformations. Such a symmetry is called a $U(1)$ *gauge symmetry* and the vector field is then called a *gauge field.* This symmetry has a geometric interpretation which we discuss now.

18.3.1 Gauge symmetry

The invariance of the fermion part of the action can be seen as a consequence of the replacement of the derivative ∂_μ of the free fermion theory by the *covariant derivative* $\partial_\mu + ieA_\mu$, which allows the transformation of the gauge field to cancel the term coming from the derivative. We are thus reminded of the concepts of covariant derivative, affine connection, curvature and parallel transport introduced for Riemannian manifolds (see Chapter 22). In particular, we note the similarity with the rotations of the local frame considered in Section 22.6. The difference is that the vectors which here are parallel-transported do not belong to the space tangential to the manifold (which is flat), but are vectors for Lie group transformations.

Let us now indicate the correspondence:

(i) $\psi(x)$ and $\bar{\psi}(x)$ are vectors for $U(1)$ transformations.

(ii) $A_\mu(x)$ is the affine connection. The connection is expected to have three indices, one, here called 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 here.

(iii) The covariant derivative is

$$\mathrm{D}_\mu = \partial_\mu + ieA_\mu\,. \tag{18.28}$$

(iv) It follows that the curvature tensor is

$$ieF_{\mu\nu} = [\mathrm{D}_\mu, \mathrm{D}_\nu] = ie\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right). \tag{18.29}$$

Finally, since the group is abelian, it is easy to write down explicitly the parallel transporter $U(C)$, which is an element of the $U(1)$ group, in terms of a line integral:

$$U(C) = \exp\left[-ie \oint_C A_\mu(s)\mathrm{d}s_\mu\right]. \tag{18.30}$$

One immediately verifies that the transformation of $U(C)$ induced by a gauge transformation (18.27) of A_μ, has indeed the expected form:

$$U(C) = U'(C) \exp\left[i \oint_C \partial_\mu \Lambda(s) \mathrm{d}s_\mu\right] = U'(C) \exp\left[i\bigl(\Lambda(y) - \Lambda(x)\bigr)\right], \tag{18.31}$$

where x and y are the end-points of C. By trying to build a renormalizable theory for a vector in four dimensions we have been naturally led to introduce a new geometric structure, an abelian gauge theory, which is just the quantum version of Maxwell's electromagnetism.

Quantization of charge. If we introduce additional charged fields, we have to assign them charges which characterize their transformation properties under $U(1)$. A delicate question arises here. Since the $U(1)$ group has the same Lie algebra as the group of translations, properties depending only on infinitesimal group transformations do not require a *quantization* of charge. In particular, in perturbation theory WT identities are true even if the charges are not rationally related and the necessity of a quantization of charges, therefore, never appears. The conventional wisdom is that QED does not imply charge quantization. Note, however, that the only known non-perturbative regularization, based on a lattice approximation, involves group elements in the form of parallel transporters and, therefore, requires charge quantization (see Section 18.5).

The question remains open.

Charged scalar fields. With the geometric ideas of parallel transport in mind, it is easy to construct an gauge action for a charged scalar field. We start from a $U(1)$ invariant action $\mathcal{S}_\mathrm{B}(\phi)$:

$$\mathcal{S}_\mathrm{B}(\phi) = \int \mathrm{d}^d x \left[|\partial_\mu \phi|^2 + U(|\phi(x)|^2)\right], \tag{18.32}$$

in which the field $\phi(x)$ is complex and replace the derivative ∂_μ by a covariant derivative. The explicit form of the covariant derivative depends on the charge we assign to the field $\phi(x)$. If we assume that ϕ couples to A_μ with a coupling constant e_B, we find,

$$\mathcal{S}_\mathrm{B}(\phi, A_\mu) = \int \mathrm{d}^d x \left[|\mathrm{D}_\mu \phi|^2 + U(|\phi|^2)\right], \tag{18.33}$$

where D_μ is now

$$\mathrm{D}_\mu = \partial_\mu + i e_\mathrm{B} A_\mu\,.$$

Note that power counting allows three independent interaction terms, two linear and one quadratic in the vector field. The gauge symmetry relates them.

18.3.2 The massless vector field limit

Geometric and physical considerations single out the theory with a massless vector field. It is the only one which is exactly gauge invariant and it describes the physics of QED because the photon is, at least to a very good approximation, massless. We could, therefore, have restricted our discussion to the massless case, as we shall do in the non-abelian case. However, considering the massless theory as a limit of the massive theory provides us with a simple resolution to several difficulties.

First we have already seen in Section 18.1 that we could not write at once a propagator for a massless vector field. In an interacting theory this difficulty persists, and if the theory is gauge invariant it is not only a disease of perturbation theory. Indeed the gauge symmetry implies that the action does not depend on one of the dynamical variables which is related to a gauge transformation. In particular, the functional integral is ill-defined because it is infinite by a factor which is the volume of the gauge group (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 introduced as a mathematical entity that enables deriving the classical field equations from a local covariant action. It contains redundant degrees of freedom and its evolution is not completely determined by the field equations.

We show in Section 18.4 how the gauge action can be quantized starting from first principles. The procedure is less straightforward and leads to non-covariant gauges, problems whose analysis we wanted to postpone. By giving a mass to the gauge field A_μ, we make all field components dynamical. The symmetry properties of the action allow the algebraic manipulations indicated in Section 18.1 (provided the action is used to calculate only gauge invariant observables, see Section 18.7). We can eventually take the zero mass limit. As we have shown, 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. For this reason 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 limit in which the mass of the vector field goes to zero, the gauge invariant limit, correlation functions calculated in Feynman's gauge have a finite limit. In Section 18.7 we derive a relation (equation (18.90)) which exhibits the gauge dependence of the bare two-point function. The corresponding relations for other correlation functions can be obtained by expanding expression (18.89). They imply that bare correlation functions have a limit in all gauges. The renormalized correlation functions will then also have a limit provided one has been careful not to choose IR divergent renormalization constants (as it is usual in massless theories).

The important question is, however, to understand whether physical, that is, gauge independent observables are IR finite. The averages of gauge invariant operators have certainly limits. On the contrary, we do not expect scattering amplitudes to be IR finite. Actually we have shown in Section 6.3.3, using the eikonal approximation, that even in quantum mechanics IR divergences appear in the phase of the scattering amplitude as a consequence of the long distance behaviour of the Coulomb force. These divergences survive in the relativistic theory. However, it can be shown that 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 gauge fields. This is physically acceptable because gauge fields which have momenta smaller than the uncertainty in the momenta measurements cannot be detected. Since this question is somewhat technical, we refer to the literature where detailed discussions of this question can be found.

18.4 Canonical Quantization and Gauge Invariance

Although we have been able to construct a gauge invariant theory as a limit of a theory of a massive vector field coupled to a conserved current, it is useful to contemplate the difficulties one encounters when one tries to quantize a gauge theory starting from first principles. Moreover, as we shall show in Appendix A19, in the case of non-abelian gauge symmetries, the massless limit is not continuous. We, therefore, show how, starting directly from the classical field equations of a gauge invariant theory, it is possible to recover the functional integral representation of the generating functional of correlation functions.

The problem can be solved by several different strategies and we shall present two of them, corresponding to so-called *Coulomb's gauge* and *temporal gauge*. We again consider the simple action:

$$\mathcal{S}(A_\mu, J_\mu) = \mathcal{S}(A_\mu) - \int \mathrm{d}^d x\, J_\mu(x) A_\mu(x), \tag{18.34}$$

$$\mathcal{S}(A_\mu) = \frac{1}{4} \int \mathrm{d}^d x\, F_{\mu\nu}^2(x)\,, \tag{18.35}$$

of a gauge field coupled to a conserved current.

18.4.1 Coulomb's gauge

We first proceed as in the massive case and eliminate the field time component from the action. Taking into account current conservation (note $J_0 = iJ_d$), we then obtain the integral of a reduced lagrangian density:

$$\mathcal{L}(A_i) = \frac{1}{2}\dot{A}_i(t,x)\left(\delta_{ij} - \frac{\partial_i\partial_j}{\partial_\perp^2}\right)\dot{A}_j(t,x) - \frac{1}{4}F_{ij}^2(t,x) + J_i(t,x)\left(\delta_{ij} - \frac{\partial_i\partial_j}{\partial_\perp^2}\right)A_j(t,x)\,.$$

In contrast with the massive case the action depends only on $\left(\delta_{ij} - \partial_i\partial_j/\partial_\perp^2\right)A_j(t,x)$. 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 $\mathbf{k}$. We recover the well-known property that a massless vector field has only $d-2$ physical components. We thus expand the vector $\tilde{A}_i(t,\hat{k})$ on a transverse basis $e_i^a(\hat{k})$, calling $\tilde{\mathfrak{A}}_a$ the corresponding $d-2$ components:

$$\hat{k}\cdot\mathbf{e}^a(\hat{k}) = 0\,, \quad \mathbf{e}^a(\hat{k})\cdot\mathbf{e}^b(\hat{k}) = \delta_{ab}, \qquad \tilde{\mathbf{A}}(t,\hat{k}) = \frac{\hat{k}\cdot\tilde{\mathbf{A}}}{\hat{k}^2}\hat{k} + \sum_{a=1}^{d-2}\mathbf{e}^a(\hat{k})\tilde{\mathfrak{A}}_a(\mathbf{k})\,,$$

and J_a the corresponding sources. The lagrangian density in these variables becomes

$$\mathcal{L}(\mathfrak{A}_a) = \sum_{a=1}^{d-2}\left[\partial_\mu\mathfrak{A}_a(t,x)\partial^\mu\mathfrak{A}_a(t,x) + J_a(t,x)\mathfrak{A}_a(t,x)\right].$$

The quantization now is straightforward. One eventually obtains a functional integral over the fields $\mathfrak{A}_a$. The corresponding action, once expressed in terms of the initial current J_μ, is, however, non-local. One can reintroduce the components A_i of the gauge field provided one multiplies the integrand by $\delta(\partial_i A_i)$. 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 euclidean generating functional in *Coulomb's gauge*:

$$\mathcal{Z}_{\mathrm{Coul.}}(J) = \int \left[\mathrm{d}A_\mu(x)\delta\left(\partial_i A_i(x)\right)\right]\exp\left[-\mathcal{S}(A) + \int \mathrm{d}^d x\, J_\mu(x)A_\mu(x)\right]. \tag{18.36}$$

Coulomb's gauge, in the abelian case, has a nice physical interpretation: only physical degrees of freedom propagate, but it leads to non-covariant calculations, and this is a serious drawback. In particular, the gauge field propagator ($\mathbf{k} \equiv \{k_d, \mathbf{k}_\perp\}$) becomes

$$\begin{cases} W^{(2)}_{dd}(k) = \dfrac{1}{\mathbf{k}_\perp^2}, \\ W^{(2)}_{id}(k) = 0, \\ W^{(2)}_{ij}(k) = \dfrac{1}{\mathbf{k}^2}\left(\delta_{ij} - \dfrac{k_i k_j}{\mathbf{k}_\perp^2}\right). \end{cases} \tag{18.37}$$

The time component does not decrease in the large time direction. Therefore, with this propagator the theory is not explicitly renormalizable by power counting.

We still have to prove that this gauge is equivalent to the covariant gauges introduced in Section 18.1, but we postpone this point and discuss before another quantization scheme.

18.4.2 The temporal (or Weyl) gauge

In the non-abelian case the quantization in Coulomb's gauge is complicated and, therefore, we now explain another, more easily generalizable, method.

The field equations in real time t that correspond to the action (18.34) are

$$\partial_\mu F^{\mu\nu}(t,x) = J^\nu(t,x), \quad x \in \mathbb{R}^{d-1}. \tag{18.38}$$

The extension of the arguments that follow to a gauge theory containing matter fields is straightforward.

The method relies on the observation that the gauge transformed of any solution of equation (18.38) 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

$$A_0(t,x) = 0, \tag{18.39}$$

in which A_0 is the time component of the field A_μ. We then rewrite equation (18.38), separating time and space components, and taking into account the condition (18.39):

$$\partial_j \dot{A}_j(t,x) = J_0(t,x), \tag{18.40}$$

$$\ddot{A}_i(t,x) - \partial_j F_{ji}(t,x) = J_i(t,x), \tag{18.41}$$

in which the indices i and j run only from 1 to $d-1$. The equation (18.41) is simply the field equation that can be derived from the classical lagrangian density $\mathcal{L}$:

$$\mathcal{L}(A_i) = \tfrac{1}{2}\dot{A}_i^2 - \tfrac{1}{4}F_{ij}^2 + J_i A_i. \tag{18.42}$$

The conjugated momentum $E_i(t,x)$ of the field $A_i(t,x)$ is the electric field

$$E_i(t,x) = \dot{A}_i(t,x). \tag{18.43}$$

The expression of the hamiltonian density follows

$$\mathcal{H}\big(E(x),A(x)\big) = \tfrac{1}{2}E_i^2(x) + \tfrac{1}{4}F_{ij}^2(x) - J_i(x)A_i(x). \tag{18.44}$$

The partition function $\mathcal{Z}(J_i)$ is then

$$\mathcal{Z}(J_i) = \int [\mathrm{d}A_i] \exp\left[-\mathcal{S}(A_i) + \int \mathrm{d}^d x\, J_i(x) A_i(x)\right], \tag{18.45}$$

in which the euclidean action $\mathcal{S}(A_i)$ is the covariant action (18.35), in which $A_d = 0$ has been set.

We still have to implement the constraint (18.40) which is Gauss's law. After quantization, it becomes a constraint on the physically acceptable states $\Psi(\mathbf{A})$. Since the conjugated momenta E_i as quantum operators are represented by differential operators $-i\delta/\delta A_i$, the condition (18.40) takes the form

$$\frac{1}{i}\partial_i \frac{\delta}{\delta A_i(x)} \Psi(\mathbf{A}) = J_0(t,x)\Psi(\mathbf{A}) . \tag{18.46}$$

We recognize in the l.h.s. of equation (18.46) the generator of time-independent gauge transformations of the field $\mathbf{A}(x)$ 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 because in the gauge (18.39) the theory has still an invariance under time-independent gauge transformations.

For a general external source, the condition (18.46) tells us how the state transforms. Consistency with quantum evolution then requires the commutation of the operator $\partial_i E_i - J_0$ with the hamiltonian. A short calculation shows that this commutation is implied by the current conservation.

Finally, we exhibit, for later purpose, a state satisfying the condition (18.46) in the case of two opposite static charges:

$$J_0(t,x) = e\left[\delta(x-x_2) - \delta(x-x_1)\right], \quad J_i(t,x) = 0 . \tag{18.47}$$

The state,

$$\Psi(\mathbf{A}) = \exp\left[-ie \oint_C A_i(s)\mathrm{d}s_i\right], \tag{18.48}$$

in which C is an arbitrary path joining x_1 to x_2, indeed verifies

$$\begin{aligned}\frac{1}{i}\partial_i \frac{\delta}{\delta A_i(x)} \Psi(\mathbf{A}) &= \frac{1}{i}\frac{\delta}{\delta \Lambda(x)}\ \Psi(\mathbf{A} - \nabla_x \Lambda)\big|_{\Lambda=0} \\ &= e\left[\delta(x-x_2) - \delta(x-x_1)\right]\Psi(\mathbf{A}),\end{aligned}$$

a result that is consistent with the equations (18.46,18.47).

Note that the representation (18.48) has the form of a parallel transporter corresponding to time-independent gauge transformations. This representation, as well as its non-abelian generalization, will be useful in Chapter 34, in the discussion of the confinement problem.

The propagator in the temporal gauge. The propagator of the gauge field in the temporal gauge (in the euclidean formalism) reads

$$W^{(2)}_{ij} = \frac{1}{k^2}\left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}_\perp^2}\right) + \frac{1}{k_d^2}\frac{k_i k_j}{\mathbf{k}_\perp^2}, \tag{18.49}$$

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 which are not explicitly renormalizable in four dimensions. Moreover, its longitudinal part has a double pole at $k_d = 0$ that also requires some regularization.

18.4.3 Equivalence with covariant quantization

We have obtained different functional integral representations of the same theory. We would like to show that, at least for gauge invariant observables, they are formally equivalent to the $O(d)$ covariant representations we have discussed in Section 18.1. We here show the equivalence between the temporal gauge and a class of gauges characterized by a gauge condition of the form:

$$n_\mu(\partial)A_\mu(x) = \nu(x), \tag{18.50}$$

where the vector n_μ is a constant or a differential operator and $\nu(x)$ an arbitrary external field. By setting the external field $\nu(x)$ to zero, one enforces the strict gauge condition $n_\mu(\partial)A_\mu(x) = 0$ but by integrating over it with a gaussian weight one can generate actions of the form (18.19). This covers all the examples met so far. The arguments easily generalize to other gauges.

We first write the expression (18.45) as an integral over a d-component vector field:

$$\mathcal{Z}(J) = \int [\mathrm{d}A_\mu] \prod_x \delta\left(A_d(x)\right) \exp\left\{-\int \mathrm{d}^d x \left[\tfrac{1}{4}F^2_{\mu\nu}(x) - J_\mu(x)A_\mu(x)\right]\right\}. \tag{18.51}$$

We then insert the identity:

$$\int [\mathrm{d}\Lambda(x)] \prod_x \delta\left[n_\mu(\partial)\left(\partial_\mu\Lambda(x) + A_\mu(x)\right) - \nu(x)\right] = \text{const.}\,, \tag{18.52}$$

inside expression (18.51):

$$\begin{aligned}\mathcal{Z} \propto \int [\mathrm{d}A_\mu \mathrm{d}\Lambda] \prod_x \delta(A_d)\delta\left[n_\mu(\partial)\left(\partial_\mu\Lambda(x) + A_\mu(x)\right) - \nu(x)\right] \\ \times \exp\left[-\mathcal{S}(A) + \int \mathrm{d}^d x\, J_\mu(x)A_\mu(x)\right].\end{aligned} \tag{18.53}$$

We perform a change of variables, $A \mapsto A'$, of the form of a gauge transformation:

$$A_\mu(x) = A'_\mu(x) - \partial_\mu\Lambda(x).$$

Since the current J_μ is conserved, only the δ-functions are modified:

$$\delta(A_d)\delta\left[n_\mu(\partial)\left(\partial_\mu\Lambda(x) + A_\mu(x)\right) - \nu(x)\right] \mapsto \delta(A_d - \dot{\Lambda})\delta\left[n_\mu(\partial)A_\mu(x) - \nu(x)\right].$$

The integration over Λ can again be performed:

$$\int [\mathrm{d}\Lambda] \prod_x \delta\left(A_d(x) - \dot{\Lambda}(x)\right) = \text{const.}\,, \tag{18.54}$$

and, therefore,

$$\mathcal{Z}(J) = \int [\mathrm{d}A_\mu]\, \delta\left[n_\mu(\partial)A_\mu(x) - \nu(x)\right] \exp\left[-\mathcal{S}(A) + \int \mathrm{d}^d x\, J_\mu(x)A_\mu(x)\right]. \tag{18.55}$$

Since the result by construction does not depend on $\nu(x)$ we can either set $\nu(x)$ to zero or integrate over $\nu(x)$ with for example the gaussian measure $\mathrm{d}\rho(\nu)$:

$$[\mathrm{d}\rho(\nu)] = [\mathrm{d}\nu(x)] \exp\left[-\frac{1}{2}\int \mathrm{d}^d x\, \nu^2(x)\right]. \tag{18.56}$$

We then obtain

$$\mathcal{Z}(J) = \int [\mathrm{d}A_\mu] \exp\left[-\mathcal{S}_{\mathbf{gauge}}(A) + \int \mathrm{d}^d x\, J_\mu(x) A_\mu(x)\right] \tag{18.57}$$

with

$$\mathcal{S}_{\mathbf{gauge}}(A) = \mathcal{S}(A) + \frac{1}{2}\int \mathrm{d}^d x\, [n_\mu(\partial) A_\mu(x)]^2 . \tag{18.58}$$

Specializing to $n_\mu = \xi^{-1/2}\partial_\mu$ we see that we have in particular demonstrated the equivalence between the *temporal gauge* $A_d = 0$ and the covariant gauges (18.19). If instead we choose $n_d = 0$ and $\mathbf{n}_\perp \equiv \partial_\perp$, and set $h = 0$, we find Coulomb's gauge.

The propagator. To the action (18.58) corresponds the gauge field propagator:

$$W^{(2)}_{\mu\nu}(k) = \frac{1}{k^2}\left[\delta_{\mu\nu} - \frac{(k_\mu n_\nu + k_\nu n_\mu)}{k\cdot n} + \frac{(k^2 + n^2)\, k_\mu k_\nu}{(k\cdot n)^2}\right]. \tag{18.59}$$

Remark. The strict gauge condition is recovered in the limit $|n_\mu| \to \infty$, which exists for the propagator but not for the action. To write explicitly the limiting action one has to introduce a Lagrange multiplier $\lambda(x)$ which implements the gauge condition:

$$\mathcal{S}_{\mathbf{gauge}}(A) = \mathcal{S}(A) + \int \mathrm{d}^d x\, \lambda(x) n_\mu(\partial) A_\mu(x) . \tag{18.60}$$

18.4.4 Interpretation: the Faddeev–Popov quantization

The result we have obtained has the following interpretation which can be justified only in the lattice approximation on a finite lattice (see Section 18.5). The problem of the gauge invariant theory 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. We, therefore, supply them with a stochastic dynamics in the sense of Chapter 16. We write the gauge field A_μ in terms of a gauge field B_μ projection of A_μ on some gauge section, that is, satisfying some gauge condition, and a gauge transformation Λ:

$$A_\mu = B_\mu + \partial_\mu \Lambda . \tag{18.61}$$

We assume that this decomposition is unique. Gauge invariance implies that the gauge action depends only on B_μ and specifies its dynamics.

To $\Lambda(x)$ we impose, for example,

$$\partial^2 \Lambda(x) + \partial_\mu B_\mu(x) = \nu(x), \tag{18.62}$$

in which $\nu(x)$ is a stochastic field with a given probability distribution.

We now impose this equation in the functional integral. This generates also the determinant of the functional derivative of the equation (18.62) with respect to the field $\Lambda(x)$. However, here this operator is just ∂^2 and, therefore, the determinant is a constant which disappears in the normalization of the functional integral.

The functional integral in the presence of sources for gauge invariant operators (polynomials in the fields that are invariant in gauge transformations) only becomes:

$$\mathcal{Z} = \int [\mathrm{d}B_\mu \mathrm{d}\Lambda]\, \delta\left[\partial^2\Lambda(x) + \partial_\mu B_\mu(x) - \nu(x)\right] \exp\left[-\mathcal{S}(B)\right]. \tag{18.63}$$

The functional measure $[\mathrm{d}B_\mu \mathrm{d}\Lambda]$ is the decomposition of the flat measure $[\mathrm{d}A_\mu]$ into a product of measures on B_μ and Λ. The action $\mathcal{S}(B)$ is the gauge invariant action $\mathcal{S}(A)$ in which equation (18.61) has been used. We now recognize that the whole expression can be rewritten in terms of A_μ as

$$\mathcal{Z} = \int [\mathrm{d}A_\mu]\, \delta\left[\partial_\mu A_\mu(x) - \nu(x)\right] \exp\left[-\mathcal{S}(A)\right]. \tag{18.64}$$

Moreover, since the result of the functional integration does not depend on the dynamics of $\Lambda(x)$, the result does not depend on the field $\nu(x)$ either and we can integrate over $\nu(x)$ with for example the gaussian measure $\mathrm{d}\rho(\nu)$.

18.5 Perturbation Theory, Regularization

From now on we consider only the covariant gauges of Section 18.1. With the propagator (18.21) power counting is the same as for a scalar field. We can, therefore, construct interacting theories renormalizable for dimensions $d \le 4$. We shall specially consider the dimension 4. Since WT identities play an essential role in gauge theories, we have first to find gauge invariant regularizations.

Dimensional regularization. We have defined dimensional regularization in Section 9.6. This regularization is well suited to perturbative calculations in QED. Examples will be given in Section 18.9. It leads to problems only in the case of chiral gauge theories, due to the γ_S problem (see Section 20.3).

18.5.1 Momentum cut-off regularization

In this chapter the special problems generated by chiral fermions will not be met, and we will be able to calculate with dimensional regularization. However, for later purpose it is instructive to also discuss momentum cut-off or Pauli–Villars's regularization, specially 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 the new problem that arises in presence of a gauge field is that only covariant derivatives are allowed, because gauge invariance is essential for the physical consistency of the theory. The regularized action in a gauge background now reads

$$\mathcal{S}_{\mathrm{F}}(\bar\psi, \psi, A) = \int \mathrm{d}^d x\, \bar\psi(x)\, (M + \not{D}) \prod_r \left(1 - \not{D}^2/M_r^2\right) \psi(x), \tag{18.65}$$

We now impose this equation in the functional integral. This generates also the determinant of the functional derivative of the equation (18.62) with respect to the field $\Lambda(x)$. However, here this operator is just ∂^2 and, therefore, the determinant is a constant which disappears in the normalization of the functional integral.

The functional integral in the presence of sources for gauge invariant operators (polynomials in the fields that are invariant in gauge transformations) only becomes:

$$\mathcal{Z} = \int [\mathrm{d}B_\mu \mathrm{d}\Lambda]\, \delta\left[\partial^2\Lambda(x) + \partial_\mu B_\mu(x) - \nu(x)\right] \exp\left[-\mathcal{S}(B)\right]. \tag{18.63}$$

The functional measure $[\mathrm{d}B_\mu \mathrm{d}\Lambda]$ is the decomposition of the flat measure $[\mathrm{d}A_\mu]$ into a product of measures on B_μ and Λ. The action $\mathcal{S}(B)$ is the gauge invariant action $\mathcal{S}(A)$ in which equation (18.61) has been used. We now recognize that the whole expression can be rewritten in terms of A_μ as

$$\mathcal{Z} = \int [\mathrm{d}A_\mu]\, \delta\left[\partial_\mu A_\mu(x) - \nu(x)\right] \exp\left[-\mathcal{S}(A)\right]. \tag{18.64}$$

Moreover, since the result of the functional integration does not depend on the dynamics of $\Lambda(x)$, the result does not depend on the field $\nu(x)$ either and we can integrate over $\nu(x)$ with for example the gaussian measure $\mathrm{d}\rho(\nu)$.

18.5 Perturbation Theory, Regularization

From now on we consider only the covariant gauges of Section 18.1. With the propagator (18.21) power counting is the same as for a scalar field. We can, therefore, construct interacting theories renormalizable for dimensions $d \le 4$. We shall specially consider the dimension 4. Since WT identities play an essential role in gauge theories, we have first to find gauge invariant regularizations.

Dimensional regularization. We have defined dimensional regularization in Section 9.6. This regularization is well suited to perturbative calculations in QED. Examples will be given in Section 18.9. It leads to problems only in the case of chiral gauge theories, due to the γ_S problem (see Section 20.3).

18.5.1 Momentum cut-off regularization

In this chapter the special problems generated by chiral fermions will not be met, and we will be able to calculate with dimensional regularization. However, for later purpose it is instructive to also discuss momentum cut-off or Pauli–Villars's regularization, specially 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 the new problem that arises in presence of a gauge field is that only covariant derivatives are allowed, because gauge invariance is essential for the physical consistency of the theory. The regularized action in a gauge background now reads

$$\mathcal{S}_{\mathrm{F}}(\bar\psi, \psi, A) = \int \mathrm{d}^d x\, \bar\psi(x)\, (M + \not{D}) \prod_r \left(1 - \not{D}^2/M_r^2\right) \psi(x), \tag{18.65}$$

which replaces the expression (9.38). Note that up to this point the regularization, unlike dimensional or lattice regularizations, preserves a possible chiral symmetry for $M = 0$.

As a consequence the higher order derivatives of the regularization 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

$$\mathcal{Z}(\bar\eta, \eta; A) = \int \left[\mathrm{d}\psi(x)\mathrm{d}\bar\psi(x)\right] \exp\left[-\mathcal{S}_{\mathrm{F}}(\bar\psi, \psi, A) + \int \mathrm{d}^d x\, \bar\eta(x)\psi(x) + \bar\psi(x)\eta(x)\right], \tag{18.66}$$

where $\bar\eta, \eta$ are Grassmann sources. Integrating over fermions explicitly, we obtain

$$\mathcal{Z}(\bar\eta, \eta; A) = \mathcal{Z}_0(A) \exp\left[-\int \mathrm{d}^d x\, \mathrm{d}^d y\, \bar\eta(y)\Delta_{\mathrm{F}}(A; y, x)\eta(x)\right], \tag{18.67}$$

$$\mathcal{Z}_0(A) = \mathcal{N} \det\left[(M + \not{D}) \prod_r \left(1 - \not{D}^2/M_r^2\right)\right],$$

where $\mathcal{N}$ is a gauge field-independent normalization and $\Delta_{\mathrm{F}}(A; y, x)$ the fermion propagator in an external gauge field.

Diagrams constructed from $\Delta_{\mathrm{F}}(A; y, x)$ (figure 18.1) belong to loops with gauge field propagators, and, therefore, can be rendered finite if the gauge field propagator can be improved, a condition that we check below.

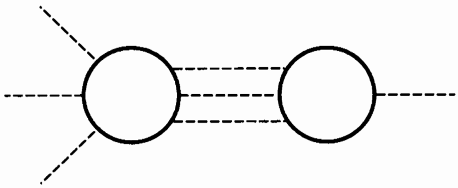


Fig. 18.1 Example of a multi-loop diagram.

The other problem involves the determinant that generates closed fermion loops in gauge background (like in figure 18.2). Using $\ln\det = \operatorname{tr}\ln$, we find

$$\ln \mathcal{Z}_0(A) = \operatorname{tr}\ln(M + \not{D}) + \sum_r \operatorname{tr}\ln\left(1 - \not{D}^2/M_r^2\right) - (A = 0),$$

or using the anticommutation of γ_5 with $\not{D}$,

$$\det(\not{D} + M) = \det \gamma_5(\not{D} + M)\gamma_5 = \det(M - \not{D}),$$

$$\ln \mathcal{Z}_0(A) = \tfrac{1}{2}\operatorname{tr}\ln\left(M^2 - \not{D}^2\right) + \sum_r \operatorname{tr}\ln\left(1 - \not{D}^2/M_r^2\right) - (A = 0),$$

We see that the regularization has no effect from the point of view of power counting on the determinant, and, therefore, on 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

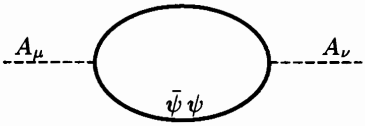


Fig. 18.2 One-loop contribution to the gauge field two-point function.

expression for the determinant of operators of the form $\not{D} + M$ in the continuum and at fixed dimension (see Section 20.3).

The fermion determinant. The fermion determinant can finally be regularized by adding to the action a bosonic regulator field with fermion spin, and, therefore, a propagator similar to Δ_{F} but with different masses

$$\mathcal{S}_{\mathrm{B}}(\bar\phi, \phi; A) = \int \mathrm{d}^d x\, \bar\phi(x)\left(M_0^{\mathrm{B}} + \not{D}\right) \prod_{r=1} \left(1 - \not{D}^2/(M_r^{\mathrm{B}})^2\right) \phi(x). \tag{18.68}$$

The Section 32.2.3 will provide an explicit example.

The integration over the boson ghost fields $\bar\phi, \phi$ adds to $\ln \mathcal{Z}_0$ the quantity

$$\delta \ln \mathcal{Z}_0(A) = -\tfrac{1}{2}\,\mathrm{tr}\ln\left((M_0^{\mathrm{B}})^2 - \not{D}^2\right) - \sum_{r=1} \mathrm{tr}\ln\left(1 - \not{D}^2/(M_r^{\mathrm{B}})^2\right) - (A = 0).$$

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}. 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 we obtain the determinant of the transformed operator

$$\mathrm{e}^{i\theta\gamma_5(x)}\,\not{D}\,\mathrm{e}^{i\theta\gamma_5(x)}(\not{D} + M_0^{\mathrm{B}})^{-1}.$$

For $\theta(x)$ constant $\mathrm{e}^{i\theta\gamma_5}$ anticommutes with $\not{D}$ and cancels. Otherwise, a non-trivial contribution remains. The method thus suggests possible difficulties with space-dependent chiral transformations.

Since actually the problem reduces to the study of a determinant in an external background, one can study it directly, as we will, starting with Section 20.3. One verifies whether it is possible to define some regularized form in a way consistent with chiral symmetry. One then inserts the one-loop renormalized diagrams in the general diagrams regularized by the preceding cut-off methods.

The gauge field propagator. For the free gauge action in a covariant gauge usual derivatives can be used because in an abelian theory the gauge field is neutral. The tensor $F_{\mu\nu}$ is gauge invariant and the action for the scalar $\chi(x)$ of equation (18.14) and thus $\partial_\mu A_\mu$ is arbitrary. 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 made superficially convergent.

Finally, the determinant $\det \mathrm{D}_\mu^2$ generated by integrating over charged scalar fields in gauge background can be regularized by Schwinger's proper time method (see Appendix A18.3). In this way the determinant is expressed in terms of the evolution operator corresponding to a non-relativistic hamiltonian in a magnetic field.

18.5.2 Lattice regularization

To understand how to construct a lattice regularization of a gauge theory, it is essential to remember the geometric interpretation of the gauge field as a connection. 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 Chapters 22,34).

Gauge invariant interaction terms on the lattice then have the form:

$$\bar{\psi}(x + an_\mu)\gamma_\mu U(x + an_\mu, x)\psi(x), \tag{18.69}$$

in which U is given by equation (18.30), the curve C being the link joining the points x to $x + an_\mu$ on the lattice. We have called n_μ the unit vector in μ direction and a the lattice spacing.

The link variable $U(x, y)$, linking site x to y is an element of the $U(1)$ group, which can, therefore, be parametrized in terms of an angle θ_{xy}, and is such that:

$$U(x, y) \equiv U_{xy} = \mathrm{e}^{i\theta_{xy}} = (U_{yx})^{-1}. \tag{18.70}$$

We show in Section 22.4 that the curvature tensor is associated with parallel transport around a closed curve. This suggests that we should take as a regularized form of $\int \mathrm{d}x\, F_{\mu\nu}^2$ the product of link variables on a closed curve on the lattice, the simplest being a square on a hypercubic lattice. Such a product is clearly gauge invariant. We then obtain the well-known ***plaquette action***, each square forming a plaquette (for details see Chapter 34):

$$\sum_{\text{all plaquettes}} U_{xy}U_{yz}U_{zt}U_{tx}\,. \tag{18.71}$$

We have denoted symbolically by x, y, z, t four sites forming a square on the lattice.

The typical gauge invariant lattice action corresponding to the continuum action of a gauge field coupled to fermions then has the form:

$$\mathcal{S}(U, \bar{\psi}, \psi) = \beta \sum_{\text{plaquettes}} U_{xy}U_{yz}U_{zt}U_{tx} - \kappa \sum_{\text{links}} \bar{\psi}_y \gamma_{yx} U_{yx}\psi_x - \sum_{\text{sites}} M\bar{\psi}_x\psi_x\,. \tag{18.72}$$

We have denoted by x, y,... the lattice sites, β and κ are the coupling constants. The action (18.72) is invariant under independent $U(1)$ transformations on each lattice site. These transformations are the lattice equivalents of the gauge transformations of the continuum theory. The measure of integration over the gauge variables is the group invariant measure, that is, the flat measure $\mathrm{d}\theta_{xy}$. Note that on the lattice and in a finite volume the gauge invariant action leads to a well-defined partition function because the $U(1)$ group is compact: the volume of the gauge group is $(2\pi)^\Omega$ if Ω is the number of lattice sites. However, in the continuum limit, the compact character of the group is lost. It is, therefore, necessary to fix the gauge on the lattice in order to be able to construct a regularized perturbation theory. Since we shall devote Chapter 34 to lattice gauge theories, we postpone the discussion of this problem. Finally, it is possible to add to the gauge invariant lattice action a term of the form $\sum \cos\theta_{xy}$ to give a mass to the vector field.

18.6 WT Identities, Renormalization

In gauge theories WT identities play an essential role because it is necessary to maintain the gauge symmetry in order to ensure that the theory which is not explicitly unitary, is equivalent to a unitary theory, at least for the "physical" observables, that is, gauge invariant observables (including S-matrix elements).

Their derivation is rather straightforward. We take the example of the action (18.26) and call $J_\mu(x)$, $\eta(x)$ and $\bar\eta(x)$ the sources for the fields $A_\mu(x)$, $\bar\psi(x)$ and $\psi(x)$, respectively. We make infinitesimal gauge transformations (18.24,18.27) on the action in presence of sources. In this section, for convenience, we assume dimensional regularization. The terms that are not invariant are the A_μ mass term, the gauge fixing term and the sources:

$$\delta\left[\mathcal{S}-\text{source terms}\right]=-\frac{1}{e}\int \mathrm{d}^d x\,\Lambda(x)\left\{\left(\partial^2/\xi-m^2\right)\partial_\mu A_\mu(x)+\partial_\mu J_\mu(x)\right.$$
$$\left.+ie\left[\bar\eta(x)\psi(x)-\bar\psi(x)\eta(x)\right]\right\}. \tag{18.73}$$

This leads, following the usual arguments, to an equation for the generating functionals $\mathcal{Z}(J,\bar\eta,\eta)$ and thus $\mathcal{W}(J,\bar\eta,\eta)$ which has the form:

$$\left\{\left(m^2-\partial^2/\xi\right)\partial_\mu\frac{\delta}{\delta J_\mu(x)}-ie\left[\bar\eta(x)\frac{\delta}{\delta\bar\eta(x)}-\eta(x)\frac{\delta}{\delta\eta(x)}\right]\right\}\mathcal{W}(J,\bar\eta,\eta)=\partial_\mu J_\mu(x). \tag{18.74}$$

This equation is equivalent to a set of identities for correlation functions. The equation for the gauge field two-point function is special and after Fourier transformation takes the form:

$$k_\mu W^{(2)}_{\mu\nu}(k)=\xi\frac{k_\nu}{k^2+\xi m^2}\,. \tag{18.75}$$

Correlation functions with only matter fields satisfy:

$$\begin{aligned}&(k^2/\xi+m^2)k_\mu W^{(2n+1)}_\mu(k;p_1,\ldots,p_n;q_1,\ldots,q_n)\\&\quad=e\sum_i\Big[W^{(2n)}(p_1,\ldots,p_i+k,\ldots,p_n;q_1,\ldots,q_n)\\&\qquad-W^{(2n)}(p_1,\ldots,p_n;q_1,\ldots,q_i+k,\ldots,q_n)\Big]\,,\end{aligned} \tag{18.76}$$

in which k is the gauge field momentum, and p_i and q_i the momenta of ψ and $\bar\psi$ fields, respectively. The presence of additional external gauge fields does not modify the identities.

The equation (18.74) is a linear first-order partial differential equation. The Legendre transformation is simple and yields an equation for the generating functional of proper vertices $\Gamma(A_\mu,\bar\psi,\psi)$:

$$\left(\partial^2/\xi-m^2\right)\partial_\mu A_\mu(x)+\partial_\mu\frac{\delta\Gamma}{\delta A_\mu(x)}+ie\left[\psi(x)\frac{\delta\Gamma}{\delta\psi(x)}-\bar\psi(x)\frac{\delta\Gamma}{\delta\bar\psi(x)}\right]=0\,. \tag{18.77}$$

It can be verified that the equations (18.74) and (18.77) have the same content as the quantum equations of motion of the χ-field of Section 18.1.

The general solution of equation (18.77) can be written as

$$\Gamma=\Gamma_{\text{sym.}}+\frac{1}{2}\int\left[m^2A_\mu^2(x)+\left(\partial_\mu A_\mu\right)^2/\xi\right]\mathrm{d}^d x\,, \tag{18.78}$$

where $\Gamma_{\text{sym.}}$ is gauge invariant.

Renormalization. We perform a loop expansion of the functional Γ. Because the equation (18.77) 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 we find for $\ell > 0$,

$$\partial_\mu \frac{\delta \Gamma_\ell}{\delta A_\mu(x)} + ie\left[\psi(x)\frac{\delta\Gamma_\ell}{\delta\psi(x)} - \bar\psi(x)\frac{\delta\Gamma_\ell}{\delta\bar\psi(x)}\right] = 0 .$$

Therefore, the generating functional Γ_ℓ of ℓ-loop proper vertices 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 Γ_ℓ^{div} is gauge invariant.

The conclusion is that the action can be completely renormalized by adding gauge invariant counterterms. As in the case of the linear symmetry breaking in Section 13.3, one can say that the terms which break the gauge invariance, the gauge field mass term and the gauge fixing term, are not renormalized since they are not modified by counterterms.

The full renormalized action can then be written as

$$\mathcal{S}_{\text{r}}(A_\mu,\bar\psi,\psi) = \int \mathrm{d}^d x \left[\tfrac14 Z_A F_{\mu\nu}^2 + \tfrac12 m^2 A_\mu^2 + \tfrac{1}{2\xi}\left(\partial_\mu A_\mu\right)^2 \right.$$
$$\left. - Z_\psi \bar\psi \left(\not\partial + M + \delta M + ie\not A\right)\psi\right], \tag{18.79}$$

where Z_A is the gauge field, Z_ψ and δM are the ψ field and mass renormalization constants. We note that there is no special renormalization constant for the charge e. Indeed if we introduce the bare fields $\psi^0, \bar\psi^0$, A_μ^0,

$$\psi^0 = Z_\psi^{1/2}\psi\,, \quad \bar\psi^0 = Z_\psi^{1/2}\bar\psi\,, \quad A_\mu^0 = Z_A^{1/2}A_\mu\,, \tag{18.80}$$

we see that we must set

$$e_0 = Z_e^{1/2} e\,, \quad \text{with} \quad Z_A Z_e = 1\,. \tag{18.81}$$

In other words, the transformation law has not been renormalized:

$$\partial_\mu + ie_0 A_\mu^0 = \partial_\mu + ieA_\mu\,. \tag{18.82}$$

Gauge invariance relates the renormalization of the charge and the gauge field.

18.7 Gauge Dependence

To understand the physical properties of the theory, it is interesting to characterize the gauge dependence, here the dependence on the parameter ξ which appears in the gauge fixing term, of correlation functions.

We, therefore, calculate the variation of correlation functions when we add a term proportional $\int \left(\partial_\mu A_\mu\right)^2 \mathrm{d}^d x$ to the renormalized action.

The fermion two-point function. We first consider the fermion two-point function. Using twice the WT identity (18.76) we obtain a relation which, in Fourier space, reads:

$$W^{(2)}_{(\partial_\mu A_\mu)^2}(p) = 2e^2 \int \frac{\mathrm{d}^d k}{(2\pi)^d}\frac{\xi^2}{\left(k^2+\xi m^2\right)^2}\left[W^{(2)}(p+k) - W^{(2)}(p)\right]. \tag{18.83}$$

Two remarks are now in order: in perturbation theory only the second term in the r.h.s. is divergent. Moreover, only the second term has a pole. Therefore, on the fermion mass-shell ($p^2 \to -M^2$), the only effect of the change of the gauge fixing term is a fermion field renormalization: the physical mass M is gauge independent (a gauge dependence of the mass would have been led to a double pole at $p^2 = -M^2$). From equation (18.83) we deduce the additional renormalization needed when changing ξ:

$$\frac{\partial \ln Z_\psi}{\partial \xi} = -\frac{e^2}{(2\pi)^d} \int \frac{\mathrm{d}^d k}{(k^2 + \xi m^2)^2}, \tag{18.84}$$

and, therefore, integrating over ξ,

$$Z_\psi(\xi) = Z_\psi(0) \exp\left[-\frac{\xi e^2}{(2\pi)^d} \int \frac{\mathrm{d}^d k}{k^2\,(k^2 + \xi m^2)}\right]. \tag{18.85}$$

General correlation functions. We could use the WT identities in the same way to exhibit the ξ-dependence of all correlation functions. However, since the identities (18.74) legitimate at the perturbative level the set of manipulations of Section 18.1, we use the latter to prove a general identity. We add to the action $\mathcal{S}(A, \bar\psi, \psi)$ a new term depending on a scalar field χ,

$$\mathcal{S}_\xi(A, \bar\psi, \psi) \mapsto \mathcal{S}_\xi(A, \bar\psi, \psi, \chi) = \mathcal{S}_\xi(A, \bar\psi, \psi) - \tfrac{1}{2}\xi m^2 \int \mathrm{d}^d x\, \chi^2(x).$$

Integrating over χ, we see that only the, irrelevant, normalization of the partition function has changed. We now shift $\chi(x)$,

$$\chi(x) \mapsto \chi(x) + \frac{1}{\xi m}\partial_\mu A_\mu(x). \tag{18.86}$$

After the change of variables and an integration by parts the action becomes

$$\begin{aligned}\mathcal{S}(A, \bar\psi, \psi) = \mathcal{S}_\xi(A, \bar\psi, \psi) - \tfrac{1}{2}\xi m^2 \int \mathrm{d}^d x\, \chi^2(x) + m \int \mathrm{d}^d x\, A_\mu(x)\partial_\mu\chi(x) \\ - \frac{1}{2\xi}\int \mathrm{d}^d x\, (\partial_\mu A_\mu)^2 .\end{aligned} \tag{18.87}$$

The last term cancels the gauge fixing term, and the second term in the r.h.s. can be eliminated by a gauge transformation:

$$A_\mu \mapsto A_\mu - \partial_\mu\chi/m \quad \text{and} \quad \psi \mapsto \psi\, \mathrm{e}^{ie\chi/m}, \quad \bar\psi \mapsto \bar\psi\, \mathrm{e}^{-ie\chi/m}. \tag{18.88}$$

The functional integral representation of $\mathcal{Z}(J, \bar\eta, \eta)$ then takes the form:

$$\begin{aligned}\mathcal{Z}(J, \bar\eta, \eta) = \int [\mathrm{d}A_\mu \mathrm{d}\bar\psi \mathrm{d}\psi] \exp\left[-\mathcal{S}(A, \bar\psi, \psi) - \tfrac{1}{2}m^2 \int \mathrm{d}^d x\, A_\mu^2\right] \int [\mathrm{d}\chi] \\ \times \exp\left\{\int \mathrm{d}^d x \left[\tfrac{1}{2}(\partial_\mu\chi)^2 + \tfrac{1}{2}\xi m^2\chi^2 - \chi\partial_\mu J_\mu/m + \bar\eta\, \mathrm{e}^{ie\chi/m}\psi + \bar\psi\, \mathrm{e}^{-ie\chi/m}\eta\right]\right\},\end{aligned} \tag{18.89}$$

in which $\mathcal{S}(A,\bar{\psi},\psi)$ is the gauge invariant action. Expanding in powers of η and $\bar{\eta}$, we can integrate over χ. For example, in the case of the $\bar{\psi}\psi$ two-point function we obtain the ratio of bare functions corresponding to two different values of ξ:

$$W_{\xi}^{(2)}(x,y)=\exp\left[\frac{\xi e^2}{(2\pi)^d}\int \mathrm{d}^d k \frac{\mathrm{e}^{ik(x-y)}-1}{k^2\left(k^2+\xi m^2\right)}\right] W_{(\xi=0)}^{(2)}(x,y). \tag{18.90}$$

Introducing the fermion field renormalization $Z_{\psi}(\xi)$, we infer the relation between renormalized correlation functions:

$$W_{\xi}^{(2)}(x,y)=Z_{\psi}^{-1}(\xi)Z_{\psi}(0)\exp\left[\frac{\xi e^2}{(2\pi)^d}\int \mathrm{d}^d k \frac{\mathrm{e}^{ik(x-y)}-1}{k^2\left(k^2+\xi m^2\right)}\right] W_{(\xi=0)}^{(2)}(x,y). \tag{18.91}$$

In particular, with the choice (18.85) for $Z_{\psi}(\xi)$, the relation becomes finite:

$$W_{\xi}^{(2)}(x,y)=\exp\left[\frac{\xi e^2}{(2\pi)^d}\int \frac{\mathrm{d}^d k\ \mathrm{e}^{ik(x-y)}}{k^2\left(k^2+\xi m^2\right)}\right] W_{(\xi=0)}^{(2)}(x,y). \tag{18.92}$$

The equation (18.92) has additional consequences: no other ξ-dependent renormalization has been needed. It follows that the renormalization constant Z_e and thus, from equation (18.81), also Z_A are gauge independent. In the renormalized action (18.79) only Z_{ψ} is gauge dependent.

Other ψ-field correlation functions have a similar structure. The ξ dependence, which comes only from the χ integration, factorizes. If we now consider correlation functions with external gauge fields, the previous analysis still applies provided the gauge field is transverse, that is, the source J_{μ} satisfies $\partial_{\mu}J_{\mu}=0$: transverse external gauge fields do not introduce any new gauge dependence.

Note that we can obtain the gauge dependence of the full gauge field correlation functions by integrating over χ after making a translation of a term proportional to $\partial_{\mu}J_{\mu}(x)$. However, identities (18.74) directly relate correlation functions with external longitudinal gauge fields (in Fourier space $A_{\mu}(k)$ proportional to k_{μ}) to fermion field correlation functions.

Unitarity. Since on the mass-shell the propagator of the gauge field is a transverse projector,

$$\left(\delta_{\mu\nu}+(\xi-1)\frac{k_{\mu}k_{\nu}}{k^2+\xi m^2}\right)k_{\nu}=\frac{\xi\left(k^2+m^2\right)}{k^2+\xi m^2}k_{\mu}=0 \quad \text{for } k^2+m^2=0\,,$$

we have obtained the gauge dependence of all S-matrix elements. In the mass-shell limit, the ξ-dependence of the singular part of correlation functions becomes simply a multiplicative renormalization (see Section A18.2). Therefore, the properly normalized S-matrix elements are gauge independent. Being gauge independent, they cannot have the ξ-dependent singularities which we have introduced with the χ-field to make the theory renormalizable. 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 the averages of products of gauge invariant operators, that is, local polynomials in the field invariant under the transformations (18.24,18.27).

The simplest such operators are $F_{\mu\nu}$ which select 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 (18.92) shows explicitly 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 has to add to the action sources for them. The form of WT identities is not modified. The arguments of Section 18.6 are still valid: the counterterms are gauge invariant. This 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 (18.7) for the gauge fields, all correlation functions are "physical", but the theory is not renormalizable. However, we have been able to construct some correlation functions of the theory, the gauge invariant correlation functions, which have a large cut-off limit. The relation (18.90), in presence of a cut-off, leads in the large ξ limit to an explanation for this surprising property:

$$W^{(2)}_{\infty}(x,y)=\exp\left[-\frac{e^2}{(2\pi)^d}\int \mathrm{d}^d k\frac{1-\mathrm{e}^{ik(x-y)}}{m^2k^2}\right]W^{(2)}_{0}(x,y). \tag{18.93}$$

For $|x-y|\neq 0$, the dominant term in the large cut-off limit in the exponential is

$$\frac{e^2}{(2\pi)^d}\int\frac{\mathrm{d}^d k}{m^2k^2}\sim e^2\frac{S_d}{(2\pi)^d}\frac{\Lambda^{d-2}}{m^2}. \tag{18.94}$$

Therefore, in the physical representation all non-gauge invariant correlation functions vanish. The 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.

18.8 Renormalization Group Equations

In this section we derive RG equations in the case of the action (18.26), 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 1PI correlation functions corresponding to l gauge fields, and n fermion pairs ψ and $\bar{\psi}$. The relation between bare and renormalized correlation functions is

$$\Gamma^{(l,n)}_{\mathrm{B}}(p_i,q_j;\alpha_0,\xi_0,m_0,M_0)=Z_A^{-l/2}Z_\psi^{-n}\Gamma^{(l,n)}(p_i,q_j;\mu,\alpha,\xi,m,M), \tag{18.95}$$

in which μ is the renormalization scale, and we have called α the loop expansion parameter:

$$\alpha=e^2/4\pi\,. \tag{18.96}$$

Differentiating equation (18.95) with respect to μ at bare parameters fixed, we find the RG equations:

$$\Bigg[\mu\frac{\partial}{\partial\mu}+\beta(\alpha)\frac{\partial}{\partial\alpha}+\delta(\alpha)\xi\frac{\partial}{\partial\xi}+\eta_m(\alpha)m\frac{\partial}{\partial m}+\eta_M(\alpha)M\frac{\partial}{\partial M} \\ -\frac{l}{2}\eta_A(\alpha)-n\eta_\psi(\alpha)\Bigg]\Gamma^{(l,n)}(p_i,q_j;\mu,\alpha,\xi,m,M)=0\,. \tag{18.97}$$

The equation (18.81) relates Z_A and Z_e, the gauge field and coupling constant renormalization constants,

$$Z_AZ_e=1\,.$$

Therefore,

$$\alpha=Z_A\alpha_0\,. \tag{18.98}$$

Moreover, we have shown in Section 18.6 that the parameters m and ξ are not renormalized. It follows that,

$$m_0^2=m^2Z_A^{-1},\qquad \xi_0=\xi Z_A\,. \tag{18.99}$$

Finally, in Section 18.7, we have shown that the renormalization constant Z_A can be chosen to be independent of ξ (the minimal subtraction scheme satisfies this requirement). The equations (18.98,18.99) then imply a set of relations between RG functions:

$$\beta(\alpha)=\eta_A(\alpha), \tag{18.100}$$

$$\delta(\alpha)=-\beta(\alpha), \tag{18.101}$$

$$\eta_m(\alpha)=\beta(\alpha)/2\,. \tag{18.102}$$

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 (18.85) it is even possible to determine the gauge dependence of η_ψ. A short calculation leads in the minimal subtraction scheme to

$$\eta_\psi(\alpha,\xi)=\eta_\psi(\alpha,0)-\alpha\xi/2\pi\,. \tag{18.103}$$

18.9 The One-Loop β-Function

Fermion contribution. We now calculate the β-function at one-loop order in the case of the action (18.26) which describes the interaction of a gauge field with charged fermions. We first evaluate the gauge field renormalization constant and then deduce the coupling constant renormalization from relation (18.81). Dimensional regularization will be used in the calculation.

The one-loop contribution to the generating functional of proper vertices coming from the fermion integration is

$$\Gamma_{1\,\text{loop}}(A_\mu)=-\operatorname{tr}\ln\left(\not\partial+ie\not A+M\right). \tag{18.104}$$

Differentiating twice with respect to A_μ, we obtain the one-loop contribution to the renormalized 1PI two-point function:

$$\Gamma^{(2)}_{\mu\nu}(p)=Z_A\left(\delta_{\mu\nu}p^2-p_\mu p_\nu\right)+p_\mu p_\nu/\xi+\Sigma_{\mu\nu}(p)\,, \tag{18.105}$$

with

$$\Sigma_{\mu\nu}(p) = e^2 \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\operatorname{tr}\left[\gamma_\mu \left(\not{k} + iM\right) \gamma_\nu \left(\not{k} - \not{p} + iM\right)\right]}{(k^2 + M^2)\left[(p-k)^2 + M^2\right]}.$$

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

$$m \frac{\mathrm{d}}{\mathrm{d}m} \int \frac{\mathrm{d}^d k}{k^2 + m^2} = (d-2) \int \frac{\mathrm{d}^d k}{k^2 + m^2} = -2m^2 \int \frac{\mathrm{d}^d k}{(k^2 + m^2)^2}, \tag{18.106}$$

we obtain,

$$\Sigma_{\mu\nu}(p) = 2e^2 \operatorname{tr} \mathbf{1} (p^2 \delta_{\mu\nu} - p_\mu p_\nu) \int_0^1 \mathrm{d}s\, s(1-s) \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{\left[k^2 + s(1-s)p^2 + M^2\right]^2}.$$

In particular, the divergent part of $\Sigma_{\mu\nu}(p)$ is

$$\Sigma_{\mu\nu}(p) = e^2 \operatorname{tr} \mathbf{1} \frac{N_d}{3\varepsilon} (p^2 \delta_{\mu\nu} - p_\mu p_\nu) + O(1), \tag{18.107}$$

in which N_d is the usual loop factor. This determines the A_μ field renormalization Z_A and thus, from equation (18.81), also Z_e:

$$Z_A = 1 - N_d \operatorname{tr} \mathbf{1} \frac{e^2}{3\varepsilon} + O\left(e^4\right), \tag{18.108}$$

$$Z_e = 1 + N_d \operatorname{tr} \mathbf{1} \frac{e^2}{3\varepsilon} + O\left(e^4\right). \tag{18.109}$$

Replacing N_d and $\operatorname{tr} \mathbf{1}$ by their values for $d = 4$ ($N_4 = 1/8\pi^2$, $\operatorname{tr} \mathbf{1} = 4$), we obtain the corresponding β-function at one-loop order:

$$\beta(\alpha) = -\varepsilon \left[\frac{\mathrm{d} \ln (\alpha Z_e)}{\mathrm{d}\alpha}\right]^{-1} = -\varepsilon\alpha + \frac{4}{3} \frac{\alpha^2}{2\pi} + O\left(\alpha^3\right). \tag{18.110}$$

Scalar boson contribution. We now consider the action

$$S(A_\mu, \phi) = \int \mathrm{d}^d x \left(\frac{1}{4} F_{\mu\nu}^2 + \left|\mathrm{D}_\mu \phi\right|^2 + M^2 \left|\phi\right|^2 + \frac{1}{6} g \left|\phi\right|^4\right). \tag{18.111}$$

Again we calculate the one-loop contribution to the gauge field propagator. An alternative calculation of the divergent part is given in Appendix A18.3. Since there are now two interaction terms of the form $\phi^2 A_\mu^2$ and $\phi^2 A_\mu$, two diagrams contribute at one-loop order (see figure 18.3).

Fig. 18.3 Charged bosons: the two one-loop diagrams contributing to the A_μ two-point function.

The two terms can also be obtained from the expansion of

$$\Gamma_{1\,\mathrm{loop}}(A_\mu) = \operatorname{tr}\ln\left[(\partial_\mu + ieA_\mu)^2 - M^2\right]\left(\partial^2 - M^2\right)^{-1}. \tag{18.112}$$

One finds:

$$\Gamma^{(2)}_{\mu\nu,1\,\mathrm{loop}} = 2e^2\delta_{\mu\nu}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{k^2+M^2} - e^2\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{(2k-p)_\mu(2k-p)_\nu}{(k^2+M^2)\left[(k-p)^2+M^2\right]}\,. \tag{18.113}$$

One again verifies that the result is transverse. Introducing Feynman parameters, and playing with the same identity (18.106), one finds

$$\Gamma^{(2)}_{\mu\nu,1\,\mathrm{loop}} = e^2(p^2\delta_{\mu\nu} - p_\mu p_\nu)\int_0^1 \mathrm{d}s(1-2s)^2\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{\left[k^2+s(1-s)p^2+M^2\right]^2}.$$

In particular, the divergent part is

$$\Gamma^{(2)}_{\mu\nu,1\,\mathrm{loop}} = \frac{N_d}{3\varepsilon}e^2(p^2\delta_{\mu\nu} - p_\mu p_\nu) + O(1). \tag{18.114}$$

The last part of the calculation is the same as in the fermion case. The difference between fermions and bosons comes only 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. Calling n_F and n_B the number of charged fermions and scalar bosons respectively, we finally obtain,

$$\beta(\alpha) = -\varepsilon\alpha + \frac{1}{3}\left(4n_\mathrm{F} + n_\mathrm{B}\right)\frac{\alpha^2}{2\pi} + O\left(\alpha^3\right). \tag{18.115}$$

The sign of the β-function. Note that in the domain of validity of expansion (18.108) (e^2 small), Z_A satisfies

$$Z_A \le 1\,.$$

This property also holds in the ϕ^4-field theory. The field renormalization constant Z_ϕ satisfies (equation (11.69)) $Z_\phi \le 1$. Both are consequences of the Källen–Lehmann representation for the two-point function (see Section 6.9). In the case of the gauge field the property is true because Z_A is related to the transverse part of the two-point function to which unphysical 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 for α small enough.

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, $\det(\not\partial + ie\not A + M)$, which is generated by the integration over the fermion fields, and use the property of the charge conjugation matrix C introduced in Section A8.1.7:

$$\begin{aligned}\det\left(\not\partial + ie\not A + M\right) &= \det\left[{}^T\!\left(\not\partial + ie\not A + M\right)\right] = \det C^{-1}\left[{}^T\!\left(\not\partial + ie\not A + M\right)\right]C\\ &= \det\left(\not\partial - ie\not A + M\right).\end{aligned}$$

Therefore, the interaction between gauge fields generated by the fermions is even in A_μ. Note in particular the implication for Feynman diagrams: fermion loops with an odd number of external gauge fields can be omitted.

18.10 The Abelian Higgs Model

As an introduction to Chapter 19 we now discuss the physics of a gauge field coupled to a charged scalar field $\phi(x)$ in an unusual phase. We start from the action (18.33). Renormalizability implies that the scalar field self-interaction is of the $|\phi|^4$ type:

$$\mathcal{S}(A_\mu, \phi) = \int \mathrm{d}^d x \left(\frac{1}{4} F_{\mu\nu}^2 + |\mathrm{D}_\mu \phi|^2 + M^2 |\phi|^2 + \frac{1}{6} g |\phi|^4 \right). \tag{18.116}$$

The field $\phi(x)$ is complex and the covariant derivative D_μ is defined by equation (18.28).

In the classical limit, if the gauge symmetry is exact, and the $U(1)$ symmetry is not spontaneously broken, the gauge field is massless and the scalar field ϕ has two real components with equal mass.

However, we know that another phase is possible in which the $U(1)$ symmetry is spontaneously broken, and $\phi(x)$ has a non-vanishing expectation value v, which for convenience we assume real (we shall comment later on the significance of $\langle\phi\rangle$ which is not gauge invariant). This is a situation we have discussed in Section 13.4 and we have concluded that the SSB of a continuous symmetry implies the presence of a massless state, a Goldstone particle. This result can be derived, in the classical limit, by parametrizing the field ϕ as (see Section 13.4)

$$\phi(x) = \frac{1}{\sqrt{2}} [v + \rho(x)] \, \mathrm{e}^{i\theta(x)} . \tag{18.117}$$

As a consequence of the symmetry the resulting action then depends only on $\partial_\mu \theta$, the field θ, therefore, is massless.

In a gauge theory, however, the transformation $\phi(x) \mapsto v + \rho(x)$ has the form of a gauge transformation. If we perform the corresponding transformation on the gauge field $A_\mu(x)$,

$$A_\mu(x) = A'_\mu(x) + \frac{1}{e} \partial_\mu \theta(x), \tag{18.118}$$

we eliminate the field θ completely from the action. After this transformation the action $\mathcal{S}(A_\mu, \phi)$ indeed becomes

$$\mathcal{S}(A_\mu, \rho) = \int \mathrm{d}^d x \left[\frac{1}{4} F_{\mu\nu}^2 + \frac{1}{2} (\partial_\mu \rho)^2 + \frac{1}{2} e^2 A_\mu^2 (\rho + v)^2 + \frac{1}{2} M^2 (\rho + v)^2 + \frac{g}{24} (\rho + v)^4 \right]. \tag{18.119}$$

In the tree approximation the spectrum of the theory thus contains a massive vector particle of mass $e^2 v^2$ and a massive scalar called the Higgs particle:

$$m^2(A) = e^2 v^2, \qquad m^2(\rho) = \tfrac{1}{3} g v^2. \tag{18.120}$$

As a consequence of gauge invariance, no Goldstone field has been generated. This is a most remarkable property, which is also at the basis of the Meissner effect in superconductivity. It is induced by the long range forces (in the non-relativistic model the electro-magnetic force) generated by the massless gauge field.

Note that the total number of physical degrees of freedom has not changed between the symmetric phase and the spontaneously broken phase since one degree of freedom of the scalar field has been transferred to the vector field.

From the technical point of view, the action (18.119) has a surprising property: in this so-called unitary gauge (the representation (18.119)), the theory contains only physical fields, can be trivially quantized but is not renormalizable. On the other hand, if we start from the action (18.116) and quantize it in the same way as in the symmetric phase, the theory will contain unphysical degrees of freedom, but we expect it to be renormalizable. We are reminded of the massive vector field coupled to a conserved current, discussed in the first part of the chapter. There exists actually a relation between the massive vector field and the Higgs model: if we take the formal non-linear σ-model limit of the action (18.116), that is, a limit in which the bare mass of the Higgs field becomes infinite at fixed expectation value v, we recover the action (18.7) with the identification $m = ev$.

In order to be able to calculate gauge invariant observables and S-matrix elements, we, therefore, return to the action (18.116). We fix the gauge by adding a term proportional to $(\partial_\mu A_\mu)^2$. This amounts to couple the phase field $\theta(x)$ which plays the role of the $\Lambda(x)$ field of Section 18.1.

As a final remark we note for later purposes that the mechanism of spontaneous symmetry breaking can also be used to give a mass to fermions in a chiral invariant theory (Section 13.6).

18.11 Quantization of the Higgs Model

We now start from the action

$$\mathcal{S}(A,\phi) = \int \mathrm{d}^d x \left[\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2\xi}\left(\partial_\mu A_\mu\right)^2 + \frac{1}{2}m^2 A_\mu^2 + |\mathrm{D}_\mu\phi|^2 + M^2|\phi|^2 + \frac{g}{6}|\phi|^4\right], \tag{18.121}$$

in which a mass has been given to the vector field to provide an IR cut-off.

We assume that ϕ has a real expectation value v at the classical level. We introduce the real and imaginary parts of ϕ and set

$$\phi(x) = \frac{1}{\sqrt{2}}\left[v + \varphi(x) + i\chi(x)\right]. \tag{18.122}$$

The quadratic part $\mathcal{S}_2$ of the action is then

$$\begin{aligned}\mathcal{S}_2(A,\phi) = \int \mathrm{d}^d x \Big[&\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2\xi}\left(\partial_\mu A_\mu\right)^2 + \frac{1}{2}\left(e^2v^2 + m^2\right)A_\mu^2 \\ &-ev\chi\partial_\mu A_\mu + \frac{1}{2}\left(\partial_\mu\varphi\right)^2 + \frac{1}{6}gv^2\varphi^2 + \frac{1}{2}\left(\partial_\mu\chi\right)^2\Big].\end{aligned} \tag{18.123}$$

We see that $\partial_\mu A_\mu$ is coupled to the Goldstone field χ. The corresponding propagators are

$$\begin{aligned} W^{(2)}_{\mu\nu}(k) &= \frac{\delta_{\mu\nu} - k_\mu k_\nu/k^2}{k^2 + e^2v^2 + m^2} + \xi\frac{k_\mu k_\nu}{k^2(k^2+\xi m^2)}, \\ W^{(2)}_{\mu\chi}(k) &= -\xi\frac{iev k_\mu}{k^2(k^2+\xi m^2)}, \\ W^{(2)}_{\chi\chi}(k) &= \frac{1}{k^2} + \frac{\xi e^2 v^2}{k^2(k^2+\xi m^2)}, \end{aligned} \tag{18.124}$$

and

$$W^{(2)}_{\varphi\varphi}(k) = \frac{1}{k^2 + gv^2/3}. \tag{18.125}$$

The spectrum of the theory contains three physical states and the usual state with negative norm coming from the regulator. We see that in the absence of a mass term for the vector field in the action (18.121), 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 corresponds really to a Goldstone mode. Even in the physical gauge a massless field is present and coupled.

18.11.1 WT identities and renormalization

It follows from the combined analysis of Chapter 13 and Section 18.6 that after renormalization the correlation functions satisfy the equivalent of WT identities (18.74) and (18.77). As a consequence the dependence of correlation functions on the parameter ξ can be determined as in Section 18.7. In particular, only correlation functions of gauge invariant operators and S-matrix elements are gauge independent.

The explicit form of the WT for correlation functions is now rather complicated. We write here only the identities corresponding to the (A_μ, χ) two-point proper vertices. Calling v the expectation value of the renormalized φ-field, we obtain by differentiating equation (18.77) with respect to A_μ:

$$k_\nu \Gamma^{(2)}_{\mu\nu}(k) + iev\,\Gamma^{(2)}_{\mu\chi}(k) = k_\mu \left(k^2/\xi + m^2\right). \tag{18.126}$$

Differentiating then with respect to χ, we find,

$$k_\mu \Gamma^{(2)}_{\mu\chi} - iev\,\Gamma^{(2)}_{\chi\chi} = 0. \tag{18.127}$$

We parametrize the different functions as

$$\Gamma^{(2)}_{\mu\nu}(k) = a(k^2)\delta_{\mu\nu} - b(k^2)k_\mu k_\nu, \quad \Gamma^{(2)}_{\mu\chi}(k) = iev\,c(k^2)k_\mu, \quad \Gamma^{(2)}_{\chi\chi}(k) = d(k^2). \tag{18.128}$$

In the tree approximation the values of a, b, c and d are

$$\begin{cases} a(k^2) = e^2v^2 + m^2 + k^2, & b(k^2) = 1 - 1/\xi, \\ c(k^2) = 1, & d(k^2) = k^2. \end{cases} \tag{18.129}$$

We now express the identity (18.127),

$$d(k^2) = k^2 c(k^2). \tag{18.130}$$

The identity (18.126) then leads to

$$a(k^2) - k^2 b(k^2) - e^2v^2 c(k^2) = k^2/\xi + m^2. \tag{18.131}$$

In the $k = 0$ limit, this equation in particular implies

$$a(0) - e^2v^2 c(0) = m^2. \tag{18.132}$$

The corresponding connected correlation functions are

$$\begin{aligned} W^{(2)}_{\mu\nu}(k) &= \frac{1}{a}\left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) + \frac{\xi k_\mu k_\nu}{k^2(k^2 + \xi m^2)}, \\ W^{(2)}_{\mu\chi}(k) &= -\frac{\xi iev k_\mu}{k^2(k^2 + \xi m^2)}, \\ W^{(2)}_{\chi\chi}(k) &= \frac{1}{ck^2}. \end{aligned} \tag{18.133}$$

RG β functions. For completeness and later purpose we give here the RG β-functions at one-loop for $d = 4$ in a more general model with N charged scalars:

$$\beta_g = \frac{1}{24\pi^2}\left[(N+4)g^2 - 18ge^2 + 54e^4\right], \quad \beta_{e^2} = \frac{1}{24\pi^2}Ne^4. \tag{18.134}$$

The origin $e^2 = g = 0$ is a stable IR fixed point only for $N \geq 183$.

18.11.2 Decoupling gauge

The quantization method we have used above 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, gets rid of the Goldstone modes), it is possible to circumvent this difficulty. In the notation of the action (18.121) we impose the condition:

$$\partial_\mu (B_\mu + \partial_\mu \Lambda) + \lambda e v \,\mathrm{Im}\left(\phi\, \mathrm{e}^{-ie\Lambda}\right) = \xi^{1/2} \nu(x), \tag{18.135}$$

in which λ is a constant which will be adjusted eventually. The important new feature is that the operator $\mathbf{M}$, functional derivative of equation (18.135) with respect to Λ, now depends on the dynamical fields:

$$\langle y | \mathbf{M} | x \rangle = \left[\partial^2 + \lambda e v \,\mathrm{Re}\left(\phi\, \mathrm{e}^{-ie\Lambda}\right)\right] \delta(x-y), \tag{18.136}$$

and the associated determinant $\det \mathbf{M}$ is no longer a constant. This is the source of difficulties of a kind we have already encountered in Chapters 16 and 17: we have to introduce spinless fermion fields to write $\det \mathbf{M}$ in local form, and then use the induced BRS symmetry to show that renormalization preserves the form of the action. We postpone this analysis until Chapter 21, and discuss here only the tree approximation.

As before we integrate over $\nu(x)$ with the distribution (18.56) and use the gauge invariance of the initial action:

$$\mathcal{S}_{\mathrm{sym}}(B_\mu, \phi) = \mathcal{S}_{\mathrm{sym}}(A_\mu, \phi\, \mathrm{e}^{-ie\Lambda}). \tag{18.137}$$

Changing then variables in the functional integral $\phi\, \mathrm{e}^{-ie\Lambda} \mapsto \phi$, we obtain the quantized action:

$$\begin{aligned} \mathcal{S}_{\mathrm{qu}}(A_\mu, \phi, \bar{C}, C) = \int \mathrm{d}^d x \left\{ \frac{1}{4} F_{\mu\nu}^2 + \frac{1}{2\xi} \left(\partial_\mu A_\mu + \lambda\, \mathrm{e}\, v \,\mathrm{Im}\, \phi\right)^2 + \left|D_\mu \phi\right|^2 \right. \\ \left. + M^2 \left|\phi\right|^2 + \frac{1}{6} g \left|\phi^4\right| - \bar{C} \left(\partial^2 + \lambda\, \mathrm{e}\, v \,\mathrm{Re}\, \phi\right) C \right\}, \end{aligned} \tag{18.138}$$

in which C and $\bar{C}$ are two scalar fermion fields which generate the determinant $\det \mathbf{M}$. As we have explained on an example in Section 8.5, scalar fermions cannot be interpreted as physical particles but are of a nature similar to Pauli–Villars regulator fields. We now use the parametrization (18.122) and choose λ:

$$\lambda = \xi \sqrt{2}\,. \tag{18.139}$$

This is of course the value only at leading order. The propagators are then

$$\begin{aligned} W^{(2)}_{\mu\nu} &= \frac{\delta_{\mu\nu}}{k^2 + e^2 v^2} + \frac{(\xi - 1) k_\mu k_\nu}{(k^2 + e^2 v^2)(k^2 + \xi e^2 v^2)}, \\ W^{(2)}_{\chi\chi} &= \frac{1}{k^2 + \xi e^2 v^2}, \\ W^{(2)}_{\bar{C}C} &= \frac{1}{k^2 + \xi e^2 v^2}. \end{aligned} \tag{18.140}$$

The advantages of this gauge (introduced by 't Hooft) are that by construction there is no $A_\mu \chi$ propagator and that all unphysical 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 BRS symmetry. We shall examine this question in Chapter 21 in detail.

18.12 Physical Observables. Unitarity of the S-Matrix

The unphysical 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 18.7, or directly by using the whole set of WT identities and showing explicitly that the pole coming from $W^{(2)}_{\mu\nu}$ cancels the contribution coming from $W^{(2)}_{\mu\chi}$ in the intermediate state in unitarity relations. As the expressions (18.133) 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.

In the limit $m = 0$, we expect also the pole at $k^2 = 0$ to cancel in physical observables. According to relation (18.132), for $k^2 \to 0$ the different propagators behave like

$$\begin{aligned} W^{(2)}_{\mu\nu} &\sim \frac{k_\mu k_\nu}{k^2}\left(\frac{1}{m^2} - \frac{1}{m^2 + e^2 v^2 c(0)}\right), \\ W^{(2)}_{\mu\chi} &\sim -iev\frac{k_\mu}{k^2 m^2}, \\ W^{(2)}_{\chi\chi} &\sim \frac{1}{c(0)k^2}. \end{aligned} \tag{18.141}$$

Again a direct argument based on WT identities for connected correlation functions and unitarity relations allows to prove that in the $m^2 = 0$ limit the χ-field decouples from physical observables. Here we do not have an alternative proof based upon gauge dependence. However, we shall construct one in a more general context by using a different gauge.

Gauge invariant operators. When the action (18.121) is used, only averages of gauge invariant operators are physical correlation functions. The simplest examples are $F_{\mu\nu}$ and $\sigma(x) = (v + \rho(x))^2 = \frac{1}{2}\phi(x)\phi^*(x)$. In particular, the parameter v is really the square root of $\langle\sigma\rangle$. On the other hand, the action (18.119) can be rewritten in terms of this field $\sigma(x)$. One attractive feature of this representation is that the measure of integration in the functional integral is just the flat measure $[\mathrm{d}\sigma(x)]$. 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 $\sigma(x)$ fields does not provide a direct explanation for the finiteness, after the introduction of finite number of counterterms, of the correlation functions of $F_{\mu\nu}$ and σ.

18.13 Stochastic Quantization: The Example of Gauge Theories

As suggested by Parisi stochastic dynamic equations can 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 this 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.

Let us explain the idea of stochastic quantization in the simplest example of the abelian gauge field without matter. 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 17):

$$\dot{A}_\mu(t,x) = -\tfrac{1}{2}\Omega\partial_\nu F_{\mu\nu}(t,x) + N_\mu(t,x), \tag{18.142}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,,$$

and $N_\mu(x,t)$ is a gaussian noise with measure

$$[\mathrm{d}\rho(N)] = [\mathrm{d}N_\mu]\exp\left[-\int \mathrm{d}^d x\,\mathrm{d}t\,N_\mu^2(x,t)/2\Omega\right]. \tag{18.143}$$

The equation is clearly invariant under time-independent gauge transformations

$$A_\mu(x,t) \mapsto A_\mu(x,t) + \partial_\mu\Lambda(x),$$

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 $\mathrm{e}^{-\mathcal{A}}$, 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,

$$\begin{aligned} A_\mu(t,k) = {} & \left(\delta_{\mu\nu} - k_\mu k_\nu/k^2\right)\mathrm{e}^{-\Omega k^2 t/2}\left[A_\nu(0,k) + \int_0^t \mathrm{e}^{\Omega k^2 t'/2}\,N_\nu(t',k)\mathrm{d}t'\right] \\ & + \left(k_\mu k_\nu/k^2\right)\left[A_\nu(0,k) + \int_0^t N_\nu(t',k)\mathrm{d}t'\right]. \end{aligned}$$

We immediately notice 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. Calling A_L and A_T the component of A_μ respectively along and perpendicular to k_μ, we find,

$$\begin{aligned} \langle A_\mathrm{T}(t,k)A_\mathrm{T}(t,-k)\rangle &= \frac{1}{k^2} + O\left(\mathrm{e}^{-\Omega k^2 t}\right), \\ \langle A_\mathrm{L}(t,k)A_\mathrm{L}(t,-k)\rangle &= A_\mathrm{L}(0,k)A_\mathrm{L}(0,-k) + \Omega t\,. \end{aligned}$$

We conclude that, due to gauge invariance, the Langevin equation (18.142) does not generate an equilibrium distribution, though 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 is interacting with matter, and in the non-abelian case (see Chapter 19).

The next problems are the formal relations with the standard quantization procedure by gauge fixing as described before in the chapter and in Chapter 19, and the problems of power counting and renormalization.

The gauge field propagator obtained from the effective dynamic action in the Fourier representation reads

$$\Delta_{\mu\nu}(\omega,k) = \frac{\delta_{\mu\nu} - k_\mu k_\nu/k^2}{\omega^2 + \Omega k^4/4} + \frac{k_\mu k_\nu}{k^2\omega^2}.$$

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.

We see also that the longitudinal propagator does not decrease at large momentum for fixed ω. An analogous problem appears in the quantization with non-covariant gauges: the theory is not renormalizable by power counting. One solution to this problem is to add to the Langevin equation a non-conservative drift force of the form $\mathrm{D}_\mu V(A)$ where D_μ is the covariant derivative, and $V(A)$ a linear function of A, for example $\partial_\mu A_\mu$. It is easy to verify that such a term does not contribute to the evolution equations (equation (4.35)) for equal-time gauge invariant correlation functions, and thus these functions are not modified. However, with this term the Langevin equation is no longer gauge invariant, an equilibrium distribution is generated and with a suitable covariant choice of $\mathrm{D}_\mu V(A)$ the theory is renormalizable by power counting. It can also be shown that it is equivalent to the theories obtained by canonical quantization procedure, but with a non-local gauge fixing term.

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. The hope of course remains that the method will help to quantize other systems where no alternative method has yet been found.

Bibliographical Notes

Many early articles about QED are reprinted in

J. Schwinger ed., *Quantum Electrodynamics* (Dover, New York 1958).

There exists of course a vast literature devoted to QED. Among the classical textbooks let us quote

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); C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York 1980).

and references therein.

For infrared divergences see for, example,

D.R. Yennie, S.C. Frautschi and H. Suura, *Ann. Phys. (NY)* 13 (1961) 379; T. Kinoshita, *J. Math. Phys.* 3 (1962) 650; T.D. Lee and M. Nauenberg, *Phys. Rev.* 133B (1964) 1549.

WT identities are derived in

J.C. Ward, *Phys. Rev.* 78 (1950) 182; Y. Takahashi, *Nuovo Cimento* 6 (1957) 371.

Quantum Electrodynamics still provides the most accurate tests of Quantum Field Theory. For example, the calculation of the eighth order contribution to the electron anomalous magnetic moment has been completed:

T. Kinoshita and W.B. Lindquist, *Phys. Rev.* D42 (1990) 636.

For the muon magnetic moment, which is more sensitive to Strong and Weak Interaction corrections, see

T. Kinoshita, B. Nizic and Y. Okamoto, *Phys. Rev.* D41 (1990) 593; T. Kinoshita, *Phys. Rev.* D47 (1993) 5013.

The abelian Higgs mechanism is discussed in

B.W. Lee, *Phys. Rev.* D5 (1972) 823.

A general class of gauges is constructed in

G. 't Hooft, *Nucl. Phys.* B35 (1971) 167.

The statistical physics analogue of the abelian Higgs mechanism is superconductivity. For the theory of superconductivity see

P.G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York 1966); D.R. Tilley and J. Tilley, *Superfluidity and Superconductivity* (Adam Hilger, Bristol 1986).

Application of stochastic quantization to gauge theories is discussed in

G. Parisi and Y.S. Wu, *Acad. Sci. Sin.* 24 (1981) 483; D. Zwanziger, *Nucl. Phys.* B192 (1981) 259; B209 (1982) 336; J. Zinn-Justin and D. Zwanziger, *Nucl. Phys.* B295 [FS21] (1987) 297.

The corresponding RG β-function is given by

K. Okano, *Nucl. Phys.* B289 (1987) 109.

See also the review

P.H. Damgaard and H. Hüffel, *Phys. Rep.* 152 (1987) 227.

Effects of vacuum fluctuations are discussed in

H.G.B. Casimir, *Proc. Kon. Ned. Akad. Wetenschap.* B51 (1948) 793.

APPENDIX A18

A18.1 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.

A18.1.1 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 we obtain 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 Fourier variables,

$$-\ln\mathcal{Z}=\tfrac{1}{2}\operatorname{tr}\ln\left[(k^2+m^2)\delta_{\mu\nu}-k_\mu k_\nu\right]=\tfrac{1}{2}(d-1)\operatorname{tr}\ln(k^2+m^2)+\tfrac{1}{2}\operatorname{tr}\ln m^2.\quad (A18.1)$$

Up to an irrelevant constant we find $(d-1)$ times the vacuum energy of a free scalar field of mass m, a result which is not surprising since the massive vector field has $d-1$ degrees of freedom. The result can be verified directly by diagonalizing the hamiltonian (18.6).

If we repeat the calculation with the action $\mathcal{S}_\xi$ of equation (18.19) we find instead

$$-\ln\mathcal{Z}=\tfrac{1}{2}(d-1)\operatorname{tr}\ln(k^2+m^2)+\tfrac{1}{2}\operatorname{tr}\ln(k^2+\xi m^2)/\xi\,.$$

This gauge-dependent result for the vacuum energy is clearly incorrect. The reason can be simply understood. In the various algebraic manipulations which have led to the action (18.19) 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 (18.14). Normalizing correctly the χ integral cancels the additional unwanted factor.

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 (18.44) (in a vanishing source). The hamiltonian is partially diagonalized by a Fourier transformation (see Section 6.6.2)

$$\mathcal{H}=\tfrac{1}{2}\left[(2\pi)^{1-d}\tilde{E}_i(\hat{p})\tilde{E}_i(-\hat{p})+(2\pi)^{d-1}\tilde{A}_i(\hat{p})\left(\hat{p}^2\delta_{ij}-\hat{p}_i\hat{p}_j\right)\tilde{A}_j(-\hat{p})\right],\qquad (A18.2)$$

where $\hat{p}$ is a space momentum. We immediately see 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 quantum mechanical 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 $d-2$ times the vacuum energy of a free massless scalar particle.

Covariant gauge. Evaluating instead the energy in a covariant gauge, using the corresponding partition function, naively we find

$$-\ln \mathcal{Z} = \tfrac{1}{2}\operatorname{tr}\ln\left[k^2\delta_{\mu\nu} - k_\mu k\nu(1-1/\xi)\right] = \tfrac{1}{2}d\operatorname{tr}\ln k^2 - \tfrac{1}{2}\operatorname{tr}\ln\xi\,.$$

Up to constant we find 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 18.4.4, if we want the multiplication by the integral over $\Lambda(x)$ to be an identity, we have to multiply the integral (18.63) by $\det(-\partial^2)$. 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.

A18.1.2 Casimir effect

In Chapter 37 we discuss quantum field theory in a finite volume or more generally in restricted geometries. However, the simplest example of finite size effects is the Casimir effect, that is, the attractive force between two parallel perfectly conducting plates in the vacuum, due to the change in vacuum energy produced by a change in boundary conditions. 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, we thus integrate over fields which have the Fourier representation,

$$A_\mu(\mathbf{x}_\perp, z) = \int \mathrm{d}^{d-1}p_\perp \sum_{n\geq 1} \mathrm{e}^{i\mathbf{p}_\perp\cdot\mathbf{x}_\perp}\sin(\pi z/L)\tilde{A}(\mathbf{p}_\perp, n), \qquad (A18.3)$$

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 $(d-2)$, that is, 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 action for a scalar field $\phi(x)$:

$$\mathcal{S}(\phi) = \frac{1}{2}\int \mathrm{d}^d x\left(\partial_\mu\phi(x)\right)^2,$$

where the physical situation of 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 know that the vacuum energy $\mathcal{E}$ is simply given by

$$L_\perp \mathcal{E} = \sum_{n\geq 1}\sum_{\mathbf{n}_\perp} \frac{1}{2}\ln\left(\frac{4\pi^2\mathbf{n}_\perp^2}{L_\perp^2} + \frac{\pi^2 n^2}{L^2}\right). \qquad (A18.4)$$

This expression is only meaningful in the presence of a UV cut-off. Since

$$\int_{\varepsilon}^{\infty} \frac{\mathrm{d}t}{t}\, \mathrm{e}^{-st} \underset{\varepsilon\to 0}{=} -\ln(\varepsilon s) + \psi(1) + O(\varepsilon),$$

we replace it by the regularized form:

$$L_\perp \mathcal{E} = -\sum_{n\geq 1}\sum_{\mathbf{n}_\perp} \frac{1}{2}\int_{a^2}^{\infty} \frac{\mathrm{d}t}{t} \exp\left[-t\left(\frac{4\pi^2 \mathbf{n}_\perp^2}{L_\perp^2} + \frac{\pi^2 n^2}{L^2}\right)\right],$$

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:

$$L_\perp \mathcal{E} = -\frac{1}{2}\sum_{n\geq 1} \int_{a^2}^{\infty} \frac{\mathrm{d}t}{t} \int \mathrm{d}^{d-1}\mathbf{n}_\perp \exp\left[-t\left(\frac{4\pi^2 \mathbf{n}_\perp^2}{L_\perp^2} + \frac{\pi^2 n^2}{L^2}\right)\right].$$

The $\mathbf{n}_\perp$ integral is gaussian. We then obtain the energy per unit area

$$\mathcal{E}/L_\perp^{d-2} = -\tfrac{1}{2}(4\pi)^{(1-d)/2} \int_{a^2}^{\infty} \frac{\mathrm{d}t}{t^{(d+1)/2}} \sum_{n\geq 1} \exp\left(-\pi^2 n^2 t/L^2\right). \qquad (A18.5)$$

We evaluate the sum, using identity ($A37.21$):

$$\sum_{n\geq 1} \exp\left(-\pi^2 n^2 t/L^2\right) = \frac{L}{2\sqrt{\pi t}} - \frac{1}{2} + \frac{L}{\sqrt{\pi t}} \sum_{n\geq 1} \exp\left(-L^2 n^2/t\right).$$

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 it takes the form

$$\left(\mathcal{E} - \mathcal{E}_0\right)/L_\perp^{d-2} = \text{const.} - A(d)L^{1-d}, \qquad (A18.6)$$

with

$$A(d) = \frac{\Gamma(d/2)}{(4\pi)^{d/2}}\zeta(d). \qquad (A18.7)$$

For $d = 4$, one finds

$$A(4) = \frac{\pi^2}{1440}.$$

Casimir effect. The resulting force between plates is, therefore, attractive. 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 the energy and force F per unit area we find, restoring the physical units,

$$\left(\mathcal{E} - \mathcal{E}_0\right)/L_\perp^2 = \text{const.} - \frac{\pi^2}{720}\frac{\hbar c}{L^3}, \quad \Rightarrow\ F = -\frac{1}{L_\perp^2}\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}L} = -\frac{\pi^2}{240}\frac{\hbar c}{L^4}. \qquad (A18.8)$$

This quantum relativistic effect is very small but measurable. It is remarkable because, though electromagnetic in nature, it is independent of the value of the electric charge.

A18.2 Gauge Dependence

To characterize the gauge dependence of a general correlation function with only matter fields (or with transverse gauge field) one has to integrate explicitly over the field χ. Setting

$$W_\xi^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n) = W_\infty^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n) \times U_\xi^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n),$$

we find,

$$U_\xi^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n) = \left\langle \exp\left[\frac{ie}{m}\sum_{i=1,n}\chi(x_i)-\chi(y_i)\right]\right\rangle,$$

where <> means gaussian integration over the χ-field as defined by equation (18.89). The integral can be calculated (for details see, for example, Section 32.1) and yields

$$U_\xi^{(2n)} = \exp\left\{\frac{e^2}{m^2}\sum_{i,j}\left[\tfrac{1}{2}\Delta_\xi(x_i-x_j)+\tfrac{1}{2}\Delta_\xi(y_i-y_j)-\Delta_\xi(x_i-y_j)\right]\right\}, \qquad (A18.9)$$

Δ_ξ being the χ-field propagator:

$$\Delta_\xi(x) = \frac{1}{(2\pi)^d}\int\frac{\mathrm{d}^d k\, \mathrm{e}^{ipx}}{k^2+\xi m^2}. \qquad (A18.10)$$

As indicated in Section 18.7, it is convenient to take the ratio of correlation functions corresponding to two finite values of ξ. The ratio between, for example, the generic ξ and Landau's gauge ($\xi=0$) is then still given by an expression of the form (A18.9) but with the propagator (A18.10) replaced by

$$m^2 K_\xi(x) = \Delta_\xi(x)-\Delta_0(x) = -\frac{\xi m^2}{(2\pi)^d}\int\frac{\mathrm{d}^d k\, \mathrm{e}^{ipx}}{k^2\,(k^2+\xi m^2)}.$$

In expression (A18.9), 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 (18.85):

$$\frac{U_\xi^{(2n)}}{U_0^{(2n)}} = \left(\frac{Z_\psi(\xi)}{Z_\psi(0)}\right)^n \exp\left\{e^2\left[\sum_{i<j}\left(K_\xi(x_i-x_j)+K_\xi(y_i-y_j)\right)-\sum_{i,j}K_\xi(x_i-y_j)\right]\right\}. \qquad (A18.11)$$

If, after renormalization, we expand the ratio between correlation functions in powers of e^2 we note 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. Therefore, when we go on mass-shell these terms do not have the corresponding poles and thus do not contribute to the S-matrix, as indicated in Section 18.7.

A18.3 Divergences at One-Loop with Schwinger's Representation

The one-loop gauge action obtained after integration over the charged scalar fields is

$$\mathcal{S}(A)=\frac{1}{4}\int \mathrm{d}^d x\, F_{\mu\nu}^2+\operatorname{tr}\ln(-\mathrm{D}^2)-\operatorname{tr}(-\partial^2). \qquad (A18.12)$$

To determine the divergent part of the determinant, we use Schwinger's representation and expand for t small the matrix elements of the operator $U=\mathrm{e}^{t\mathrm{D}^2}$,

$$(\partial_\mu+ieA_\mu)^2\langle x|\,U(t)\,|x'\rangle=\partial_t\langle x|\,U(t)\,|x'\rangle\,. \qquad (A18.13)$$

We set

$$\langle x|\,U(t)\,|x'\rangle=\mathrm{e}^{-\sigma(x,x';t)}\,. \qquad (A18.14)$$

Equation ($A18.13$) then takes the form:

$$\nabla^2\sigma-ie\partial_\mu A_\mu-(\partial_\mu\sigma-ieA_\mu)^2=\partial_t\sigma\,. \qquad (A18.15)$$

The function σ has for $t\to 0$ an expansion of the form:

$$\sigma=\frac{1}{4t}(x-x')^2+\frac{d}{2}\ln 4\pi t+\sigma_0+\sigma_1 t+\sigma_2 t^2+O\left(t^3\right). \qquad (A18.16)$$

Therefore,

$$\begin{aligned}\partial_\mu\sigma&=\frac{1}{2t}(x-x')_\mu+\partial_\mu\sigma_0+\partial_\mu\sigma_1 t+\partial_\mu\sigma_2 t^2+O\left(t^3\right)\\ \partial_t\sigma&=-\frac{1}{4t^2}(x-x')^2+\frac{d}{2t}+\sigma_1+2t\sigma_2\,.\end{aligned}$$

It follows

$$\begin{aligned}(\partial_\mu\sigma)^2&=\frac{1}{4t^2}(x-x')^2+\frac{1}{t}(x-x')_\mu\partial_\mu\sigma_0+(\partial_\mu\sigma_0)^2+(x-x')_\mu\partial_\mu\sigma_1\\ &\quad+2t\partial_\mu\sigma_0\partial_\mu\sigma_1+t(x-x')_\mu\partial_\mu\sigma_2\,,\\ \nabla^2\sigma&=\frac{d}{2t}+\nabla^2\sigma_0+t\nabla^2\sigma_1+t^2\nabla^2\sigma_2\,.\end{aligned}$$

The order t^{-1} yields the first non-trivial equation

$$(x-x')_\mu\bigl(\partial_\mu\sigma_0-ieA_\mu(x)\bigr)=0\,.$$

The solution is

$$\sigma_0(x,x')=ie\int_0^1 \mathrm{d}s(x-x')_\mu A_\mu\bigl(x'+s(x-x')\bigr).$$

It has the expected gauge transformations

$$\delta A_\mu(x)=\partial_\mu\theta(x)\quad\Rightarrow\quad\delta\sigma_0(x,x')=ie\bigl(\theta(x)-\theta(x')\bigr).$$

The term of order t^0 yields

$$\sigma_1+(x-x')_\mu\partial_\mu\sigma_1=\partial_\mu\bigl(\partial_\mu\sigma_0-ieA_\mu(x)\bigr)-\bigl(\partial_\mu\sigma_0-ieA_\mu(x)\bigr)^2. \qquad (A18.17)$$

Because

$$\partial_\mu \sigma_0 - ieA_\mu(x) = ie \int_0^1 \mathrm{d}s\, s(x-x')_\nu F_{\mu\nu}\big(x' + s(x-x')\big)$$

is gauge invariant, σ_1 is also gauge invariant. Then,

$$\partial_\mu \big(\partial_\mu \sigma_0 - ieA_\mu(x)\big) = ie \int_0^1 \mathrm{d}s\, s^2 (x-x')_\nu \partial_\mu F_{\mu\nu}\big(x' + s(x-x')\big).$$

Let us solve in general:

$$\sigma_1 + (x-x')_\mu \partial_\mu \sigma_1 = X(x,x') \ \Rightarrow\ \sigma_1(x,x') = \int_0^1 \mathrm{d}s\, X\big(x' + s(x-x'), x'\big).$$

The first term in the r.h.s. of equation ($A18.17$) yields

$$ie \int_0^1 \mathrm{d}s\, s(1-s)(x-x')_\nu \partial_\mu F_{\mu\nu}\big(x' + s(x-x')\big).$$

The second contribution is

$$2e^2 \int u\mathrm{d}u\, v\mathrm{d}v\, (x-x')_\nu (x-x')_\rho F_{\mu\nu}\big(x' + u(x-x')\big) F_{\mu\rho}\big(x + v(x'-x)\big),$$

with $u \ge 0$ $v \ge 0$ and $v + u \le 1$.

We need the last equation coming from the coefficient of t only for $x = x'$. Thus,

$$2\sigma_2 = \nabla^2 \sigma_1 .$$

The contribution to σ_1 linear in e vanishes in this limit. The solution is

$$\sigma_2(x,x) = \tfrac{1}{12} e^2 F^2_{\mu\nu}(x).$$

In addition $\sigma_0(x,x) = \sigma_1(x,x) = 0$. The divergent contribution to the gauge field two-point function at one-loop generated by charged scalars can thus be obtained from

$$\begin{aligned}
\operatorname{tr}\ln(-\mathrm{D}^2 + M^2) - \operatorname{tr}\ln(-\nabla^2 + M^2) &= \int \frac{\mathrm{d}t}{t}\, \mathrm{e}^{-M^2 t} \operatorname{tr}\left[\mathrm{e}^{t\nabla^2} - U(t)\right] \\
&\sim \frac{e^2}{12} \frac{1}{(4\pi)^{d/2}} \int \frac{\mathrm{d}t}{t^{1+d/2}}\, \mathrm{e}^{-M^2 t}\, t^2 \int \mathrm{d}^d x\, F^2_{\mu\nu}(x).
\end{aligned}$$

In dimensional regularization

$$\int_0 \mathrm{d}t\, t^{1-d/2} \sim \frac{2}{\varepsilon},$$

and, therefore, the divergent term is

$$\operatorname{tr}\ln(-\mathrm{D}^2 + M^2) - \operatorname{tr}\ln(-\nabla^2 + M^2) \sim \frac{1}{8\pi^2} \frac{e^2}{3\varepsilon} \frac{1}{4} \int \mathrm{d}^d x\, F^2_{\mu\nu}(x),$$

in agreement with the Feynman diagram calculation.

In the case of fermions one can relate the determinant to the boson determinant

$$\begin{aligned}
\operatorname{tr}\ln(\not{D} + M) &= \frac{1}{2} \operatorname{tr}\ln\left[M^2 - (\not{D})^2\right] = \tfrac{1}{2} \operatorname{tr}\ln\left[M^2 - \mathrm{D}^2 + \frac{1}{2} e F_{\mu\nu}\sigma_{\mu\nu}\right], \\
&= \frac{1}{2} \operatorname{tr}\mathbf{1}\, \operatorname{tr}' \ln\left(M^2 - \mathrm{D}^2\right) - \frac{1}{8} \operatorname{tr}\mathbf{1}\, e^2 \operatorname{tr}' F_{\mu\nu}\partial^{-2} F_{\mu\nu}\partial^{-2} + O(A^4), \\
&\sim \frac{1}{2} \operatorname{tr}\mathbf{1}\, e^2 \frac{1}{8\pi^2} \left(\frac{1}{3\varepsilon} - \frac{1}{\varepsilon}\right) \frac{1}{4} \int \mathrm{d}^d x\, F^2_{\mu\nu}(x),
\end{aligned}$$

where tr' means trace over space variables only. Taking into account the minus sign in front of the fermion determinant, we find $\operatorname{tr}\mathbf{1}$ times the boson result, in agreement with Feynman diagram calculations.

19 NON-ABELIAN GAUGE THEORIES: INTRODUCTION

In Chapter 18 we have described the structure and the formal properties of abelian gauge theories which provide a framework for the discussion of Quantum Electrodynamics. However, to be able to describe other fundamental interactions, Weak and Strong Interactions, it is necessary to generalize the concept of gauge theories to non-abelian groups. In this chapter we, therefore, construct a field theory invariant under *local*, that is, space-dependent, transformations of a general compact Lie group G. Inspired by the abelian example of Chapter 18, we immediately introduce the geometric concept of parallel transport, a concept discussed more extensively in Chapter 22 in the example of Riemannian manifolds. All the required mathematical quantities then appear quite naturally. We quantize gauge theories and study some of the formal properties of the quantum theory like the BRS symmetry. We show how perturbation theory can be regularized, a somewhat non-trivial problem. We finally discuss general aspects of the Higgs mechanism.

In the appendix we quantize massive non-abelian gauge fields and briefly explain the non-renormalizability of the field theory.

19.1 Geometric Construction

We consider a scalar field $\phi(x)$ transforming under a linear unitary or orthogonal representation $\mathcal{R}(G)$ of a compact group G. We want to construct a field theory which has a *local* G-symmetry, that is, a theory where the action is invariant under space-dependent group transformations, also called *gauge transformations*. Denoting by $\mathbf{g}$ a matrix belonging to the representation $\mathcal{R}(G)$, we write the ϕ-field transformation:

$$\phi'(x) = \mathbf{g}(x)\phi(x). \tag{19.1}$$

If we consider only products of fields taken at the same point, global invariance ($\mathbf{g}$ constant) implies local 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. The analogue of gauge transformations are there reparametrizations of Riemannian manifolds and we refer to Chapter 22 for a more detailed geometric and algebraic discussion. To solve the problem it is necessary to introduce parallel transporters $\mathbf{U}(C)$ which are curve-dependent elements of the representation $\mathcal{R}(G)$. If C is a curve joining point y to x, and $\mathbf{g}(x)$ a space-dependent group element, we write the transformation of $\mathbf{U}(C)$:

$$\mathbf{U}'(C) = \mathbf{g}(x)\mathbf{U}(C)\mathbf{g}^{-1}(y). \tag{19.2}$$

It then follows that the vector ϕ_U defined by

$$\phi_U = \mathbf{U}(C)\phi(y), \tag{19.3}$$

transforms by $\mathbf{g}(x)$ instead of $\mathbf{g}(y)$ and the quantity $\phi^\dagger(x)\mathbf{U}(C)\phi(y)$ is gauge invariant.

In the limit of an infinitesimal differentiable curve,

$$y_\mu = x_\mu + \mathrm{d}x_\mu\,, \tag{19.4}$$

we can parametrize $\mathbf{U}(C)$ in terms of the *connection* $\mathbf{A}_\mu(x)$, which is a vector from the point of view of space transformations, and a matrix (antisymmetric or anti-hermitian) belonging to the representation of the Lie algebra of G:

$$\mathbf{U}(C) = \mathbf{1} + \mathbf{A}_\mu(x)\mathrm{d}x_\mu + o\left(\|\mathrm{d}x_\mu\|\right). \tag{19.5}$$

The transformation properties of $\mathbf{A}_\mu(x)$ are obtained by expanding equation (19.2) at first order in $\mathrm{d}x_\mu$,

$$\mathbf{A}'_\mu(x) = \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x). \tag{19.6}$$

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 being affine.

Covariant derivative. To the connection $\mathbf{A}_\mu(x)$ is associated a covariant derivative $\mathbf{D}_\mu$, whose explicit form depends on the tensor on which it is acting. To obtain its expression when acting on $\boldsymbol{\phi}(x)$ we consider in equation (19.3) the limit (19.4) of an infinitesimal curve. The equation (19.3) becomes

$$\begin{aligned}\boldsymbol{\phi}_U &= \left(\mathbf{1} + \mathbf{A}_\mu(x)\mathrm{d}x_\mu\right)\left(\boldsymbol{\phi}(x) + \partial_\mu\boldsymbol{\phi}(x)\mathrm{d}x_\mu\right) + o\left(\|\mathrm{d}x_\mu\|\right) \\ &= \left(\mathbf{1} + \mathrm{d}x_\mu\mathbf{D}_\mu\right)\boldsymbol{\phi}(x) + o\left(\|\mathrm{d}x_\mu\|\right)\end{aligned} \tag{19.7}$$

with

$$\mathbf{D}_\mu = \mathbf{1}\,\partial_\mu + \mathbf{A}_\mu\,. \tag{19.8}$$

$\mathbf{D}_\mu$ is both a differential operator acting on space variables and a matrix. The identity

$$\mathbf{g}(x)\left(\mathbf{1}\,\partial_\mu + \mathbf{A}_\mu\right)\mathbf{g}^{-1}(x) = \mathbf{1}\,\partial_\mu + \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu\mathbf{g}^{-1}(x), \tag{19.9}$$

shows that $\mathbf{D}_\mu$ is a tensor, since $\mathbf{D}'_\mu$, the transform of $\mathbf{D}_\mu$ under the gauge transformation (19.6), is

$$\mathbf{D}'_\mu = \mathbf{g}(x)\mathbf{D}_\mu\,\mathbf{g}^{-1}(x). \tag{19.10}$$

In the equations (19.9,19.10) the products have to be understood as products of differential and multiplicative operators.

Infinitesimal gauge transformations. Setting,

$$\mathbf{g}(x) = \mathbf{1} + \boldsymbol{\omega}(x) + o\left(\|\boldsymbol{\omega}\|\right), \tag{19.11}$$

in which $\boldsymbol{\omega}(x)$ belongs to the Lie algebra of $\mathcal{R}(G)$, we derive from equation (19.6) the form of the infinitesimal gauge transformation of the field $\mathbf{A}_\mu$,

$$-\delta\mathbf{A}_\mu(x) = \partial_\mu\boldsymbol{\omega} + \left[\mathbf{A}_\mu, \boldsymbol{\omega}\right] \equiv \mathbf{D}_\mu\boldsymbol{\omega}\,. \tag{19.12}$$

In equation (19.8) we have given the form of the covariant derivative corresponding to the representation $\mathcal{R}(G)$. The equation (19.12) yields the form of the covariant derivative in the adjoint representation. One verifies:

$$\partial_\mu \omega' + \left[\mathbf{A}'_\mu, \omega'\right] = \mathbf{g}(x)\left\{\partial_\mu \omega + [\mathbf{A}_\mu, \omega]\right\}\mathbf{g}^{-1}(x), \tag{19.13}$$

in which $\mathbf{A}'_\mu$ is given by equation (19.6) and ω' by

$$\omega'(x) = \mathbf{g}(x)\omega(x)\mathbf{g}^{-1}(x). \tag{19.14}$$

Curvature tensor. The commutator of two covariant derivatives is no longer a differential operator:

$$\mathbf{F}_{\mu\nu}(x) = [\mathbf{D}_\mu, \mathbf{D}_\nu] = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu]. \tag{19.15}$$

It is again an element of the Lie algebra of $\mathcal{R}(G)$ and transforms, as a consequence of equation (19.10), as

$$\mathbf{F}'_{\mu\nu}(x) = \mathbf{g}(x)\mathbf{F}_{\mu\nu}(x)\mathbf{g}^{-1}(x). \tag{19.16}$$

$\mathbf{F}_{\mu\nu}$ is a tensor, the curvature tensor, generalization of the electromagnetic field of QED. As we show in Chapter 22 in a different context, the curvature tensor is associated with parallel transport along an infinitesimal closed curve.

Expressions in component form. In many cases it is useful to write previous expressions in component form. We expand $\mathbf{A}_\mu(x)$ on a basis $\{t^a\}$ of generators of the Lie algebra in the representation $\mathcal{R}$,

$$\mathbf{A}_\mu(x) = A^a_\mu(x)t^a. \tag{19.17}$$

The covariant derivative (19.8) then reads,

$$(\mathbf{D}_\mu)_{ij} = \partial_\mu \delta_{ij} + A^a_\mu(x)t^a_{ij}. \tag{19.18}$$

The equation (19.12) involves the structure constants f_{abc} of the Lie algebra,

$$-\delta A^a_\mu(x) = \partial_\mu \omega_a(x) + f_{bca} A^b_\mu(x)\omega_c(x). \tag{19.19}$$

This equation yields also the form of the covariant derivative in the adjoint representation.

Finally, the curvature tensor can also be expanded on the basis:

$$\mathbf{F}_{\mu\nu}(x) = F^a_{\mu\nu}(x)t^a, \tag{19.20}$$

and, therefore,

$$F^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + f_{bca}A^b_\mu(x)A^c_\nu(x). \tag{19.21}$$

This last expression is independent of the group representation.

Forms and gauge fields. It is sometimes convenient to use the language of differential forms and to associate with gauge fields one-forms, and two-forms with the curvature tensor:

$$\mathbf{A} = \mathbf{A}_\mu \mathrm{d}x_\mu\,, \quad \mathbf{F} = \mathbf{F}_{\mu\nu}\mathrm{d}x_\mu \wedge \mathrm{d}x_\nu = 2\left(\mathrm{d}\mathbf{A} + \mathbf{A}^2\right), \tag{19.22}$$

where the differential operator d acting on forms has been defined in Section 1.4: $\mathrm{d} \equiv \mathrm{d}x_\mu \partial_\mu$.

19.2 Gauge Invariant Actions

Matter fields. For boson fields transforming by (19.1), and taking into account the transformation (19.10) of the covariant derivatives, one verifies that the action

$$\mathcal{S}_{\mathrm{B}}(\boldsymbol{\phi}) = \int \mathrm{d}^d x \left[\left(\mathbf{D}_\mu \boldsymbol{\phi}(x)\right)^\dagger \mathbf{D}_\mu \boldsymbol{\phi}(x) + V\left(\boldsymbol{\phi}(x)\right) \right],$$

is gauge invariant if $V(\boldsymbol{\phi})$ is a group invariant function of the scalar field $\boldsymbol{\phi}$.

Similarly for fermions transforming by $\mathcal{R}(G)$ the action

$$\mathcal{S}_{\mathrm{F}}(\bar{\psi}, \psi) = -\int \mathrm{d}^d x \, \bar{\psi}(x) \left(\not{\mathbf{D}} + M\right) \psi(x),$$

is gauge invariant.

Gauge field. The simplest gauge invariant action $\mathcal{S}(\mathbf{A}_\mu)$ function of the gauge field $\mathbf{A}_\mu$, and generalization of the abelian action (18.35), has the form

$$\mathcal{S}(\mathbf{A}_\mu) = -\frac{1}{4e^2} \int \mathrm{d}^d x \ \operatorname{tr} \mathbf{F}^2_{\mu\nu}(x). \tag{19.23}$$

Note that we have not added a mass term for the gauge field as in Section 18.1. We indeed show in Appendix A19 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 or antisymmetric.

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 (19.6) 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 (19.23) is no longer a free field action; the gauge field has self-interactions and even the spectrum of the pure gauge action is non-perturbative (some analytic results can be obtained in dimension two). We indicate in Chapter 34 how lattice gauge theory provides a framework for non-perturbative investigations. The difference between the abelian and the non-abelian case is reminiscent of the non-linear chiral model.

(ii) As 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 which correspond to gauge transformations and, therefore, some *gauge fixing* is required.

19.3 Hamiltonian Formalism. Quantization

In this section we deal only with a pure gauge theory, because the extension of all arguments to a general gauge invariant theory is straightforward.

We first show that non-abelian gauge theories can be quantized, using a simple hamiltonian formalism, by the method explained in the abelian case in Section 18.4. This leads to a field theory that, at least at the formal level, is unitary because it corresponds to a hermitian hamiltonian.

19.3.1 Temporal gauge

Classical field equations. We first consider real time field theory, we denote 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 $(\mathbf{A}_i, x_i)$.

To the continuation to real time of the action (19.23) corresponds a classical field equation:

$$\mathbf{D}_\mu \mathbf{F}^{\mu\nu}(x) = 0\,, \tag{19.24}$$

in which the explicit form of $\mathbf{D}_\mu$ is given by equation (19.12). The equation (19.24) 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 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 $\left[\delta_{ij} - \mathrm{D}_i(\mathrm{D}_\perp^2)^{-1}\mathrm{D}_j\right]\dot{\mathbf{A}}_j$ appears ($\mathrm{D}_\perp^2$ 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 only discuss the quantization in the temporal (or Weyl) gauge.

Temporal gauge. As in the abelian case we first note that if $\mathbf{A}_\mu(t,x)$ is a solution of equation (19.24), any gauge transform of $\mathbf{A}_\mu(t,x)$ is also a solution. To describe all solutions we can thus make a gauge section, that is, consider a section in the space of all gauge fields which intersects once all gauge orbits. We then represent all solutions by an element of the section and a gauge transformation. One choice of gauge condition is specially well-suited to the construction of a hamiltonian formalism,

$$\mathbf{A}_0(t,x) = 0\,, \tag{19.25}$$

which defines the *temporal gauge*.

The euclidean equation (19.24), separating the space and time components, is equivalent to

$$\begin{cases} \mathbf{D}_0\mathbf{F}_{0k} - \mathbf{D}_l\mathbf{F}_{lk} = 0\,, \\ \mathbf{D}_l\mathbf{F}_{l0} = 0\,. \end{cases} \tag{19.26}$$

The indices k, l vary only from 1 to $d-1$ and correspond to space components. In the gauge (19.25) the equations simplify and become

$$\dot{\mathbf{E}}_k = \mathbf{D}_l\mathbf{F}_{lk}\,, \tag{19.27}$$

$$\mathbf{D}_l\mathbf{E}_l = 0 \tag{19.28}$$

with

$$\mathbf{E}_k = \dot{\mathbf{A}}_k\,. \tag{19.29}$$

The equation (19.27) is a dynamical equation that can be directly derived from the initial lagrangian in which the condition (19.25) has been used:

$$\mathcal{L}(\mathbf{A}_k) = -\operatorname{tr}\int \mathrm{d}^{d-1}x\left[\dot{\mathbf{A}}_k^2(t,x) - \frac{1}{4e^2}\mathbf{F}_{kl}^2(t,x)\right]. \tag{19.30}$$

The expression (19.30) defines a conventional lagrangian for the space components of the gauge field: $\mathbf{E}_k$ is the conjugated momentum of $\mathbf{A}_k$ and the corresponding hamiltonian takes the form:

$$\mathcal{H}(\mathbf{E},\mathbf{A}) = -\operatorname{tr}\int \mathrm{d}^{d-1}x \left[\mathbf{E}_k^2(x) + \frac{1}{4e^2}\mathbf{F}_{kl}^2(x)\right]. \tag{19.31}$$

The equation (19.28) instead is a constraint equation, non-abelian generalization of Gauss's law. The only relevant solutions of the field equations are those which satisfy the constraint. One verifies that the equation (19.28) is compatible with the classical motion, that is, that the Poisson brackets of the equation with the hamiltonian vanish. The reason is easy to understand: the gauge condition (19.25) is left invariant by time-independent gauge transformations. Therefore, time-independent gauge transformations form a symmetry group of the lagrangian (19.30) and thus of the hamiltonian (19.31). The quantities $\mathbf{D}_l\mathbf{E}_l$ are the generators, in the sense of Poisson brackets, of the symmetry group.

These considerations immediately generalize to the quantum theory, the quantum operators $\mathbf{D}_l\mathbf{E}_l$ generators of a symmetry group, commute with the hamiltonian. The space of admissible physical states $\Psi(\mathbf{A})$ is restricted by the quantum generalization of Gauss's law:

$$\mathbf{D}_l\mathbf{E}_l\Psi(\mathbf{A}) \equiv \mathbf{D}_l\frac{1}{i}\frac{\delta}{\delta\mathbf{A}_l(x)}\Psi(\mathbf{A}) = 0\,. \tag{19.32}$$

The equation implies that physical states are gauge invariant, that is, belong to the invariant sector of the symmetry group, a subspace which is left invariant by quantum evolution.

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

$$\mathcal{Z} = \int [\mathrm{d}\mathbf{A}_\mu]\,\delta(\mathbf{A}_d)\exp\left[\frac{1}{4e^2}\int \mathrm{d}^d x\ \operatorname{tr}\mathbf{F}_{\mu\nu}^2(x)\right]. \tag{19.33}$$

Note that at zero temperature the perturbative vacuum is automatically gauge invariant and Gauss's law plays no role. This is no longer the case at finite temperature.

Remarks. 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 18.4). In particular, in the temporal gauge the theory is not renormalizable in the sense of power counting. Indeed the propagator in this gauge

$$W_{ij}^{(2)}(\mathbf{k}_\perp, k_d) = \frac{1}{k^2}\left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}_\perp^2}\right) + \frac{1}{k_d^2}\frac{k_i k_j}{\mathbf{k}_\perp^2}, \tag{19.34}$$

in which $\mathbf{k}_\perp$ is the "space" part of $\mathbf{k}$, does not decrease at k_d fixed for large spatial momenta $|\mathbf{k}_\perp|$.

These problems are solved by showing that gauge invariant observables can equivalently be calculated from another quantum action which leads to a theory which is explicitly covariant and renormalizable by power counting.

19.3.2 Covariant gauge

We would like to work with a covariant gauge constraining $\partial_\mu \mathbf{A}_\mu$, rather than with the temporal gauge (19.33). For this purpose we now use the analysis of Sections 16.2–16.6. We start from the equation for the space-dependent group element $\mathbf{g}(x)$

$$E(\mathbf{g}) \equiv \partial_\mu \mathbf{A}^g_\mu(x) - \boldsymbol{\nu}(x) = 0\,, \tag{19.35}$$

where $\mathbf{A}^g_\mu$ is the gauge transform (equation (19.6)) by $\mathbf{g}$ of $\mathbf{A}_\mu$ and $\boldsymbol{\nu}(x)$ an arbitrary field which belongs to the Lie algebra of $\mathcal{R}(G)$. If $\boldsymbol{\nu}(x)$ is a stochastic field the equation (19.35) gives to $\mathbf{g}(x)$ a stochastic dynamics.

It is convenient to set

$$\mathbf{B}_\mu = \mathbf{A}^g_\mu\,. \tag{19.36}$$

We assume that the equation (19.35) has a unique solution, which is equivalent to assert that in the space of gauge fields the surface $\partial_\mu \mathbf{B}_\mu(x) - \boldsymbol{\nu}(x) = 0$ intersects once and only once all gauge orbits. This is certainly possible for small fields and therefore in perturbation theory (see the remark at the end of the section concerning this problem beyond perturbation theory).

We need 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,

$$\delta E(\mathbf{g}) = -\partial_\mu \mathbf{D}_\mu(\mathbf{B})\boldsymbol{\omega}(x). \tag{19.37}$$

We now introduce spinless fermions $\bar{\mathbf{C}}$ and $\mathbf{C}$, the Faddeev–Popov "ghosts", and a boson field $\boldsymbol{\lambda}$ all transforming under the adjoint representation. This allows to write the identity (a version of equation (16.18))

$$1 = \int [\mathrm{d}\mathbf{g}\,\mathrm{d}\bar{\mathbf{C}}\,\mathrm{d}\mathbf{C}\,\mathrm{d}\boldsymbol{\lambda}] \exp\left[-\mathcal{S}_{\mathrm{gauge}}(\mathbf{A}^g_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})\right], \tag{19.38}$$

with

$$\mathcal{S}_{\mathrm{gauge}}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \int \mathrm{d}^d x \ \mathrm{tr}\left\{\boldsymbol{\lambda}(x)\left[\partial_\mu \mathbf{A}_\mu(x) - \boldsymbol{\nu}(x)\right] + \mathbf{C}(x)\partial_\mu \mathbf{D}_\mu(\mathbf{A})\bar{\mathbf{C}}(x)\right\}. \tag{19.39}$$

We introduce the identity (19.38) in the representation (19.33) of the partition function and obtain

$$\mathcal{Z} = \int [\mathrm{d}\mathbf{g}\,\mathrm{d}\bar{\mathbf{C}}\,\mathrm{d}\mathbf{C}\,\mathrm{d}\boldsymbol{\lambda}\,\mathrm{d}\mathbf{A}_\mu]\delta(\mathbf{A}_d) \exp\left[\frac{1}{4e^2}\int \mathrm{d}^d x \ \mathrm{tr}\,\mathbf{F}^2_{\mu\nu}(x) - \mathcal{S}_{\mathrm{gauge}}(\mathbf{A}^g_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})\right]. \tag{19.40}$$

We then change variables $\mathbf{A}^g_\mu \mapsto \mathbf{A}_\mu$. The action (19.23) is gauge invariant. Only the gauge condition $\delta(\mathbf{A}_d)$ is affected. Changing g into g^{-1} we find

$$\mathcal{Z} = \int [\mathrm{d}\mathbf{g}\,\mathrm{d}\bar{\mathbf{C}}\,\mathrm{d}\mathbf{C}\,\mathrm{d}\boldsymbol{\lambda}\,\mathrm{d}\mathbf{A}_\mu]\delta(\mathbf{A}^g_d) \exp\left[\frac{1}{4e^2}\int \mathrm{d}^d x \ \mathrm{tr}\,\mathbf{F}^2_{\mu\nu}(x) - \mathcal{S}_{\mathrm{gauge}}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})\right]. \tag{19.41}$$

We now integrate over the group field $\mathbf{g}(x)$. The result of the integral $\int [\mathrm{d}\mathbf{g}]\,\delta(\mathbf{A}^g_d)$ is gauge invariant. We can thus calculate it only for fields satisfying the condition $\mathbf{A}_d = 0$. We have to evaluate the determinant coming from the variation of $\mathbf{A}^g_d$ at $\mathbf{g} = \mathbf{1}$. We find,

$$\delta \mathbf{A}_d = \mathbf{D}_d \delta\mathbf{g}\,.$$

Since $\mathbf{A}_d = 0$ the covariant derivative $\mathbf{D}_d$ is equal to the ordinary derivative ∂_d and the corresponding determinant is a constant independent of $\mathbf{A}_d$.

This leads to the functional integral representation

$$\mathcal{Z} = \int \left[\mathrm{d}\mathbf{A}_\mu \, \mathrm{d}\bar{\mathbf{C}} \, \mathrm{d}\mathbf{C} \, \mathrm{d}\boldsymbol{\lambda}\right] \exp\left[-\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})\right], \tag{19.42}$$

where the quantum action $\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ is given by

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \int \mathrm{d}^d x \; \mathrm{tr}\left[-\frac{1}{4e^2}\mathbf{F}_{\mu\nu}^2 + \boldsymbol{\lambda}(x)\left[\partial_\mu \mathbf{A}_\mu(x) - \boldsymbol{\nu}(x)\right] + \mathbf{C}(x)\partial_\mu \mathbf{D}_\mu \bar{\mathbf{C}}(x)\right]. \tag{19.43}$$

Since the partition function is independent of $\boldsymbol{\nu}$ one can average over the "noise" field $\boldsymbol{\nu}(x)$ with a gaussian distribution

$$[\mathrm{d}\rho(\boldsymbol{\nu})] = [\mathrm{d}\boldsymbol{\nu}] \exp\left[\frac{1}{2\xi e^2}\int \mathrm{d}^d x \; \mathrm{tr}\, \boldsymbol{\nu}^2(x)\right]. \tag{19.44}$$

The averaged partition function $\mathcal{Z}$ then reads

$$\mathcal{Z} = \int \left[\mathrm{d}\mathbf{A}_\mu \, \mathrm{d}\bar{\mathbf{C}} \, \mathrm{d}\mathbf{C} \, \mathrm{d}\boldsymbol{\lambda}\right] \exp\left[-\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda})\right], \tag{19.45}$$

where $\mathcal{S}$ is a local action:

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}) = \int \mathrm{d}^d x \; \mathrm{tr}\left[-\frac{1}{4e^2}\mathbf{F}_{\mu\nu}^2 + \frac{\xi e^2}{2}\boldsymbol{\lambda}^2(x) + \boldsymbol{\lambda}(x)\partial_\mu \mathbf{A}_\mu(x) + \mathbf{C}(x)\partial_\mu \mathbf{D}_\mu \bar{\mathbf{C}}(x)\right]. \tag{19.46}$$

Except in the limit in which ξ vanishes, it is also possible to integrate over $\boldsymbol{\lambda}(x)$ to find a new quantum action:

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}) = \int \mathrm{d}^d x \; \mathrm{tr}\left\{-\frac{1}{e^2}\left[\frac{1}{4}\mathbf{F}_{\mu\nu}^2 + \frac{1}{2\xi}\left(\partial_\mu \mathbf{A}_\mu\right)^2\right] + \mathbf{C}(x)\partial_\mu \mathbf{D}_\mu \bar{\mathbf{C}}(x)\right\}. \tag{19.47}$$

This form of the quantum action is better suited for perturbative calculations, although geometric properties of the action are more apparent in expression (19.46), in particular the BRS symmetry.

We have, therefore, established the formal equivalence between the two expressions (19.47) and (19.33) 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, therefore, unphysical.

Remark. As pointed out by Gribov, in contrast with the abelian case, depending on the value of the gauge field $\mathbf{A}_\mu(x)$, the gauge condition (19.35) has not always a unique solution in $\mathbf{g}(x)$, a problem called Gribov's ambiguity. When two solutions merge, the operator $\partial_\mu \mathbf{D}_\mu(\mathbf{A})$ has zero eigenvalues. This implies that the representation (19.45) is not meaningful beyond perturbation theory. The same ambiguity has been shown to arise for a large class of gauge conditions.

19.3.3 BRS symmetry

One consequence of the quantization procedure is that the action (19.47) is no longer gauge invariant. On the other hand we know from the analysis of Chapter 16 applied to the quantization procedure that the action now has a BRS symmetry, consequence of the stochastic dynamics given to the degrees of freedom of the gauge group variables. To understand the form of the BRS transformations, it is convenient to separate the gauge group degrees of freedom which induce the BRS symmetry from the other degrees of freedom of the gauge field which play no role

$$\mathbf{A}_\mu(x) = \mathbf{B}_\mu^g(x),$$

where $\mathbf{B}_\mu(x)$ satisfies the gauge condition (19.35): $\partial_\mu \mathbf{B}_\mu(x) = \boldsymbol{\nu}(x)$. From the analysis of Section 16.4 we know the form of the BRS transformations in the case of group manifolds:

$$\begin{cases} \delta\mathbf{g}(x) = \varepsilon\bar{\mathbf{C}}(x)\mathbf{g}(x), & \delta\bar{\mathbf{C}}(x) = \varepsilon\bar{\mathbf{C}}^2(x), \\ \delta\mathbf{C}(x) = \varepsilon\boldsymbol{\lambda}(x), & \delta\boldsymbol{\lambda}(x) = 0\,. \end{cases} \tag{19.48}$$

The field $\mathbf{B}_\mu(x)$ has a dynamics provided by the gauge action and is not affected by BRS transformations:

$$\delta\mathbf{B}_\mu(x) = 0\,. \tag{19.49}$$

Calculating the effect of a BRS transformation on the field $\mathbf{A}_\mu$ we then find:

$$\delta\mathbf{A}_\mu(x) = \delta\mathbf{B}_\mu^g(x) = -\varepsilon\mathbf{D}_\mu\bar{\mathbf{C}}(x)\,. \tag{19.50}$$

This result is not surprising: the transformations (19.48) correspond to an infinitesimal change in the gauge group degrees of freedom and thus to an infinitesimal gauge transformation for $\mathbf{A}_\mu$. We recognize also in equation (19.50) the transformation of the current associated with group elements as discussed in Section 17.3 (equations (17.37–17.39)).

Introducing the BRS differential operator

$$\mathcal{D} = \int \mathrm{d}^d x \left[-\mathbf{D}_\mu\bar{\mathbf{C}}(x)\frac{\delta}{\delta\mathbf{A}_\mu(x)} + \bar{\mathbf{C}}^2(x)\frac{\delta}{\delta\bar{\mathbf{C}}(x)} + \lambda(x)\frac{\delta}{\delta\mathbf{C}(x)}\right], \tag{19.51}$$

we can also express the BRS symmetry of the quantized action by the equation

$$\mathcal{D}\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{C}}, \mathbf{C}, \boldsymbol{\lambda}) = 0\,. \tag{19.52}$$

Moreover, we have shown quite generally in Section 16.5.1 that $\mathcal{S}_{\mathbf{gauge}}$ is BRS exact. Here,

$$\mathcal{S}_{\mathbf{gauge}} = \mathcal{D}\int \mathrm{d}^d x\, \mathbf{C}(x)\left[\partial_\mu\mathbf{A}_\mu(x) - \boldsymbol{\nu}(x)\right],$$

or after integration over the $\boldsymbol{\nu}$ field

$$\mathcal{S}_{\mathbf{gauge}} = \mathcal{D}\int \mathrm{d}^d x\, \mathbf{C}(x)\left[\partial_\mu\mathbf{A}_\mu(x) + \xi e^2\boldsymbol{\lambda}(x)\right]. \tag{19.53}$$

We show in Chapter 21, in a more general framework, that WT identities associated with the BRS symmetry (19.48,19.50) imply the structural stability of the quantum action (19.46) under renormalization.

19.4 Perturbation Theory, Regularization

Compared with the abelian case, the new features of the action (19.47) are the presence of gauge field self-interactions and ghost terms. Let us first write the different terms of the gauge action in component form to establish conventions and normalizations. The gauge action takes the form:

$$\mathcal{S}(A^a_\mu) = \frac{1}{4e^2} \int \mathrm{d}^d x \, F^a_{\mu\nu} F^a_{\mu\nu}, \tag{19.54}$$

where the curvature tensor is given by equation (19.21) and the trace of the unit matrix has been swallowed into a redefinition of the coupling constant. In the covariant gauge of Section 19.3.2, the gauge field propagator is (equation (18.22))

$$[\Delta_\xi]^{ab}_{\mu\nu}(k) = e^2 \delta ab \left[\frac{\delta_{\mu\nu}}{k^2} + (\xi - 1)\frac{k_\mu k_\nu}{(k^2)^2}\right]. \tag{19.55}$$

In four dimensions, as in the abelian case, the gauge field has dimension 1. The ghost field action becomes

$$\mathcal{S}_{\text{ghost}} = \int \mathrm{d}^d x \, C^a \partial_\mu \left(\partial_\mu \delta_{ac} + f_{bca} A^b_\mu\right) \bar{C}^c. \tag{19.56}$$

The ghost fields thus have a simple δ_{ab}/p^2 propagator and 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 of course the same as in the abelian case.

Compared to the abelian case the non-abelian theory has three new vertices coming from the gauge field self-interactions and the interaction with the ghost fields. The gauge field three-point function at leading order is

$$[\Gamma^{(3)}]^{abc}_{\mu\nu\rho}(p,q,r) = \frac{i}{e^2} f_{abc} \left[(r-q)_\mu \delta_{\nu\rho} + (p-r)_\nu \delta_{\rho\mu} + (q-p)_\rho \delta_{\mu\nu}\right]. \tag{19.57}$$

The gauge field four-point function is given by

$$\begin{aligned}[\Gamma^{(4)}]^{abcd}_{\mu\nu\rho\sigma} = \frac{1}{e^2} \big[& f_{eab} f_{ecd} \left(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}\right) + f_{eac} f_{ebd} \left(\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}\right) \\ & + f_{ead} f_{ecb} \left(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\nu}\delta_{\sigma\rho}\right)\big]. \end{aligned} \tag{19.58}$$

All terms are obtained from the first by exchanging the indices to make the correlation function totally symmetric. Finally, the ghost gauge field vertex is

$$\left\langle C^a(p) \bar{C}^b(q) A^c_\mu(r)\right\rangle = -i f_{abc} p_\mu \,. \tag{19.59}$$

Notice 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. There is, however, a special case $\xi = 0$ called Landau's gauge, in which, because $\partial_\mu \mathbf{A}_\mu$ 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 some geometric factors corresponding to group indices. For example, the coupling to fermions generated by the covariant derivative (19.18) is simply $\gamma_\mu t^a_{ij}$.

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 which is obviously 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 which is both unitary and renormalizable (for more details see Appendix A19). On the other hand, we want eventually to prove the gauge independence of the theory and therefore we must be able to define it 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 which will be given in Chapter 37.

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. We review the three regularization methods which we have always considered in this work. Dimensional regularization is the most convenient for practical calculations and works in the absence of chiral fermions. Lattice regularization, which is also relevant for non-perturbative calculations can be used generally (see Chapter 34 for details) since recently a method for handling chiral fermions has been discovered (related to Ginsparg–Wilson's relation). Finally, momentum or Pauli–Villars's type regularizations work partially in geometric models, in the sense that they regularize all diagrams except one-loop diagrams. We verify this property again here. The regularized gauge action takes the form:

$$\mathcal{S}(\mathbf{A}_\mu) = \int \mathrm{d}^d x \, \operatorname{tr} \mathbf{F}_{\mu\nu} P(\mathbf{D}^2/\Lambda^2) \mathbf{F}_{\mu\nu} \,, \tag{19.60}$$

in which P is a polynomial of arbitrary degree. In the same way the gauge function $\partial_\mu \mathbf{A}_\mu$ is changed into

$$\partial_\mu \mathbf{A}_\mu \longmapsto Q(\partial^2/\Lambda^2) \partial_\mu \mathbf{A}_\mu \,, \tag{19.61}$$

in which Q is a polynomial of same degree as P. As a consequence both the gauge field propagator and the ghost propagator can be made arbitrarily convergent. However, as in the abelian case, the covariant derivatives generate new interactions which are more singular. It is easy to verify that the power counting of one-loop diagrams is unchanged while higher order diagrams can be made convergent by taking the degrees of P and Q large enough.

For matter fields the situation is the same as in the abelian case, for example, massive fermions contributions can be regularized by adding a set of regulator fields, massive fermions and bosons with spin.

Again in the case of chiral fermions, global chiral properties can be preserved, but problems arise with local chiral transformations. However, the problem of the compatibility between the gauge symmetry and the quantum corrections is reduced to an explicit verification of the WT identities for the one-loop diagrams. Note that the preservation of gauge symmetry is necessary for the cancellation of unphysical states in physical amplitudes, and thus essential to the physical consistency of the quantum field theory.

WT identities and renormalization. From the BRS symmetry corresponding to the transformations (19.48–19.50) follows WT identities. Their form is somewhat complicated and we postpone the discussion to Chapter 21, where we derive the form of the renormalized action for a general gauge theory. We give here the result only in the example of the pure gauge action in the covariant gauge. We can assume that the gauge group G is simple. Then the renormalized form of the action (19.47) is given by the substitution:

$$\begin{cases} e^2 \longmapsto Z_e e^2 \,, & \mathbf{A}_\mu \longmapsto Z_A^{1/2} \mathbf{A}_\mu \,, \\ \xi \longmapsto Z_A Z_e^{-1} \xi \,, & \mathbf{C}\bar{\mathbf{C}} \longmapsto Z_C \mathbf{C}\bar{\mathbf{C}} \,. \end{cases} \tag{19.62}$$

This result has a simple interpretation: the gauge structure (19.47) is preserved and the coefficient of $(\partial_\mu \mathbf{A}_\mu)^2$ is unrenormalized exactly as in the abelian case. However, unlike the abelian case, the gauge transformation of the gauge field and, more generally the form of the covariant derivative, are modified by the gauge field renormalization.

19.5 The Non-Abelian Higgs Mechanism

We have already discussed the Higgs mechanism in the abelian case. 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 without generating massless Goldstone modes. Simply, because the group structure is richer, a number of different situations may arise.

We consider a classical gauge invariant action for a gauge field coupled to a scalar boson $\boldsymbol{\phi}$ transforming under an orthogonal representation of the symmetry group:

$$\mathcal{S}(\mathbf{A}_\mu, \boldsymbol{\phi}) = \int \mathrm{d}^d x \left[-\frac{1}{4e^2} \operatorname{tr} \mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2} \mathbf{D}_\mu \boldsymbol{\phi}(x) \cdot \mathbf{D}_\mu \boldsymbol{\phi}(x) + V\big(\boldsymbol{\phi}(x)\big) \right] . \tag{19.63}$$

We assume that the symmetric potential $V(\boldsymbol{\phi})$ 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 $\boldsymbol{\phi}$ we consider here only two types of examples.

19.5.1 Simple Lie groups

We first assume that G, the symmetry group of the action, is simple and is thus also the gauge group. Moreover, we assume for simplicity that the field $\boldsymbol{\phi}$ belongs to an irreducible representation. We call H the stabilizer of the vector $\mathbf{v}$, the minimum of the potential. We separate the generators t^α of G in the matrix representation into two subsets $\alpha \le p$ 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 $\boldsymbol{\phi}(x)$ by

$$\boldsymbol{\phi}(x) = \exp\left[\sum_{\alpha>p} t^\alpha \theta^\alpha(x)\right] \big(\mathbf{v} + \boldsymbol{\rho}(x)\big), \tag{19.64}$$

in which the vectors $\boldsymbol{\rho}$ and $\{t_i^\alpha v_i\}$ span two orthogonal subspaces. The transformation:

$$\boldsymbol{\phi}(x) \longmapsto \{\theta^\alpha(x), \quad \boldsymbol{\rho}(x)\}, \tag{19.65}$$

is such that the new fields $\boldsymbol{\rho}(x)$ and $\theta^\alpha(x)$ can be expanded in powers of $\boldsymbol{\phi}(x) - \mathbf{v}$. In the absence of gauge fields, we have used the representation (19.64) to show that the fields $\theta^\alpha(x)$ correspond to massless Goldstone modes induced by the spontaneous breaking of the G-symmetry.

Here the equation (19.64) can also be viewed as a local group transformation relating the two fields $\boldsymbol{\phi}$ and $\boldsymbol{\rho} + \mathbf{v}$. If we perform on the field $\mathbf{A}_\mu$ a gauge transformation of the form (19.6) with

$$\mathbf{g}(x) = \exp\left[\sum_{\alpha>p} t^\alpha \theta^\alpha(x)\right], \tag{19.66}$$

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 $\rho = 0$ it reduces to a mass term for the gauge field:

$$\tfrac{1}{2}\mathbf{D}_\mu\boldsymbol{\phi}\cdot\mathbf{D}_\mu\boldsymbol{\phi}|_{\rho=0} = \tfrac{1}{2}\sigma_{\alpha\beta}A_\mu^\alpha A_\mu^\beta\,, \tag{19.67}$$

with the mass matrix

$$\sigma_{\alpha\beta} = t_{ij}^\alpha v_j t_{ik}^\beta v_k\,. \tag{19.68}$$

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 would-be Goldstone bosons. We conclude that the spontaneous breaking of the G-symmetry generates no Goldstone bosons but instead 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 (19.66), the set of massive vector fields can be quantized in a completely standard way. As in the abelian case, the quantized theory, however, is not renormalizable.

19.5.2 The $G \times G$ symmetry

Another possibility is that the symmetry group of the action is the direct product of the gauge group G by another group G'. We here consider only the simplest example where the symmetry group is $G \times G$ and G is simple. We assume that the scalar boson field $\boldsymbol{\phi}$ is a matrix transforming under $G \times G$ by

$$\boldsymbol{\phi}' = \mathbf{g}_1\boldsymbol{\phi}\mathbf{g}_2^{-1}, \tag{19.69}$$

in which $\mathbf{g}_1$ and $\mathbf{g}_2$ are two elements of G in a matrix representation. We further assume that one minimum of the potential is proportional to the unit matrix $\boldsymbol{\phi} = v\mathbf{1}$ 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/H is then a symmetric space (see Appendix A15.4). As above the would-be Goldstone bosons correspond to gauge transformations and can thus be eliminated from the action. In this example all components of the gauge field acquire a mass since the symmetry corresponding to the gauge field is completely broken. Furthermore, all components of the gauge field have the same mass m_A:

$$\operatorname{tr}{}^T(\mathbf{D}_\mu\boldsymbol{\phi})\,(\mathbf{D}_\mu\boldsymbol{\phi})\,|_{\boldsymbol{\phi}=v} = -v^2\operatorname{tr}\mathbf{A}_\mu^2 \;\Rightarrow\; m_A^2 = e^2v^2. \tag{19.70}$$

The action obtained after the gauge transformation specified by equation (19.66) contains only 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 field theory is, however, not renormalizable by power counting. The difference with the massive vector field theory we examine in Appendix A19 is that the suitable addition of some scalar fields makes this theory equivalent, at least for physical observables, with a renormalizable theory with additional unphysical degrees of freedom.

Remark. 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 Appendix A19).

The $SU(2)$ *example.* We discuss more specifically the important example of the $SU(2)$ group because it can be considered as a simplified version of the Standard Model of weak-electromagnetic interactions which will be described in Section 20.1. We take for scalar field $\boldsymbol{\phi}$ a 2×2 complex matrix transforming under the (1/2,1/2) representation of $SU(2)\times SU(2)$ (see also Section 13.6). The simplest action then reads:

$$\mathcal{S}(\mathbf{A}_\mu,\boldsymbol{\phi})=\int \mathrm{d}^d x\left[\frac{1}{4e^2}\mathbf{F}^2_{\mu\nu}+\operatorname{tr}\left(\mathbf{D}_\mu\boldsymbol{\phi}\right)^\dagger \mathbf{D}_\mu\boldsymbol{\phi}+r\operatorname{tr}\boldsymbol{\phi}\boldsymbol{\phi}^\dagger+\frac{\lambda}{6}\left(\operatorname{tr}\boldsymbol{\phi}\boldsymbol{\phi}^\dagger\right)^2\right]. \tag{19.71}$$

We have represented the gauge field as a three-component vector $\mathbf{A}_\mu$. The curvature tensor reads:

$$\mathbf{F}_{\mu\nu}=\partial_\mu\mathbf{A}_\nu-\partial_\nu\mathbf{A}_\mu-\mathbf{A}_\mu\times\mathbf{A}_\nu\,, \tag{19.72}$$

and the covariant derivative acts like

$$\mathbf{D}_\mu\boldsymbol{\phi}=\left(\partial_\mu+\tfrac{1}{2}i\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right)\boldsymbol{\phi}\,. \tag{19.73}$$

The $\boldsymbol{\tau}$ matrices are identical to the σ matrices defined in Section A8.1.4. We choose the expectation value of $\boldsymbol{\phi}$ proportional to the unit matrix. It is then convenient to parametrize $\boldsymbol{\phi}$ in the form (see Section 13.6):

$$\boldsymbol{\phi}=\tfrac{1}{2}\left(\sigma+i\boldsymbol{\tau}\cdot\boldsymbol{\pi}\right), \tag{19.74}$$

in which σ and $\boldsymbol{\pi}$ are real fields. An infinitesimal gauge transformation in this representation takes the form:

$$\begin{cases}\delta\mathbf{A}_\mu=\partial_\mu\boldsymbol{\omega}-\mathbf{A}_\mu\times\boldsymbol{\omega}\,,\\ \delta\sigma=\tfrac{1}{2}\boldsymbol{\omega}\cdot\boldsymbol{\pi}\,,\\ \delta\boldsymbol{\pi}=-\tfrac{1}{2}\sigma\boldsymbol{\omega}+\tfrac{1}{2}\boldsymbol{\omega}\times\boldsymbol{\pi}\,.\end{cases} \tag{19.75}$$

The scalar field action in these variables reads:

$$\mathcal{S}_{\text{scalar}}=\int \mathrm{d}^d x\left[\tfrac{1}{2}\left(\partial_\mu\sigma-\tfrac{1}{2}\boldsymbol{\pi}\cdot\mathbf{A}_\mu\right)^2+\tfrac{1}{2}\left(\partial_\mu\boldsymbol{\pi}+\tfrac{1}{2}\sigma\mathbf{A}_\mu-\tfrac{1}{2}\mathbf{A}_\mu\times\boldsymbol{\pi}\right)^2+\tilde{V}(\sigma^2+\boldsymbol{\pi}^2)\right], \tag{19.76}$$

with

$$\tilde{V}(s)=\tfrac{1}{2}rs+\tfrac{1}{24}\lambda s^2.$$

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 $\boldsymbol{\phi}$ has a non-zero expectation value. Without loss of generality we choose the component σ to have a non-zero expectation value:

$$\langle\sigma\rangle=v\,.$$

Then the symmetry $SU(2)\times SU(2)$ is broken down to the diagonal $SU(2)$ group. 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

$$\mathcal{S}(\mathbf{A}_\mu,\sigma)=\int \mathrm{d}^d x\left[\frac{1}{4e^2}\mathbf{F}^2_{\mu\nu}+\frac{1}{2}(\partial_\mu\sigma)^2+\frac{1}{8}\sigma^2\mathbf{A}^2_\mu+\frac{r}{2}\sigma^2+\frac{\lambda}{24}\sigma^4\right]. \tag{19.77}$$

This action has an $O(3)$-invariant. 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. The gauge field mass m_A is given in the classical approximation in terms of the σ-field expectation value v by

$$m_A=ev/2\,. \tag{19.78}$$

19.5.3 Gauge fixing of the Higgs model in a covariant gauge

If we consider the contribution (19.76) to the action, we see that when the σ-field has an expectation value, a term of the form $\partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$ is generated which introduces a mixing between the would-be Goldstone boson $\boldsymbol{\pi}$ and the longitudinal part of the vector field. This is a feature which we have already encountered in the abelian case (Section 18.10). As suggested by 't Hooft, it is possible to use the gauge function to eliminate such a term. In the $SU(2)$ example we can take as gauge function F:

$$F(\mathbf{A}_\mu, \boldsymbol{\pi}) = \partial_\mu \mathbf{A}_\mu + \tfrac{1}{2}\lambda\xi\boldsymbol{\pi} \,. \tag{19.79}$$

After a gaussian integration, the corresponding contribution to the action is

$$\mathcal{S}_F = \mathcal{S}_{\mathrm{gauge}} + \mathcal{S}_{\mathrm{ghost}} \tag{19.80}$$

with

$$\mathcal{S}_{\mathrm{gauge}} = \frac{1}{2\xi e^2} \int \mathrm{d}^d x \left(\partial_\mu \mathbf{A}_\mu + \tfrac{1}{2}\lambda\xi\boldsymbol{\pi}\right)^2 \tag{19.81}$$

and

$$\mathcal{S}_{\mathrm{ghost}} = \int \mathrm{d}^d x \left[\partial_\mu \mathbf{C} \cdot \left(\partial_\mu \bar{\mathbf{C}} - \mathbf{A}_\mu \times \bar{\mathbf{C}}\right) + \frac{\lambda\xi}{4}\mathbf{C}\left(\sigma\bar{\mathbf{C}} + \boldsymbol{\pi} \times \bar{\mathbf{C}}\right)\right]. \tag{19.82}$$

At leading order the term $\partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$ is eliminated by the choice

$$\lambda = e^2 v \,. \tag{19.83}$$

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 the propagators (equations (18.140)) have no poles at zero momentum and no IR problems are encountered:

$$\begin{aligned}
W^{(2)}_{\mu\nu} &= \frac{e^2\delta_{\mu\nu}}{k^2 + m_A^2} + \frac{e^2(\xi - 1)k_\mu k_\nu}{(k^2 + m_A^2)(k^2 + \xi m_A^2)}, \\
W^{2)}_{\boldsymbol{\pi}\boldsymbol{\pi}} &= \frac{1}{k^2 + \xi m_A^2}, \\
W^{(2)}_{\mathbf{C}\bar{\mathbf{C}}} &= \frac{1}{k^2 + \xi m_A^2},
\end{aligned} \tag{19.84}$$

in which m_A is the mass of $\mathbf{A}_\mu$ in the tree approximation (equation (19.78)). Furthermore, all unphysical states have a mass which 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 one must include all particles both physical and unphysical. By showing that the S-matrix does not depend on the gauge, one proves simultaneously that the contributions of unphysical states cancels in the intermediate states and thus the S-matrix is unitary in the physical subspace. A general proof of this kind will be given in Chapter 21.

19.5.3 Gauge fixing of the Higgs model in a covariant gauge

If we consider the contribution (19.76) to the action, we see that when the σ-field has an expectation value, a term of the form $\partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$ is generated which introduces a mixing between the would-be Goldstone boson $\boldsymbol{\pi}$ and the longitudinal part of the vector field. This is a feature which we have already encountered in the abelian case (Section 18.10). As suggested by 't Hooft, it is possible to use the gauge function to eliminate such a term. In the $SU(2)$ example we can take as gauge function F:

$$F(\mathbf{A}_\mu, \boldsymbol{\pi}) = \partial_\mu \mathbf{A}_\mu + \tfrac{1}{2}\lambda\xi\boldsymbol{\pi} \,. \tag{19.79}$$

After a gaussian integration, the corresponding contribution to the action is

$$\mathcal{S}_F = \mathcal{S}_{\text{gauge}} + \mathcal{S}_{\text{ghost}} \tag{19.80}$$

with

$$\mathcal{S}_{\text{gauge}} = \frac{1}{2\xi e^2}\int \mathrm{d}^d x \left(\partial_\mu \mathbf{A}_\mu + \tfrac{1}{2}\lambda\xi\boldsymbol{\pi}\right)^2 \tag{19.81}$$

and

$$\mathcal{S}_{\text{ghost}} = \int \mathrm{d}^d x \left[\partial_\mu \mathbf{C}\cdot\left(\partial_\mu \bar{\mathbf{C}} - \mathbf{A}_\mu \times \bar{\mathbf{C}}\right) + \frac{\lambda\xi}{4}\mathbf{C}\left(\sigma\bar{\mathbf{C}} + \boldsymbol{\pi}\times\bar{\mathbf{C}}\right)\right]. \tag{19.82}$$

At leading order the term $\partial_\mu \boldsymbol{\pi} \cdot \mathbf{A}_\mu$ is eliminated by the choice

$$\lambda = e^2 v \,. \tag{19.83}$$

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 the propagators (equations (18.140)) have no poles at zero momentum and no IR problems are encountered:

$$\begin{aligned}
W^{(2)}_{\mu\nu} &= \frac{e^2\delta_{\mu\nu}}{k^2+m_A^2} + \frac{e^2(\xi-1)k_\mu k_\nu}{(k^2+m_A^2)(k^2+\xi m_A^2)}, \\
W^{2)}_{\boldsymbol{\pi}\boldsymbol{\pi}} &= \frac{1}{k^2+\xi m_A^2}, \\
W^{(2)}_{\mathbf{C}\bar{C}} &= \frac{1}{k^2+\xi m_A^2},
\end{aligned} \tag{19.84}$$

in which m_A is the mass of $\mathbf{A}_\mu$ in the tree approximation (equation (19.78)). Furthermore, all unphysical states have a mass which 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 one must include all particles both physical and unphysical. By showing that the S-matrix does not depend on the gauge, one proves simultaneously that the contributions of unphysical states cancels in the intermediate states and thus the S-matrix is unitary in the physical subspace. A general proof of this kind will be given in Chapter 21.

Bibliographical Notes

Non-abelian gauge fields were first discussed in

C.N. Yang and R.L. Mills, *Phys. Rev.* 96 (1954) 191; R. Utiyama, *Phys. Rev.* 101 (1956) 1597; M. Gell-Mann and S. Glashow, *Ann. Phys.* (NY) 15 (1961) 437.

Original articles on the quantization of gauge theories include

R.P. Feynman, *Acta Phys. Polon.* 24 (1963) 697; B.S. DeWitt, *Phys. Rev.* 162 (1967) 1195, 1239; L.D. Faddeev and V.N. Popov, *Phys. Lett.* 25B (1967) 29; S. Mandelstam, *Phys. Rev.* 175 (1968) 1580; M. Veltman, *Nucl. Phys.* B21 (1970) 288; G. 't Hooft, *Nucl. Phys.* B33 (1971) 173;

The axial gauge was proposed in

R.L. Arnowitt and S.I. Fickler, *Phys. Rev.* 127 (1962) 1821.

General linear gauges were proposed in

G. 't Hooft, *Nucl. Phys.* B35 (1971) 167;

and further discussed in

K. Fujikawa, B.W. Lee and A.I. Sanda, *Phys. Rev.* D6 (1972) 2923; B.W. Lee and J. Zinn-Justin, *Phys. Rev.* D7 (1973) 1049.

The quantization of constrained systems is discussed in

L.D. Faddeev, *Theor. Math. Phys.* 1 (1969) 3.

The BRS symmetry of the quantized action was discovered by

C. Becchi, A. Rouet and R. Stora, *Comm. Math. Phys.* 42 (1975) 127.

The first discussion of the ambiguity of gauge conditions is due to

V.N. Gribov, *Nucl. Phys.* B139 (1978) 1.

The Higgs mechanism was introduced in

P.W. Higgs, *Phys. Lett.* 12 (1964) 132; *Phys. Rev.* 145 (1966) 1156; F. Englert and R. Brout, *Phys. Rev. Lett.* 13 (1964) 321; G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, *Phys. Rev. Lett.* 13 (1964) 585; T.W.B. Kibble, *Phys. Rev.* 155 (1967) 1554.

It was cast into the framework of renormalizable field theories in

G. 't Hooft, *Nucl. Phys.* B35 (1971) 167.

The corresponding WT identities were derived in general form in

B.W. Lee and J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3137,

an article whose Appendix C has directly inspired Appendix A19.

A selection of original articles has been reprinted in

Gauge Theory of Weak and Electromagnetic Interactions, C.H. Lai ed. (World Scientific, Singapore 1981).

An early review is

E.S. Abers and B.W. Lee, *Phys. Rep.* 9C (1973) 1.

Finally, among the textbooks see, for example, the works of Itzykson and Zuber, and Faddeev and Slavnov, quoted in the preface.

APPENDIX A19
MASSIVE YANG–MILLS FIELDS

For completeness, we briefly explain why, in contrast with the abelian case, it is not possible to construct a renormalizable field theory in which a mass is given to the gauge field by directly adding a mass term to the action. We, therefore, consider the real time lagrangian density:

$$\mathcal{L}(\mathbf{A}_\mu) = -\frac{1}{e^2}\operatorname{tr}\left[-\tfrac{1}{4}\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}(t,x) + \tfrac{1}{2}m^2\mathbf{A}_\mu\mathbf{A}^\mu(t,x)\right]. \tag{A19.1}$$

The first problem we meet 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. It is actually algebraically convenient to first determine the hamiltonian, and use the equation of motion afterwards. The hamiltonian is here obtained by performing a Legendre transformation only on space components. The conjugated momenta $\mathbf{E}_i$ are (roman indices mean space components)

$$\mathbf{E}_i = \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{A}}_i} = -\frac{1}{e^2}\mathbf{F}_{0i}\,.$$

After an integration by parts the hamiltonian density $\mathcal{H}$ can be written

$$\begin{aligned}\mathcal{H} &= \operatorname{tr}\mathbf{E}_i\dot{\mathbf{A}}_i - \mathcal{L}\\ &= -\operatorname{tr}\left[\frac{1}{2}e^2\mathbf{E}_i^2 + \partial_i\mathbf{E}_i\mathbf{A}_0 + \frac{m^2}{2e^2}\mathbf{A}_0^2 - \frac{1}{e^2}\left(\tfrac{1}{4}\mathbf{F}_{ij}^2 + \tfrac{1}{2}m^2\mathbf{A}_i^2\right)\right].\end{aligned} \tag{A19.2}$$

Using a familiar property of the Legendre transformation we obtain the $\mathbf{A}_0$ equation of motion

$$0 = -\frac{\partial\mathcal{L}}{\partial\mathbf{A}_0} = \frac{\partial\mathcal{H}}{\partial\mathbf{A}_0} = \frac{m^2}{e^2}\mathbf{A}_0 + \partial_i\mathbf{E}_i\,. \tag{A19.3}$$

This equation determines $\mathbf{A}_0$ which can be eliminated from the hamiltonian. We then write the functional integral representation of the evolution operator in terms of the reduced hamiltonian, integrating over $\{\mathbf{E}_i, \mathbf{A}_i\}$. We note, however, that in this expression the reduced hamiltonian can be replaced by the initial hamiltonian ($A19.2$) 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 ($A19.3$) and yields a constant determinant because the coefficient of $\mathbf{A}_0^2$ in expression ($A19.2$) is field-independent. Finally, integrating with the initial hamiltonian over $\mathbf{E}_i$ at $\mathbf{A}_\mu$ fixed, we recover the initial action. Thus, the partition function can be expressed in terms of the simple euclidean action

$$\mathcal{S}(\mathbf{A}_\mu) = -\frac{1}{e^2}\int \mathrm{d}^d x\operatorname{tr}\left[\frac{1}{4}\mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2}m^2\mathbf{A}_\mu^2\right], \tag{A19.4}$$

with a flat field integration measure. In contrast with the massless gauge-invariant situation this functional integral is well-defined.

Remark. In the massless case one can follow the same strategy. The $\mathbf{A}_0$ field equation ($A19.3$), however, 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 above. The reason is that, due to space gauge invariance, Gauss's law commutes with the hamiltonian. Therefore, if it is satisfied at one time it is satisfied at all later times. An integration over $\mathbf{A}_0(t,x)$ will be infinitely redundant, as we immediately verify since this procedure leads to the initial action and thus to an undefined functional integral.

The massless limit. Provided we only consider gauge invariant correlation functions, we can introduce a gauge condition and follow all the algebraic steps of Section 19.3.2. The only modification, which is 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 (A19.4) by the substitution $\mathbf{A}_\mu \mapsto \mathbf{A}_\mu^g$ (equation (19.6)):

$$\mathbf{A}_\mu^g(x) = \mathbf{g}(x)\mathbf{A}_\mu(x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_\mu \mathbf{g}^{-1}(x) = \mathbf{g}(x)\mathbf{D}_\mu \mathbf{g}^{-1}(x) \qquad (A19.5)$$

with

$$\mathbf{D}_\mu = \partial_\mu + \mathbf{A}_\mu(x)\,. \qquad (A19.6)$$

We thus find:

$$\mathcal{S}(\mathbf{A}_\mu, \mathbf{g}) = -\frac{1}{e^2}\int \mathrm{d}^d x \,\mathrm{tr}\left[\frac{1}{4}\mathbf{F}_{\mu\nu}^2(x) - \frac{1}{2}m^2 \mathbf{D}_\mu \mathbf{g}^{-1}\mathbf{D}_\mu \mathbf{g}\right]. \qquad (A19.7)$$

We recognize a $G \times G/G$ non-linear σ-model (see Section 15.4) in which one of the G components of the symmetry group has been gauged. In contrast with the abelian case the scalar field has a self-coupling and is coupled to the gauge field.

The non-linear σ-model action is not renormalizable for dimensions $d > 2$ (see Chapter 14). If we assume that the theory has a cut-off Λ, we see that the $\mathbf{g}$-field fluctuations will not be very much 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, we do not expect the scalar field to decouple (in perturbation theory we will get IR divergences).

We also note that the complete action can be considered as the limit of a Higgs model action in which the bare mass of the Higgs field has been sent to infinity. Calling ϕ the scalar field we can view the action (A19.7) as the formal limit, when g goes to zero, of:

$$\mathcal{S}(\mathbf{A}_\mu, \phi) = \int \mathrm{d}^d x \,\mathrm{tr}\left[-\frac{1}{4e^2}\mathbf{F}_{\mu\nu}^2(x) + \frac{1}{2}\mathbf{D}_\mu\phi^\dagger \mathbf{D}_\mu\phi + \frac{1}{g}\left(\phi^\dagger\phi - (m/e)^2\right)^2\right]. \qquad (A19.8)$$

Recalling the equivalence between the non-linear σ model and the ϕ^4 field theory (see Chapter 31), we may speculate that, beyond perturbation theory, the actions (A19.7) and (A19.8) lead to the same renormalized correlation functions.

0 THE STANDARD MODEL. ANOMALIES

In Chapter 19 we have discussed the structure and the formal properties of non-abelian gauge theories. We now apply this formalism to the description of some general properties of the Standard Model of Weak, Electromagnetic and Strong Interactions. In particular, we calculate the RG β-function of the strong sector (Quantum Chromodynamics or QCD) and verify the property of asymptotic freedom.

The weak–electromagnetic theory (w.e.m.) with three quark generations, 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. Only the Higgs particle remains to be discovered (best fits of present data suggest a mass below 300 GeV) and this is the goal of new colliders like LHC at CERN. Most important will be the determination of the interactions of the Higgs particle since they are responsible for the mass of all other particles, and the mass spectrum is one most mysterious feature of the Standard Model.

Note, however, that in recent years other experiments are offered increasingly convincing evidence of solar and atmospheric neutrino oscillations. This implies at least a small modification of the Standard Model by the addition of neutrino Dirac or Majorana masses.

The situation for QCD is somewhat different. High-energy so-called inclusive results can be predicted, as a consequence of the property of large momentum asymptotic freedom (see Chapter 35). However, low energy properties cannot be derived from perturbation theory, the effective coupling being too large. Therefore, evidence for the validity of the quark confinement scheme relies on non-perturbative numerical investigations of lattice gauge theories (see Chapter 34).

In some cases, when gauge fields are coupled to axial currents, the WT identities which are necessary to prove the consistency of gauge theories are not satisfied beyond the classical approximation. They are spoiled by *anomalies*. Therefore, the second part of this chapter is devoted to the discussion of this physically important problem. Results are illustrated by some physical consequences like the π_0 decay and the solution to the $U(1)$ problem.

20.1 The Standard Model of Weak–Electromagnetic Interactions

We first briefly describe the Standard Model of weak and electromagnetic interactions which provides a physical application of the non-abelian Higgs mechanism. We restrict the discussion to one *generation* and two *flavours*, indicating eventually how it generalizes to three generations. For detailed phenomenological applications of the model the reader is referred to the literature.

The gauge group of the model is $SU(2) \times U(1)$. The form of the action for the Higgs field sector can be obtained from the action considered in the previous section by gauging a $U(1)$ subgroup of the remaining non-gauge $SU(2)$ symmetry. The scalar field is thus a $SU(2)$-doublet. The pattern of symmetry breaking is the same as before but the consequence now is that an unbroken $U(1)$ gauge symmetry remains whose generator is a linear combination of the original $U(1)$ generator and one of the $SU(2)$ generators.

It is associated with electromagnetic interactions. All fermions are either singlets or doublets. 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.

The gauge field action is a simple sum:

$$\mathcal{S}(\mathbf{A}_\mu, B_\mu) = \frac{1}{4}\int \mathrm{d}^4x \left(\mathbf{F}_{\mu\nu}^2 + B_{\mu\nu}^2\right), \tag{20.1}$$

in which $\mathbf{F}_{\mu\nu}$ is the curvature corresponding to the $SU(2)$ component and the gauge field $\mathbf{A}_\mu$ with the conventional normalization for this problem:

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu - g\mathbf{A}_\mu \times \mathbf{A}_\nu\,, \tag{20.2}$$

and $B_{\mu\nu}$ the curvature corresponding to the $U(1)$ component and the gauge field B_μ:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu\,. \tag{20.3}$$

As in Section 19.5.2, the scalar field $\boldsymbol{\phi}$ can be written as a 2×2 complex matrix forming a pair of $SU(2)$ doublets (equation (19.74)), but which has in addition a charge $g'/2$ for $U(1)$ transformations. The scalar field action is then

$$\begin{aligned}\mathcal{S}_{\text{scal.}} = \int \mathrm{d}^4x \,\Big[&\operatorname{tr}\left(\partial_\mu\boldsymbol{\phi}^\dagger - \tfrac{1}{2}ig'B_\mu\tau_3\boldsymbol{\phi}^\dagger + \tfrac{1}{2}ig\boldsymbol{\phi}^\dagger\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right)\\ &\left(\partial_\mu\boldsymbol{\phi} + \tfrac{1}{2}ig'B_\mu\tau_3\boldsymbol{\phi} - \tfrac{1}{2}ig\boldsymbol{\phi}\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right) + V(\boldsymbol{\phi})\Big]\,,\end{aligned}$$

in which $\boldsymbol{\phi}$ is a complex matrix of the form (19.74) and the potential V has the same form as in the action (19.71). Here, however, it is more convenient to parametrize $\boldsymbol{\phi}$ in terms of a complex vector $\boldsymbol{\varphi}$ forming a $SU(2)$ doublet. The scalar action then takes the form:

$$\begin{aligned}\mathcal{S}_{\text{scal.}} = \int \mathrm{d}^4x \,\Big\{&\left[\partial_\mu\boldsymbol{\varphi}^\dagger - \tfrac{1}{2}i\boldsymbol{\varphi}^\dagger\left(g'B_\mu + g\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right)\right]\left[\partial_\mu\boldsymbol{\varphi} + \tfrac{1}{2}i\left(g'B_\mu + g\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right)\boldsymbol{\varphi}\right]\\ &+V(\boldsymbol{\varphi})\Big\}\end{aligned} \tag{20.4}$$

with

$$V(\boldsymbol{\varphi}) = r\boldsymbol{\varphi}^\dagger\boldsymbol{\varphi} + \tfrac{1}{6}\lambda\left(\boldsymbol{\varphi}^\dagger\boldsymbol{\varphi}\right)^2. \tag{20.5}$$

20.1.1 The Higgs mechanism

We now assume that the potential is such that in the tree approximation the field $\boldsymbol{\varphi}$ has a non-zero expectation value proportional to the vector $(0,1)$:

$$\langle\boldsymbol{\varphi}\rangle = \frac{v}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix}. \tag{20.6}$$

With this form of the expectation value, the subgroup of $SU(2)\times U(1)$ which corresponds to the transformations,

$$\boldsymbol{\varphi} \longmapsto \mathrm{e}^{i\omega(1+\tau_3)}\,\boldsymbol{\varphi}\,, \tag{20.7}$$

leaves the expectation value of φ unchanged, and, therefore, as anticipated, $SU(2)\times U(1)$ is broken down to $U(1)$. Replacing φ by its expectation value in action (20.4) we read off the mass terms for the gauge fields in the tree approximation:

$$\frac{1}{8}v^2 \int d^4x \left[\left(g'B_\mu - gA_\mu^{(3)}\right)^2 + g^2 \left|A_\mu^{(1)} + iA_\mu^{(2)}\right|^2\right]. \tag{20.8}$$

It follows that the two components $A_\mu^{(1,2)}$ have the common mass,

$$m_W = \tfrac{1}{2}gv\,. \tag{20.9}$$

The linear combination $g'B_\mu - gA_\mu^{(3)}$ is also massive while the orthogonal combination remains massless and thus represents the photon. One defines the weak angle θ_W,

$$g'/g = \tan\theta_W\,. \tag{20.10}$$

The photon field A_μ then corresponds to

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W A_\mu^{(3)}, \tag{20.11}$$

while the massive field Z_μ is

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W A_\mu^{(3)}. \tag{20.12}$$

The Z mass is then

$$m_Z = \tfrac{1}{2}v\sqrt{g^2+g'^2}\,. \tag{20.13}$$

The components $A_\mu^{(1,2)}$ are coupled to A_μ and correspond to a charged vector field which is usually written in complex notation:

$$W_\mu^{\pm} = (A_\mu^{(1)} \pm iA_\mu^{(2)})/\sqrt{2}\,. \tag{20.14}$$

From the coupling of the charged vector bosons with the photon, we derive the relation between electric charge e and coupling constants g and g':

$$e = gg'/\sqrt{g^2+g'^2} = g\sin\theta_W = g'\cos\theta_W\,. \tag{20.15}$$

20.1.2 Coupling to leptons

We consider here only the electron and the corresponding neutrino ν_e since the couplings to other leptons (μ, ν_μ and τ, ν_τ) have exactly the same structure. In the Standard Model the neutrino and the left-handed part of the electron are combined into a left-handed doublet L of $SU(2)$:

$$L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \tag{20.16}$$

where $e_L = \frac{1}{2}(1-\gamma_5)e^-$, while the right-handed part of the electron $R = \frac{1}{2}(1+\gamma_5)e^-$ forms a $SU(2)$ singlet. In addition, L and R have different $U(1)$ charges. The lepton contribution to the action is

$$\mathcal{S}_{\text{leptons}} = -\int d^4x \left[\bar{R}\gamma_\mu\left(\partial_\mu - ig'B_\mu\right)R + \bar{L}\gamma_\mu\left(\partial_\mu - \tfrac{1}{2}ig'B_\mu + \tfrac{1}{2}ig\mathbf{A}_\mu\cdot\boldsymbol{\tau}\right)L\right]. \tag{20.17}$$

In this model, since the left-handed and right-handed components of the fermion fields are treated differently, the breaking of parity symmetry is explicit.

The $SU(2)$ symmetry forbids a mass term for the electron field. On the other hand, a coupling between the scalar field and the fermions is allowed by the symmetry:

$$\mathcal{S}_{\text{Yukawa}} = G_e \int \mathrm{d}^4 x \left(\bar{R}\varphi^\dagger L + \bar{L}\varphi R\right). \tag{20.18}$$

If we replace the φ-field by its expectation value, we see that the spontaneous breaking of the $SU(2)\times U(1)$ symmetry generates the electron mass m_e which is, therefore, calculable, but, in the absence of new dynamic principle, in terms of another arbitrary parameter, the Yukawa coupling constant G_e:

$$m_e = vG_e \,. \tag{20.19}$$

Note that this implies that the Yukawa coupling of leptons is proportional to their mass and, therefore, the perturbative approximation becomes worse for heavier leptons.

The coupling constant G (the Fermi constant), characteristic of the strength of weak interactions, is defined in terms of an effective low energy current–current and thus four-fermion interaction:

$$\frac{G}{\sqrt{2}} \int \mathrm{d}^4 x \, J_\mu(x) J_\mu^\dagger(x). \tag{20.20}$$

The contribution to the charged current J_μ coming from the electron and the neutrino has the form:

$$J_\mu(x) = \bar{e}(1-\gamma_5)\gamma_\mu \nu_{\mathrm{L}} = 2\bar{e}_{\mathrm{L}}\gamma_\mu \nu_{\mathrm{L}} \,. \tag{20.21}$$

The relation between G 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

$$\frac{1}{8} g^2 v^2 \left[\left(A_\mu^{(1)}\right)^2 + \left(A_\mu^{(2)}\right)^2\right] + \frac{ig}{2}\bar{L}\gamma_\mu \left(A_\mu^{(1)}\tau_1 + A_\mu^{(2)}\tau_2\right) L \,. \tag{20.22}$$

Completing squares we immediately obtain the result of the integration:

$$\frac{1}{2v^2}\left[\left(\bar{L}\gamma_\mu \tau_1 L\right)^2 + \left(\bar{L}\gamma_\mu \tau_2 L\right)^2\right] = \frac{2}{v^2}\bar{\nu}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}} \bar{e}_{\mathrm{L}}\gamma_\mu \nu_{\mathrm{L}} \,. \tag{20.23}$$

Comparing with definition (20.20), we conclude,

$$G/\sqrt{2} = 1/2v^2 = g^2/8m_W^2 \,. \tag{20.24}$$

The phenomenological Fermi model of low energy charged weak interactions determines all parameters of the Standard Model but two, for example, the weak angle θ_{W} and the Higgs field self-coupling λ which have to be deduced from additional experimental results. The recent direct measurements of the W and Z masses, for example, determine the parameter θ_{W}. Note that at leading order the masses can be rewritten as

$$M_W^2 = \frac{e^2}{4\sqrt{2} G \sin^2\theta_{\mathrm{W}}},$$

that is,

$$M_W = \left(\frac{\pi\alpha}{\sqrt{2}G}\right)^{1/2} = \frac{38}{\sin\theta_{\mathrm{W}}}\,\mathrm{GeV} \tag{20.25}$$

and

$$M_Z = \frac{M_W}{\cos\theta_{\mathrm{W}}}. \tag{20.26}$$

The present (1998) experimental values are $m_W = 80.4 \pm 0.1$ GeV, $m_Z = 91.19 \pm 0.01$ GeV. They lead to a weak angle $\sin^2\theta_{\mathrm{W}} = 0.225 \pm 0.004$, an estimate which is in good agreement with estimates coming from the couplings to neutral currents (an even better agreement is obtained when radiative corrections are taken into account). The parameter λ mainly determines the Higgs mass, and since the Higgs particle has not yet been observed, its value is unknown.

The coupling of the charged vector bosons $W^{\pm}$ to $e^-\nu_e$ is obtained by introducing the definition (20.14) into equation (20.17):

$$\begin{aligned}\frac{g}{2}\bar{L}\gamma_\mu\left(\tau^1 A^{(1)}_\mu + \tau^2 A^{(2)}_\mu\right)L &= \frac{g}{2\sqrt{2}}\left[\left(\bar{\nu}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}} + \bar{e}_{\mathrm{L}}\gamma_\mu\nu_{\mathrm{L}}\right)\left(W^+_\mu + W^-_\mu\right)\right.\\ &\quad\left.+\left(\bar{\nu}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}} - \bar{e}_{\mathrm{L}}\gamma_\mu\nu_{\mathrm{L}}\right)\left(W^+_\mu - W^-_\mu\right)\right]\\ &= \frac{g}{\sqrt{2}}\left(\bar{\nu}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}} W^+_\mu + \bar{e}_{\mathrm{L}}\gamma_\mu\nu_{\mathrm{L}} W^-_\mu\right).\end{aligned} \tag{20.27}$$

Using definitions (20.10–20.12) we also obtain the couplings of fermions to the neutral vector fields A_μ and Z_μ:

$$\frac{eZ_\mu}{\sin 2\theta_{\mathrm{W}}}\left(2\sin^2\theta_{\mathrm{W}}\bar{e}_{\mathrm{R}}\gamma_\mu e_{\mathrm{R}} - \cos 2\theta_{\mathrm{W}}\bar{e}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}} + \bar{\nu}_{\mathrm{L}}\gamma_\mu\nu_{\mathrm{L}}\right) - eA_\mu\left(\bar{e}_{\mathrm{R}}\gamma_\mu e_{\mathrm{R}} + \bar{e}_{\mathrm{L}}\gamma_\mu e_{\mathrm{L}}\right). \tag{20.28}$$

20.1.3 Coupling to hadrons

It is now easy to add interaction terms for quarks in the model. We first consider again only one generation corresponding to two flavours, calling **u** and **d** the corresponding quarks. 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 $\mathbf{Q}_{\mathrm{L}}$.

$$\mathbf{Q}_{\mathrm{L}} = \{\mathbf{u}_{\mathrm{L}}\ ,\ \mathbf{d}_{\mathrm{L}}\}\ . \tag{20.29}$$

All right-handed components form $SU(2)$ singlets. The coupling to quarks can be written as

$$\begin{aligned}\mathcal{S}_{\mathrm{quarks}} = -\int \mathrm{d}^4x\,&\left[\bar{\mathbf{Q}}_{\mathrm{L}}\left(\not\partial + \tfrac{1}{2}ig'Y_{\mathrm{L}}\not B + \tfrac{1}{2}ig\not\mathbf{A}\cdot\boldsymbol{\tau}\right)\mathbf{Q}_{\mathrm{L}}\right.\\ &\left.+\bar{\mathbf{Q}}_{1\mathrm{R}}\left(\not\partial + \tfrac{1}{2}ig'Y_{1\mathrm{R}}\not B\right)\mathbf{Q}_{1\mathrm{R}} + \bar{\mathbf{Q}}_{2\mathrm{R}}\left(\not\partial + \tfrac{1}{2}ig'Y_{2\mathrm{R}}\not B\right)\mathbf{Q}_{2\mathrm{R}}\right].\end{aligned} \tag{20.30}$$

Calling T_3 the eigenvalue of τ_3 the generator of $SU(2)$, Y the $U(1)$ charge and q the electric charge, we derive from the relations (20.11,20.12,20.15) for each fermion:

$$q = \tfrac{1}{2}(T_3 + Y). \tag{20.31}$$

The charge q_1 of the first component of the fermion being given, a set of constraints on the $U(1)$ charges Y_{L}, $Y_{1,2\mathrm{R}}$ follow:

$$\begin{aligned} q_1 &= \tfrac{1}{2}(Y_{\mathrm{L}}+1)\,, \qquad & q_1 &= \tfrac{1}{2}Y_{1\mathrm{R}}\,, \\ q_2 &= \tfrac{1}{2}(Y_{\mathrm{L}}-1), \qquad & q_2 &= \tfrac{1}{2}Y_{2\mathrm{R}}\,. \end{aligned} \tag{20.32}$$

We note that all $U(1)$ charges are determined by the electric charge of the first component of the fermion doublet. Let us verify these relations in the case of the leptons:

$$q_1 = 0 \Rightarrow Y_{\mathrm{L}} = -1\,,\ q_2 = -1\,,\ Y_{1\mathrm{R}} = 0\,,\ Y_{2\mathrm{R}} = -2\,. \tag{20.33}$$

The proton being a **uud** state and the neutron a **udd** state, the charges of the **u** and **d** quarks are $2/3$ and $-1/3$, respectively. The relations (20.31) then imply

$$Y_{\mathrm{L}} = 1/3\,,\ Y_{1\mathrm{R}} = 4/3\,,\ Y_{2\mathrm{R}} = -2/3\,. \tag{20.34}$$

Note that the charges of quarks are compatible with the $SU(2)$ doublet assignment. We shall verify in Section 20.5 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 in the case of 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 Yukawa couplings are

$$G_{q2}\bar{\mathbf{Q}}_{2\mathrm{R}}\varphi^\dagger\mathbf{Q}_{\mathrm{L}} + G_{q1}\bar{\mathbf{Q}}_{\mathrm{L}}i\tau_2\varphi^*\mathbf{Q}_{1\mathrm{R}} + \text{ h.c.}\,, \tag{20.35}$$

which can, therefore, provide masses for the two quarks. This is at least the situation for one generation. However, six quarks in three generations have been discovered (see table 20.1). Moreover, from the Z decays one infers that the number of generations with "light" neutrinos (i.e. with a mass below 45 GeV) is exactly three.

Therefore, in the interactions (20.35) the spinors which appear on the right and the left then need not be 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:

$$\bar{\mathbf{Q}}^\alpha_{1\mathrm{R}}M_{\alpha\beta}\mathbf{Q}^\beta_{1\mathrm{L}} + \bar{\mathbf{Q}}^\alpha_{1\mathrm{L}}M^\dagger_{\alpha\beta}\mathbf{Q}^\beta_{1\mathrm{R}}\,, \tag{20.36}$$

for the quarks of charge $2/3$, and a similar one for the charge $-1/3$ quarks. Performing independent unitary transformations $\mathbf{U}_{\mathrm{R,L}}$ on the right and left quark components it is possible to replace the matrix $\mathbf{M}$ by a real diagonal matrix $\mathcal{M}$:

$$\mathbf{U}^\dagger_{\mathrm{R}}\mathbf{M}\mathbf{U}_{\mathrm{L}} = \mathcal{M}\,, \tag{20.37}$$

In this representation the quarks are mass eigenstates. However, the weak interactions no longer have the simple form (20.30) because the unitary transformations on the quark components $\mathbf{Q}_{1\mathrm{L}}$ and $\mathbf{Q}_{2\mathrm{L}}$ 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 (because there are three generations at present), the Kobayashi–Maskawa matrix (KM) which relates the quark mass eigenstates **d**, **s** and **b** to the quarks appearing in the weak interactions:

$$\left[\mathbf{Q}_2^\alpha\right]_{\text{weak int.}} = U_{\alpha\beta}\left[\mathbf{Q}_2^\beta\right]_{\text{mass eigenst.}}\,. \tag{20.38}$$

With only two generations (**d** and **s**) it is possible to cast the matrix into the form:

$$\mathbf{U}_\mathrm{C} = \begin{pmatrix} \cos\theta_\mathrm{C} & \sin\theta_\mathrm{C} \\ -\sin\theta_\mathrm{C} & \cos\theta_\mathrm{C} \end{pmatrix},$$

in which θ_C is the celebrated Cabibbo angle, after unobservable changes of the relative phases between the quarks. In the presence of the third **b** quark, the 3×3 KM matrix can be parametrized in terms of three rotation angles and one phase parameter. This phase may be responsible for the observed direct CP violation in neutral kaon and B_0 meson decay.

Table 20.1

Quarks and leptons. The three generations (1998).

Charge 2/3 quarks	Charge $-1/3$ quarks	Charge -1 leptons	Neutrinos
u, $m = 1.5$ to 5 MeV	**d**, $m = 3$ to 9 MeV	e, $m = 0.511$ MeV	ν_e, $m < 10$ eV
c, $m = 1.1$ to 1.4 GeV	**s**, $m = 60$ to 170 MeV	μ, $m = 105.6$ MeV	ν_μ, $m < 0.17$ MeV
t, $m = 174 \pm 6$ GeV	**b**, $m = 4.1$ to 4.4 GeV	τ, $m = 1.777$ GeV	ν_τ, $m < 19$ MeV

Beyond the Standard Model. Theoretical speculations based on the search for a unifying simple group including $U(1) \times SU(2) \times SU(3)$ as a subgroup, have first 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 Higgs field. Running the three independent couplings to higher energies an apparent unification was observed at energies of order 10^{15} GeV. The non-observation of the predicted proton decay 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.

The experimental confirmation of neutrino oscillations can be explained if neutrinos have masses

$$m^2(\nu_\mu) - m^2(\nu_e) = O(10^{-5}\mathrm{eV}^2)\,, \quad m^2(\nu_\tau) - m^2(\nu_\mu) = O(10^{-3}\mathrm{eV}^2),$$

and the mass eigenstates differ from the linear combinations of neutrinos appearing in the weak interactions. This implies mixing angles like in the quark sector, as described in Section 20.1.3 (but large mixing angles seem to be favoured). A major problem comes from the very small neutrino masses. A neutrino mass matrix clearly implies a modification of the Standard Model.

20.1.4 The problem of the elementary scalar field

The Standard Model has (Higgs) scalar fields as an essential ingredient. This is the source of several difficulties. The Higgs field is responsible for the masses of all particles, but in the Standard Model these masses are all given in terms of arbitrary parameters, like for example the Yukawa couplings which determine the fermions masses. This could perhaps be expected from an effective low energy theory. Yet puzzling is the diversity of these couplings. If the couplings were "natural", that is, of order unity, all fermion masses would be in the few 100 GeV range, as the W and Z masses or the Higgs expectation value. But in this sense only the top quark (**t**) mass, which is about 174 GeV, is natural. Instead, even in the quark sector only the masses span about five order of magnitudes, and taking into account the lepton sector makes the problem much worse.

Another perhaps even more fundamental problem is related to the scalar field mass renormalization. Generically the scalar bare and physical masses are expected to be of the order of the momentum "cut-off" which gives the scale of some new physics. It is only by *fine tuning* the scalar bare mass that one can render the renormalized mass much smaller than the cut-off. In the statistical physics interpretation of the ϕ^4 theory the divergence of the correlation length (the inverse physical mass in the particle language) is obtained by adjusting the temperature, and thus 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. To get a rough idea about the severity of the problem we can use perturbation theory, since the ϕ^4 coupling λ cannot be too large without endangering other perturbative calculations. Neglecting other interactions, the one-loop mass counter-term with a momentum cut-off Λ is

$$\delta m_0^2 = \frac{\lambda}{16\pi^4}\int \frac{\mathrm{d}^4 p}{p^2(1+p^2/\Lambda^2)^2} = \frac{\lambda}{16\pi^2}\Lambda^2 .$$

The Higgs mass m_{H} at leading order is given by

$$m_{\mathrm{H}}^2 = \tfrac{1}{3}\lambda v^2 .$$

Therefore,

$$f \equiv \delta m_0^2/m_{\mathrm{H}}^2 = \frac{3}{16\pi^2}(\Lambda/v)^2 .$$

Considering the unexplained range of fermion masses, it is difficult to decide how much fine tuning is acceptable. For example, $f = 10$ cannot be excluded from such a crude argument. Then Λ is of the order of 10 TeV, about the highest energy of new collider LHC. If more fine tuning is accepted the scale of new physics could be out of reach in the new generation of experiments, at least from this argument alone. On the other hand, an absence of any new physics below Planck's scale (10^{19} GeV) is difficult to believe.

Fine tuning solutions. At the scale of new physics the scalar field problem must be cured. Two 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 name of *technicolour* (see next section). Such models have problems generating fermion masses and no widely accepted model has been proposed.

(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 17.6), 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 main difficulty with this approach is that none of the superpartners of existing particles have yet been found. Moreover, the mechanism of spontaneous supersymmetry breaking is not fully understood.

20.2 Quantum Chromodynamics: Renormalization Group

We now concentrate on Quantum Chromodynamics (QCD) a model of quarks and gluons that produces the observed hadrons and their Strong Interactions, neglecting completely the weak and electromagnetic interactions which we have described in the preceding section. Quantum Chromodynamics, as it stands today, consists in a set of quarks characterized by a *flavour quantum number*, relevant for Weak Interactions, which are also triplets of a gauged symmetry, the $SU(3)$ *colour*, realized in the symmetric phase. Their interactions are mediated by the corresponding gauge fields (called ***gluons***):

$$\mathcal{S}(\mathbf{A}_\mu, \bar{\mathbf{Q}}, \mathbf{Q}) = -\int \mathrm{d}^4 x \left[\frac{1}{4g^2} \operatorname{tr} \mathbf{F}_{\mu\nu}^2 + \sum_{\text{flavours}} \bar{\mathbf{Q}}_f \left(\not{\mathbf{D}} + m_f \right) \mathbf{Q}_f \right]. \tag{20.39}$$

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 (AF) at large momentum in four dimensions (like the non-linear σ-model in two dimensions). In Chapter 35 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 have ever been observed at large distance (but they manifest themselves indirectly in the jet physics). 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 *confinement* hypothesis (see Chapter 34).

20.2.1 RG equations in the covariant gauge

We first discuss the gauge dependence of RG equations and functions of pure gauge theories in the covariant gauge (19.47), that is, the dependence on the parameter ξ. A short discussion of the abelian case can be found in Section 18.8. We call Z_A the gauge field renormalization constant and Z_g the renormalization constant of the coupling constant α:

$$\alpha = g^2/4\pi .$$

α in Strong Interactions is in general denoted α_S to distinguish it from its QED analogue. In this section no confusion is possible. It will be shown in Section 21.6.2 that for such gauges, as in the abelian case, the gauge fixing term is not renormalized. Therefore,

$$\xi_0 = \xi Z_A / Z_g . \tag{20.40}$$

In terms of the renormalization scale μ the RG equation for the gauge field n-point function reads:

$$\left[\mu\frac{\partial}{\partial\mu}+\beta(\alpha,\xi)\frac{\partial}{\partial\alpha}+\delta(\alpha,\xi)\xi\frac{\partial}{\partial\xi}-\frac{n}{2}\eta_A(\alpha,\xi)\right]\Gamma^{(n)}(\mu,\alpha,\xi)=0\,, \tag{20.41}$$

where,

$$\delta(\alpha,\xi)\equiv\xi^{-1}\,\mu\frac{\partial}{\partial\mu}\bigg|_{\alpha_0\,,\,\xi_0\ \text{fixed}}\xi=\beta(\alpha,\xi)-\eta_A(\alpha,\xi). \tag{20.42}$$

We prove in Section 21.7 that the bare correlation functions of gauge invariant operators are gauge independent. This in particular implies that they are independent of ξ_0. The same property 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. Let us call Γ such a correlation function. It satisfies

$$\frac{\partial}{\partial\xi}\bigg|_{\alpha_0\,,\,\text{cut-off fixed}}\Gamma=\left(\frac{\partial}{\partial\xi}+\rho(\alpha,\xi)\frac{\partial}{\partial\alpha}\right)\Gamma(\mu,\alpha,\xi)=0 \tag{20.43}$$

with

$$\rho(\alpha,\xi)=\frac{\partial\alpha}{\partial\xi}\bigg|_{\alpha_0\,,\,\text{cut-off fixed}}. \tag{20.44}$$

Γ also satisfies a RG equation which we assume to be homogeneous:

$$\left[\mu\frac{\partial}{\partial\mu}+\beta(\alpha,\xi)\frac{\partial}{\partial\alpha}+\delta(\alpha,\xi)\xi\frac{\partial}{\partial\xi}-\eta_\Gamma(\alpha,\xi)\right]\Gamma(\mu,\alpha,\xi)=0\,. \tag{20.45}$$

Using equation (20.43) to eliminate $\partial/\partial\xi$, we obtain a new RG equation for Γ:

$$\left[\mu\frac{\partial}{\partial\mu}+\tilde\beta(\alpha,\xi)\frac{\partial}{\partial\alpha}-\eta_\Gamma(\alpha,\xi)\right]\Gamma(\mu,\alpha,\xi)=0 \tag{20.46}$$

with

$$\tilde\beta=\beta-\xi\delta\rho\,. \tag{20.47}$$

Writing then the compatibility condition between the two linear equations (20.43) and (20.46) we obtain two equations:

$$\left(\frac{\partial}{\partial\xi}+\rho(\alpha,\xi)\frac{\partial}{\partial\alpha}\right)\eta_\Gamma(\alpha,\xi)=0\,, \tag{20.48}$$

$$\left(\frac{\partial}{\partial\xi}+\rho(\alpha,\xi)\frac{\partial}{\partial\alpha}\right)\tilde\beta=\frac{\partial\rho}{\partial\alpha}\tilde\beta\,. \tag{20.49}$$

The first equation expresses that, as expected, the multiplicative renormalization of Γ is independent of ξ_0. The second equation shows that the zeroes of $\tilde\beta$ are gauge independent. Differentiating the equation with respect to α one also finds that the slope of β at its zeroes is gauge independent. Finally, one verifies that in a minimal subtraction scheme the function ρ vanishes. In dimensional regularization the relation between α_0 and α takes the form:

$$\alpha_0=\mu^\varepsilon\alpha Z_g=\mu^\varepsilon\alpha\left(1+\frac{Z_g^1(\alpha,\xi)}{\varepsilon}+\frac{Z_g^2(\alpha,\xi)}{\varepsilon^2}+\cdots\right). \tag{20.50}$$

The important point is that the term without pole in ε in the expansion of Z_g is ξ independent. Using the definition (20.44) of ρ we then find:

$$0 = \rho \left(1 + \frac{\partial Z_g^1}{\partial \alpha}\frac{1}{\varepsilon} + \cdots \right) + \alpha \left(\frac{\partial Z_g^1}{\partial \xi}\frac{1}{\varepsilon} + O\left(\frac{1}{\varepsilon^2}\right)\right). \tag{20.51}$$

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 minimal subtraction scheme the β-function and η_Γ are independent of ξ.

20.2.2 The RG β-function at one-loop order

We now calculate the RG β-function at leading order in a gauge theory corresponding to a simple group G in particular to verify asymptotic freedom, because no simple explanation has yet been proposed which allows one to understand the sign without explicit calculation.

The calculation can be done by various methods, for example, we could use the background field method as in the case of the models on homogeneous spaces in Section 15.6.2. Here, instead, we calculate directly the β-function from the renormalization of the gauge coupling constant as defined by the fermion-gauge field vertex. We thus need the divergent parts of the gauge field and fermion two-point functions, and the fermion gauge field three-point function. We work in the Feynman gauge and use *dimensional regularization*. The normalizations of vertices and propagators are those given in Section 19.4.

The gauge field two-point function. Four diagrams contribute to the two-point function, corresponding to the gauge field loops, the Faddeev–Popov ghost loop and the fermion loops (see figure 20.1).

Fig. 20.1 The gauge field two-point function at one-loop (dotted lines represent ghosts).

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 18.9 up to a simple geometric factor. Diagram (a) is given by

$$(\mathrm{a}) = \tfrac{1}{2} f_{acd} f_{bcd} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{N_{\mu\nu}(k,q)}{q^2 (k+q)^2}, \tag{20.52}$$

with

$$\begin{aligned} N_{\mu\nu}(k,q) = {} & \delta_{\mu\nu}(5k^2 + 2k\cdot q + 2q^2) + k_\mu k_\nu (d-6) + (q^\mu k_\nu + q^\nu k_\mu)(2d-3) \\ & + 2q_\mu q_\nu (2d-3). \end{aligned} \tag{20.53}$$

To calculate the diagrams, we project the integrand over $\delta_{\mu\nu}$ and $k_\mu k_\nu$ (see equations (9.68,9.69)), and use repeatedly the identity:

$$2k\cdot q = (k+q)^2 - k^2 - q^2.$$

We set

$$f_{acd}f_{bcd} = C(G)\delta_{ab}\,. \tag{20.54}$$

A short calculation yields the divergent part:

$$(\mathrm{a})_{\mathrm{div}} = \delta_{ab}\frac{C(G)}{12}\left(19k^2\delta_{\mu\nu} - 22k_\mu k_\nu\right)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.55}$$

Diagram (c) is given by

$$(\mathrm{c}) = -f_{acd}f_{bcd}\int\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{q_\mu(k+q)_\nu}{q^2(k+q)^2}. \tag{20.56}$$

The divergent part is

$$(\mathrm{c})_{\mathrm{div}} = \delta_{ab}\left(k^2\delta_{\mu\nu} + 2k_\mu k_\nu\right)\frac{1}{12}C(G)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.57}$$

Note that both divergent contributions are not separately transverse. By adding them we get the divergent part of the two-point function in the absence of fermions, which now is transverse as expected:

$$\left[\Gamma^{(2)ab}_{\mu\nu}\right]_{\mathrm{div}} = \delta_{ab}\left(k^2\delta_{\mu\nu} - k_\mu k_\nu\right)\frac{5}{3}C(G)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.58}$$

Calling Z_A and Z_g the renormalization constants of the gauge field and the coupling constant g^2 we obtain the relation at one-loop order:

$$\frac{Z_A}{Z_g} = 1 + \frac{5}{3}C(G)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.59}$$

Using the result (18.108), we can add the fermion contribution and finally find:

$$\frac{Z_A}{Z_g} = 1 + \left(\frac{5}{3}C(G) - \frac{4}{3}T(R)\right)\frac{g^2}{8\pi^2\varepsilon}, \tag{20.60}$$

where the fermions belong to the representation R and $T(R)$ is the trace of the square of the generators in this representation:

$$\operatorname{tr} t^a t^b = -\delta_{ab}T(R). \tag{20.61}$$

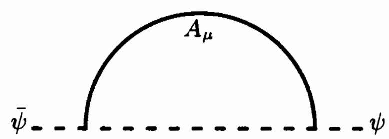


Fig. 20.2 One-loop contribution to the fermion two-point function.

The fermion two-point function. One diagram contributes to the fermion two-point function $\Gamma^{(2)}$ (see figure 20.2) which differs from its QED counterpart only by a geometric factor:

$$\Gamma^{(2)}_{1\,\mathrm{loop}}(k) = g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} t^a \gamma_\mu \frac{1}{i\not{q} + m} t^a \gamma_\mu \frac{1}{(k-q)^2} \tag{20.62}$$

(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 $\not{k}$. 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:

$$Z_\mathrm{F} = 1 - C(R)\frac{g^2}{8\pi^2\varepsilon}\,, \quad \text{with} \quad t^a t^a = -C(R)\mathbf{1}\,. \tag{20.63}$$

The gauge field fermion vertex. Two diagrams contribute at one-loop order (see figure 20.3), the first has a QED counterpart, the second being specific to a non-abelian theory.

Fig. 20.3 The gauge field fermion vertex at one-loop.

$$(\mathrm{a}) = g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} t^b \gamma_\nu \frac{1}{i\not{q} + m} t^a \gamma_\mu \frac{1}{i\not{q} + i\not{k} + m} t^b \gamma_\nu \frac{1}{(p_1 - q)^2}\,. \tag{20.64}$$

To calculate the divergent part of the integral we multiply by γ_μ and take the trace. We also use the identity:

$$t^b t^a t^b = \left(\tfrac{1}{2}C(G) - C(R)\right) t^a.$$

We then find:

$$(\mathrm{a})_{\mathrm{div.}} = \left(C(R) - \frac{1}{2}C(G)\right) t^a \gamma_\mu \frac{g^2}{8\pi^2\varepsilon}\,, \tag{20.65}$$

$$(\mathrm{b}) = i f_{abc} g^2 \int \frac{\mathrm{d}^d q}{(2\pi)^d} t^b \gamma_\nu \frac{1}{i\not{p}_1 - i\not{q} + m} t^c \gamma_\rho \frac{V_{\mu\nu\rho}(k,q,-k-q)}{q^2(k+q)^2}\,, \tag{20.66}$$

with (equation (19.57)):

$$V_{\mu\nu\rho}(k,q,r) = (r-q)_\mu \delta_{\nu\rho} + (k-r)_\nu \delta_{\rho\mu} + (q-k)_\rho \delta_{\mu\nu}\,. \tag{20.67}$$

The divergent part is

$$(\mathrm{a})_{\mathrm{div.}} = \frac{3}{2} C(G) t^a \gamma_\mu \frac{g^2}{8\pi^2\varepsilon}\,. \tag{20.68}$$

It follows that

$$Z_{\mathrm{F}} Z_A^{1/2} = 1 - (C(R) + C(G))\,\frac{g^2}{8\pi^2\varepsilon}, \tag{20.69}$$

and, therefore, (equation (20.63)):

$$Z_A = 1 - 2C(G)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.70}$$

Using finally the result (20.60) we obtain Z_g:

$$Z_g = 1 - \left(\tfrac{11}{3}C(G) - \tfrac{4}{3}T(R)\right)\frac{g^2}{8\pi^2\varepsilon}. \tag{20.71}$$

The β-function at one-loop order follows:

$$\beta(g^2) = -\varepsilon\left[\frac{\mathrm{d}\ln(g^2 Z_g)}{\mathrm{d}g^2}\right]^{-1} = -\left[\tfrac{11}{3}C(G) - \tfrac{4}{3}T(R)\right]\frac{g^4}{8\pi^2} + O\left(g^6\right). \tag{20.72}$$

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

$$C(G) = N\,, \qquad T(R) = \tfrac{1}{2}N_{\mathrm{F}}\,,$$

and, therefore,

$$\beta(g^2) = -\left(\frac{11N}{3} - \frac{2N_{\mathrm{F}}}{3}\right)\frac{g^4}{8\pi^2} + O\left(g^6\right). \tag{20.73}$$

The theory is asymptotically free, that is, the β-function is negative for small coupling for

$$N_{\mathrm{F}} < 11N/2\,, \tag{20.74}$$

which, in the case of $SU(3)$, means at most 16 flavours. If this condition is met, $g = 0$ is a UV fixed point (for details see Chapter 35).

20.3 The Abelian Anomaly

We have pointed out in Sections 18.5, 19.4 that none of the standard regularization methods can deal in a straightforward way with one-loop diagrams in the case of chiral gauge fields. We now show that indeed gauge theories with massless fermions and chiral symmetry can be found where the axial current is not conserved. The divergence of the axial current, when it does not vanish, is called an *anomaly*. This leads in particular 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 theory of weak electromagnetic interactions, the electromagnetic decay of the π_0 meson, or the $U(1)$ problem.

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 in gauge theories when chiral properties are involved. This reduces the problem to the discussion of fermions in the background of gauge fields, or equivalently to the properties of the determinant of the gauge covariant Dirac operator.

20.3.1 Abelian axial current and abelian vector gauge field

We first 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:

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^4x\,\bar{\psi}(x)\not{D}\psi(x), \quad \not{D} \equiv \not{\partial} + ie\not{A}\,, \tag{20.75}$$

and the corresponding functional integral

$$\mathcal{Z}(A_\mu) = \int \left[\mathrm{d}\psi \mathrm{d}\bar{\psi}\right] \exp\left[-\mathcal{S}(\psi,\bar{\psi})\right] = \det \not{D}\,. \tag{20.76}$$

In what follows we denote by $\langle\bullet\rangle$ expectation values with respect to the measure.

We can find regularizations which preserve gauge invariance, and since the fermions are massless, chiral symmetry. We would, therefore, naively expect the corresponding axial current to be 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

$$\psi_\theta(x) = \mathrm{e}^{i\theta(x)\gamma_5}\,\psi(x), \qquad \bar{\psi}_\theta(x) = \bar{\psi}(x)\,\mathrm{e}^{i\theta(x)\gamma_5}, \tag{20.77}$$

yields the axial current $J^5_\mu(x)$. For the action (20.75) one finds,

$$\delta\mathcal{S} = \int \mathrm{d}^4x\,\partial_\mu\theta(x)J^5_\mu(x) \quad \text{with} \quad J^5_\mu(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x). \tag{20.78}$$

After the transformation (20.77), $\mathcal{Z}(A_\mu)$ becomes

$$\mathcal{Z}(A_\mu,\theta) = \det\left[\mathrm{e}^{i\gamma_5\theta(x)}\,\not{D}\,\mathrm{e}^{i\gamma_5\theta(x)}\right]. \tag{20.79}$$

Since $\mathrm{e}^{i\gamma_5\theta}$ has a determinant which is unity, one would naively conclude that $\mathcal{Z}(A_\mu,\theta) = \mathcal{Z}(A_\mu)$ and, therefore, that the current $J^5_\mu(x)$ is conserved. This is a conclusion we now check by an explicit calculation of the expectation value of $\partial_\mu J^5_\mu(x)$ in the case of the action (20.75).

Remarks.

(i) For any regularization which is consistent with the hermiticity of γ_5

$$\left|\mathcal{Z}(A_\mu,\theta)\right|^2 = \det\,(\not{D}\not{D}^\dagger).$$

Therefore, an anomaly can appear only in the imaginary part of $\ln\mathcal{Z}$.

(ii) If the regularization is gauge invariant, $\mathcal{Z}(A_\mu,\theta)$ is also gauge invariant. Therefore, a possible anomaly will also be gauge invariant. One regularization scheme that has the required property is based on regulator fields. But as the discussion of Section 9.5.2 has shown, 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 the anomaly. The operator $\partial_\mu J^5_\mu(x)$ has dimension 4 and since a possible anomaly is a large momentum or short distance effect, $\langle \partial_\mu J^5_\mu(x)\rangle$ can only be a local function of A_μ of dimension 4. In addition parity implies that it is proportional to the completely antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$. This determines $\langle \partial_\mu J^5_\mu(x)\rangle$ up to a multiplicative constant,

$$\langle \partial_\lambda J^5_\lambda(x)\rangle \propto e^2 \epsilon_{\mu\nu\rho\sigma}\partial_\mu A_\nu(x)\partial_\rho A_\sigma(x) \propto e^2\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\,,$$

$F_{\mu\nu}$ being the electromagnetic tensor. The possible anomaly is always gauge invariant.

To find 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 $\langle \partial_\lambda J^5_\lambda(x)\rangle$ in powers of A. We define the three-point function:

$$\begin{aligned}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) &= \frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)}\left\langle J^5_\lambda(k)\right\rangle\Big|_{A=0}\,, \\ &= \frac{\delta}{\delta A_\mu(p_1)}\frac{\delta}{\delta A_\nu(p_2)} i\,\mathrm{tr}\left[\gamma_5\gamma_\lambda \not{D}^{-1}(k)\right]\Big|_{A=0}\,.\end{aligned} \tag{20.80}$$

$\Gamma^{(3)}$ is the sum of the two Feynman diagrams of figure 20.4.

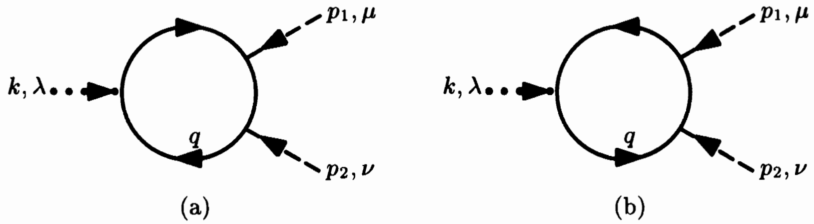


Fig. 20.4 Anomalous diagrams.

The contribution of diagram (a) is

$$(\mathrm{a}) \mapsto \frac{e^2}{(2\pi)^4}\,\mathrm{tr}\left[\int \mathrm{d}^4 q\, \gamma_5\gamma_\lambda(\not{q}+\not{k})^{-1}\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right], \tag{20.81}$$

and the contribution of diagram (b) is obtained by exchanging $p_1, \gamma_\nu \leftrightarrow p_2, \gamma_\nu$.

Power counting tells us that the function $\Gamma^{(3)}$ may have a linear divergence which, due to the presence of the γ_5 factor, must be proportional to $\epsilon_{\lambda\mu\nu\rho}$, symmetric in the exchange $p_1, \gamma_\nu \leftrightarrow p_2, \gamma_\nu$, and thus proportional to

$$\epsilon_{\lambda\mu\nu\rho}(p_1 - p_2)_\rho\,. \tag{20.82}$$

On the other hand, by commuting γ_5 we notice that $\Gamma^{(3)}$ is formally a symmetric function of the three sets of external arguments. A divergence breaks the symmetry between external arguments. Therefore, a symmetric regularization of the kind we shall adopt leads to a finite result. The result is not ambiguous because a possible ambiguity again is proportional to (20.82).

In the same way if the regularization is consistent with vector gauge invariance the WT identity

$$p_{1\mu}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = 0\,, \tag{20.83}$$

is satisfied. Applied to the divergent part it yields

$$-p_{1\mu}p_{2\rho}\epsilon_{\lambda\mu\nu\rho} = 0\,,$$

a condition that is not satisfied. Therefore, the sum of the two diagrams is finite. Different regularizations may still differ by finite quantities of the form (20.82) but again all regularizations consistent with vector gauge invariance must give the same answer.

Therefore, there are two possibilities:

(i) The divergence $k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$ in a regularization respecting the symmetry between the three arguments vanishes. Then $\Gamma^{(3)}$ is gauge invariant and the axial current is conserved.

(ii) The divergence of the symmetric regularization does not vanish. Then it is possible to add to $\Gamma^{(3)}$ a term proportional to (20.82) to restore gauge invariance but this term breaks the symmetry between external momenta: the axial current is not conserved, an anomaly is present.

20.3.2 Divergence in the regularized theory

The calculation can be done using one of the various gauge invariant regularizations, for example Pauli–Villars's regularization or dimensional regularization with γ_5 being defined as in dimension 4 and thus no longer anticommuting with other γ matrices (see Section 9.6.2). Instead we choose a regularization which preserves the symmetry between the three external arguments and global chiral symmetry, but breaks gauge invariance, replacing in the fermion propagator:

$$(\not{q})^{-1} \longmapsto (\not{q})^{-1}\rho(\varepsilon q^2),$$

where ε is the regularization parameter ($\varepsilon \to 0$), $\rho(z)$ is a positive differentiable function such that $\rho(0) = 1$, and decreasing at least like $1/z$ for $z \to +\infty$.

Then the compatibility between current conservation and gauge invariance implies that $k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$ vanishes.

It is convenient to consider directly the contribution $C^{(2)}(k)$ of order A^2 to $\langle k_\lambda J^5_\lambda(k)\rangle$ which sums the two diagrams:

$$\begin{aligned} C^{(2)}(k) = e^2 \int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\, A_\mu(p_1)A_\nu(p_2) \int \frac{\mathrm{d}^4q}{(2\pi)^4}\rho\big(\varepsilon(q+k)^2\big)\rho\big(\varepsilon(q-p_2)^2\big)\rho(\varepsilon q^2) \\ \times \operatorname{tr}\left[\gamma_5 \not{k}(\not{q}+\not{k})^{-1}\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right], \end{aligned} \tag{20.84}$$

because the calculation then suggests how the method generalizes to arbitrary even dimensions. The calculation relies on the cyclic property of the trace and the *anticommutation* of γ_5.

We transform the expression, using the identity

$$(\not{q})^{-1}\not{k}(\not{q}+\not{k})^{-1} = (\not{q})^{-1} - (\not{q}+\not{k})^{-1}, \tag{20.85}$$

and obtain

$$\begin{aligned} C^{(2)}(k) = e^2 \int \mathrm{d}^4p_1\,\mathrm{d}^4p_2\, A_\mu(p_1)A_\nu(p_2t) \int \frac{\mathrm{d}^4q}{(2\pi)^4}\rho\big(\varepsilon(q+k)^2\big)\rho\big(\varepsilon(q-p_2)^2\big) \\ \times \rho(\varepsilon q^2) \operatorname{tr}\left\{\gamma_5\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu\left[\not{q}^{-1} - (\not{q}+\not{k})^{-1}\right]\right\}. \end{aligned} \tag{20.86}$$

We separate the two contributions in the r.h.s. In the second contribution, proportional to $(\not{q}+\not{k})^{-1}$ we interchange $(p_1,\,\mu)$ and $(p_2,\,\nu)$ and shift $q \mapsto q+p_1$. Combining again the two contributions, we find,

$$C^{(2)}(k) = e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr) \\ \times \operatorname{tr}\left[\gamma_5\gamma_\mu(\not{q}-\not{p}_2)^{-1}\gamma_\nu \not{q}^{-1}\right]\left[\rho\bigl(\varepsilon(q+k)^2\bigr)-\rho\bigl(\varepsilon(q+p_1)^2\bigr)\right]. \tag{20.87}$$

We see that the two terms would cancel in the absence of regulators. This corresponds to the formal proof of current conservation. However, without regularization the integrals diverge and previous manipulations are not legitimate.

Here, instead, we find a non-vanishing sum because the regulating factors which are different. After evaluation of the trace (notice our convention ($A8.23$) for γ_5) the sum becomes

$$C^{(2)}(k) = 4e^2 \int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, A_\mu(p_1) A_\nu(p_2) \int \frac{\mathrm{d}^4 q}{(2\pi)^4} \rho\bigl(\varepsilon(q-p_2)^2\bigr)\rho\bigl(\varepsilon q^2\bigr) \\ \times \epsilon_{\mu\nu\rho\sigma}\frac{p_{2\rho} q_\sigma}{q^2(q-p_2)^2}\left[\rho\bigl(\varepsilon(q+p_1)^2\bigr)-\rho\bigl(\varepsilon(q+k)^2\bigr)\right]. \tag{20.88}$$

Contributions coming from finite values of q cancel in the $\varepsilon \to 0$ limit. Due to the cut-off, the relevant values of q are of order $\varepsilon^{-1/2}$. We can, therefore, simplify the q integrand:

$$\int \frac{\mathrm{d}^4 q}{(2\pi)^4 q^4} p_{2\rho} q_\sigma \rho^2(\varepsilon q^2)\rho'(\varepsilon q^2)\left[2\varepsilon q_\lambda (p_1-k)_\lambda\right]. \tag{20.89}$$

The identity:

$$\int \mathrm{d}^4 q\, q_\alpha q_\beta f(q^2) = \tfrac{1}{4}\delta_{\alpha\beta}\int \mathrm{d}^4 q\, q^2 f(q^2),$$

transforms the integral into

$$\tfrac{1}{2} p_{2\rho}(2p_1+p_2)_\sigma \int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2}\rho^2(\varepsilon q^2)\rho'(\varepsilon q^2). \tag{20.90}$$

The remaining integral can be calculated explicitly (we recall $\rho(0)=1$)

$$\int \frac{\varepsilon \mathrm{d}^4 q}{(2\pi)^4 q^2}\rho^2(\varepsilon q^2)\rho'(\varepsilon q^2) = \frac{1}{8\pi^2}\int_0^\infty \varepsilon q \mathrm{d}q\, \rho^2(\varepsilon q^2)\rho'(\varepsilon q^2) = -\frac{1}{48\pi^2},$$

and yields a result independent of the function ρ. We finally obtain

$$\left\langle k_\lambda J^5_\lambda(k)\right\rangle = -\frac{e^2}{12\pi^2}\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4 p_1\, \mathrm{d}^4 p_2\, p_{1\mu} A_\nu(p_1) p_{2\rho} A_\sigma(p_2). \tag{20.91}$$

From the definition (20.80) we conclude

$$k_\lambda \Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2}\epsilon_{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma}\,. \tag{20.92}$$

This non-vanishing result implies that any definition of the determinant $\det \not{D}$ breaks at least either current conservation or gauge invariance. Since gauge invariance is essential to the construction of QED we choose to break current conservation. Exchanging arguments, we obtain the value of $p_{1\mu}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)$:

$$p_{1\mu}\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{6\pi^2}\epsilon_{\lambda\nu\rho\sigma}k_\rho p_{2\sigma} \,. \tag{20.93}$$

If instead we had used a gauge invariant regularization, the result for $\Gamma^{(3)}$ would have differed by a term $\delta\Gamma^{(3)}$ proportional to (20.82):

$$\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = K\epsilon_{\lambda\mu\nu\rho}(p_1-p_2)_\rho \,. \tag{20.94}$$

The constant K then is determined by the condition of gauge invariance

$$p_{1\mu}\left[\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) + \delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2)\right] = 0 \,,$$

which yields

$$p_{1\mu}\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = -\frac{e^2}{6\pi^2}\epsilon_{\lambda\nu\rho\sigma}k_\rho p_{2\sigma} \;\Rightarrow\; K = e^2/(6\pi^2). \tag{20.95}$$

This gives an additional contribution to the divergence of the current

$$k_\lambda\delta\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \frac{e^2}{3\pi^2}\epsilon_{\mu\lambda\rho\sigma}p_{1\rho}p_{2\sigma} \,. \tag{20.96}$$

Therefore, in a QED-like gauge invariant field theory with massless fermions the axial current is not conserved: this is called the chiral anomaly. For any gauge invariant regularization one finds

$$k_\lambda\Gamma^{(3)}_{\lambda\mu\nu}(k;p_1,p_2) = \left(\frac{e^2}{2\pi^2} \equiv \frac{2\alpha}{\pi}\right)\epsilon_{\mu\nu\rho\sigma}p_{1\rho}p_{2\sigma} \,. \tag{20.97}$$

Equation (20.97) can be rewritten, after Fourier transformation, as a non-conservation equation for the axial current:

$$\partial_\lambda J^5_\lambda(x) = -i\frac{\alpha}{4\pi}\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} \,. \tag{20.98}$$

Since global chiral symmetry is not broken, the integral over the whole space of the anomalous term must vanish. This condition is indeed verified since the anomaly can immediately be written as a total derivative:

$$\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = 4\partial_\mu\left(\epsilon_{\mu\nu\rho\sigma}A_\nu\partial_\rho A_\sigma\right). \tag{20.99}$$

The space integral of the anomalous term depends only on the behaviour of the gauge field at boundaries, and this property indicates a relation between topology and anomalies.

Equation (20.98) also implies

$$\ln\det\left[\mathrm{e}^{i\gamma_5\theta(x)}\not{D}\,\mathrm{e}^{i\gamma_5\theta(x)}\right] = \ln\det\not{D} - i\frac{\alpha}{4\pi}\int \mathrm{d}^4x\,\theta(x)\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}(x)F_{\rho\sigma}(x) + O(\theta^2). \tag{20.100}$$

In the form we have presented it the calculation generalizes without difficulty to general even dimensions $2n$. Note simply that the permutation $(\mathbf{p}_1, \mu) \leftrightarrow (\mathbf{p}_2, \nu)$ in the second term of equation (20.86) is replaced by a cyclic permutation. The anomaly in the divergence of the axial current $J^S_\lambda(x)$ in general is

$$\partial_\lambda J^S_\lambda(x) = -2i \frac{e^n}{(4\pi)^n n!} \epsilon_{\mu_1\nu_1\ldots\mu_n\nu_n} F_{\mu_1\nu_1} \ldots F_{\mu_n\nu_n} , \tag{20.101}$$

where $\epsilon_{\mu_1\nu_1\ldots\mu_n\nu_n}$ is the completely antisymmetric tensor.

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 which would have both an abelian gauge vector and axial symmetry, where the action in the fermion sector would read

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^4 x \, \bar{\psi}(x) \left(\not{\partial} + ie\not{A} + i\gamma_5 \not{B}\right) \psi(x). \tag{20.102}$$

Current conservation is a WT identity in the gauge theory 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. They, therefore, also yield an anomaly which cannot be removed since the correlation function by definition is symmetric in its three arguments. This prevents the construction even of a theory with a purely axial gauge symmetry ($e = 0$).

A way to solve both problems is to cancel the anomaly by introducing another fermion of opposite chiral coupling. With more fermions other coupling combinations are possible. Note, however, that 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.

20.3.3 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 gauge group G and $\mathbf{A}_\mu$ is the corresponding gauge field. The action is

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^4 x \, \bar{\psi}(x) \not{\mathbf{D}} \, \psi(x) \tag{20.103}$$

with

$$\not{\mathbf{D}} = \not{\partial} + \not{\mathbf{A}} . \tag{20.104}$$

The axial current

$$J^5_\mu(x) = i\bar{\psi}(x)\gamma_5\gamma_\mu\psi(x),$$

is still gauge invariant. Therefore, no new calculation is needed; the result is completely determined by dimensional analysis, gauge invariance and the previous calculation which yields the term of order $\mathbf{A}^2$:

$$\partial_\lambda J^5_\lambda(x) = -\frac{i}{16\pi^2} \epsilon_{\mu\nu\rho\sigma} \operatorname{tr} \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma} , \tag{20.105}$$

in which $\mathbf{F}_{\mu\nu}$ is now the corresponding curvature. Again this expression must be a total derivative. One verifies:

$$\epsilon_{\mu\nu\rho\sigma} \operatorname{tr} \mathbf{F}_{\mu\nu} \mathbf{F}_{\rho\sigma} = 4 \, \epsilon_{\mu\nu\rho\sigma} \partial_\mu \operatorname{tr} \left(\mathbf{A}_\nu \partial_\rho \mathbf{A}_\sigma + \tfrac{2}{3} \mathbf{A}_\nu \mathbf{A}_\rho \mathbf{A}_\sigma\right) . \tag{20.106}$$

20.3.4 Anomaly and eigenvalues of the Dirac operator

We assume that the spectrum of $\not{D}$, the Dirac operator in a non-abelian gauge field (equation (20.104)), is discrete (putting temporarily the fermions in a box if necessary) and call d_n and $\varphi_n(x)$ the corresponding eigenvalues and eigenvectors:

$$\not{D}\varphi_n = d_n\varphi_n \,. \tag{20.107}$$

The eigenvalues are gauge invariant, because in a gauge transformation of unitary matrix $\mathbf{g}(x)$ the Dirac operator becomes

$$\not{D} \mapsto \mathbf{g}^{-1}(x)\not{D}\mathbf{g}(x) \;\Rightarrow\; \varphi_n(x) \mapsto \mathbf{g}(x)\varphi_n(x).$$

For a unitary or orthogonal group the massless Dirac operator is antihermitian; therefore, the eigenvalues are imaginary and the eigenvectors orthogonal. In addition we choose them with unit norm.

The anticommutation $\not{D}\gamma_5 + \gamma_5\not{D} = 0$ implies

$$\not{D}\gamma_5\varphi_n = -d_n\gamma_5\varphi_n \,. \tag{20.108}$$

Therefore, either d_n is different from zero, and $\gamma_5\varphi_n$ is an eigenvector of $\not{D}$ 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 $\not{D}$ can be chosen eigenvectors of definite chirality, that is, eigenvectors of γ_5 with eigenvalue ± 1,

$$\not{D}\varphi_n = 0\,, \quad \gamma_5\varphi_n = \pm\varphi_n \,.$$

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 $\not{D}+m$ regularized by mode truncation (mode regularization):

$$\det{}_N(\not{D}+m) = \prod_{n\le N}(d_n+m), \tag{20.109}$$

keeping the N lowest eigenvalues of $\not{D}$ (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 $\not{D}$ are gauge invariant. Note that in the truncated space the trace of γ_5 is the index of the Dirac operator:

$$\operatorname{tr}\gamma_5 = n_+ - n_- \,. \tag{20.110}$$

It does not vanish if $n_+ \neq n_-$, a situation which endangers axial current conservation.

In a chiral transformation (20.77) with θ constant, the determinant of $(\not{D}+m)$ becomes

$$\det{}_N(\not{D}+m) \mapsto \det{}_N\left(\mathrm{e}^{i\theta\gamma_5}(\not{D}+m)\,\mathrm{e}^{i\theta\gamma_5}\right).$$

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 $\not{D}+m$ by $d_n\sigma_3+m$. The determinant in the subspace is then

$$\det\left(\mathrm{e}^{i\theta\sigma_1}(d_n\sigma_3+m)\,\mathrm{e}^{i\theta\sigma_1}\right) = \det\mathrm{e}^{2i\theta\sigma_1}\det(d_n\sigma_3+m) = m^2-d_n^2,$$

because σ_1 is traceless.

In the eigenspace of vanishing eigenvalue $d_n = 0$ with positive chirality, of dimension n_+, γ_5 is diagonal with eigenvalue 1 and thus

$$m^{n_+} \mapsto m^{n_+} \,\mathrm{e}^{2i\theta n_+} .$$

Similarly, in the eigenspace $d_n = 0$ of chirality -1

$$m^{n_-} \mapsto m^{n_-} \,\mathrm{e}^{-2i\theta n_-} .$$

We conclude

$$\det{}_N \left(\mathrm{e}^{i\theta\gamma_5}(\not{D} + m)\,\mathrm{e}^{i\theta\gamma_5}\right) = \mathrm{e}^{2i\theta(n_+ - n_-)} \det{}_N(\not{D} + m).$$

The ratio of both determinants is independent of N. Taking the limit $N \to \infty$, we find

$$\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D} + m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D} + m)^{-1}\right] = \mathrm{e}^{2i\theta(n_+ - n_-)} . \tag{20.111}$$

Note that the l.h.s. of equation (20.111) is obviously 1 when $\theta = n\pi$, which implies that the coefficient of 2θ in the r.h.s. must indeed be an integer.

The variation of $\ln\det(\not{D} + m)$,

$$\ln\det\left[\left(\mathrm{e}^{i\gamma_5\theta}(\not{D} + m)\,\mathrm{e}^{i\gamma_5\theta}\right)(\not{D} + m)^{-1}\right] = 2i\theta(n_+ - n_-),$$

at first order in θ is related to the variation of the action (20.75) and thus to the expectation value of the integral of the divergence of the axial current $\left\langle \int \mathrm{d}^4x \partial_\mu J^5_\mu(x)\right\rangle$. In the limit $m = 0$ it is thus related to the space integral of the chiral anomaly (20.105):

$$-\frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\int \mathrm{d}^4x \;\mathrm{tr}\,\mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma} = n_+ - n_- . \tag{20.112}$$

Concerning this result several comments can be made:

(i) At first order in θ in the absence of regularization we have calculated ($\ln\det = \mathrm{tr}\ln$)

$$\ln\det\left[1 + i\theta\left(\gamma_5 + (\not{D} + m)\gamma_5(\not{D} + m)^{-1}\right)\right] \sim 2i\theta\,\mathrm{tr}\,\gamma_5 ,$$

where we have used the cyclic property of the trace. Since the trace of the matrix γ_5 vanishes we could expect naively a vanishing result. But trace here means trace in γ space and in coordinate space, and γ_5 really stands here for $\gamma_5\delta(x - y)$. The mode regularization yields a well-defined finite result for the ill-defined product $0 \times \delta^d(0)$.

(ii) The property that the integral (20.112) is quantized shows that the form of the anomaly is related to topological properties of the gauge field 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 (20.106) confirms.

(iii) Gauge field configurations exist for which the r.h.s. of equation (20.112) does not vanish, for example, instantons as we show in Section 41.6. We have shown above 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 20.5 and 41.6.

(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 r.h.s. of equation (20.111) we have been forced to introduce a mass to find 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 zero because the trace of γ_5 does not vanish in the eigenspace (equation (20.110)). In the limit $m \to 0$ the complete determinant vanishes but not the ratio of determinants for different values of θ because the powers of m cancel.

20.4 Non-Abelian Anomaly

We first consider the problem of conservation of a general axial current in a non-abelian vector gauge theory, and then the issue of obstruction to gauge invariance in chiral gauge theories.

20.4.1 General axial current

We now discuss the problem of the conservation of a general axial current in the example of a fermion action which 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 call t^α the generators of G. The current then has the form:

$$J^{5\alpha}_\mu(x) = -\bar{\psi}\gamma_5\gamma_\mu t^\alpha \psi \,. \tag{20.113}$$

When the gauge group is connected with the chiral group, the current conservation equation involves the gauge covariant derivative ($\mathbf{D}_\mu = \partial_\mu + [A_\mu, \bullet]$):

$$\mathbf{D}_\mu \mathbf{J}^5_\mu = 0 \,. \tag{20.114}$$

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 geometrical factor. Then the complete form of the anomaly is dictated by gauge covariance. One finds:

$$\mathbf{D}_\lambda J^{5\alpha}_\lambda(x) = -\frac{i}{16\pi^2}\epsilon_{\mu\nu\rho\sigma} \operatorname{tr} t^\alpha \mathbf{F}_{\mu\nu}\mathbf{F}_{\rho\sigma} \,. \tag{20.115}$$

In particular, if the gauge group is disconnected from the chiral group the anomaly is proportional to $\operatorname{tr} t^\alpha$ and, therefore, only different from zero for the abelian factors of G.

20.4.2 Obstruction to gauge invariance

We now want to consider a non-abelian gauge field coupled to left or right-handed fermions:

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^4 x \, \bar{\psi}(x) \tfrac{1}{2}(1 + \gamma_5) \not{D} \psi(x), \tag{20.116}$$

($\frac{1}{2}(1-\gamma_5)$ is treated in the same way). We can construct a consistent gauge theory only if the partition function

$$\mathcal{Z}(\mathbf{A}_\mu) = \int [\mathrm{d}\psi \mathrm{d}\bar{\psi}] \exp\left[-\mathcal{S}(\psi,\bar{\psi})\right] \tag{20.117}$$

is gauge invariant.

If we introduce the generators t^α of the gauge group in the fermion representation, we can write the corresponding current $\mathbf{J}_\mu$ as

$$J^\alpha_\mu(x) = -\bar{\psi}\tfrac{1}{2}(1+\gamma_5)\gamma_\mu t^\alpha \psi \,. \tag{20.118}$$

The invariance of $\mathcal{Z}(\mathbf{A}_\mu)$ under an infinitesimal gauge transformation again leads to a covariant conservation equation for the current:

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu \rangle = 0 \,.$$

The calculation of the term of degree two in the gauge field of the anomaly is straightforward: the regularization adopted for the calculation in Section 20.3.2 is also suited to the present case since the current-gauge field three-point function is symmetric in the external arguments. The group structure yields a simple geometrical factor. The global factor can be taken from the abelian calculation. It differs from the result (20.91) 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:

$$(\mathbf{D}_\mu \mathbf{J}_\mu(x))^\alpha = -\frac{i}{24\pi^2}\partial_\mu \epsilon_{\mu\nu\rho\sigma} \operatorname{tr}\left[t^\alpha \left(\mathbf{A}_\nu \partial_\rho \mathbf{A}_\sigma + \tfrac{1}{2}\mathbf{A}_\nu \mathbf{A}_\rho \mathbf{A}_\sigma\right)\right] . \tag{20.119}$$

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

$$d_{\alpha\beta\gamma} = \tfrac{1}{2}\operatorname{tr}\left[t^\alpha \left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right] . \tag{20.120}$$

For a unitary representation the generators t^α are, with our conventions, antihermitian. Therefore, the coefficients $d_{\alpha\beta\gamma}$ are purely imaginary:

$$d^*_{\alpha\beta\gamma} = \tfrac{1}{2}\operatorname{tr}\left[t^\alpha \left(t^\beta t^\gamma + t^\gamma t^\beta\right)\right]^* = -d_{\alpha\beta\gamma} \,. \tag{20.121}$$

For all real (the t^α antisymmetric) or "pseudo-real" ($t^\alpha = -S\,{}^T t^\alpha S^{-1}$) representations these coefficients vanish. It follows that the only non-abelian groups which can lead to anomalies in four dimensions are: $SU(N)$ for $N \ge 3$, $SO(6)$ and E_6.

20.4.3 Wess–Zumino consistency conditions

In Section 20.4.2 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 have argued in Section 15.3, this implies compatibility conditions, which are here constraints on the form of the anomaly. One convenient way to express these constraints is to express the nilpotency of BRS transformations (for details see Chapter 21).

In a BRS transformation the variation of the gauge field $\mathbf{A}_\mu$ takes the form (equation (19.50)):

$$\delta \mathbf{A}_\mu(x) = \mathbf{D}_\mu \mathbf{C}(x)\bar{\varepsilon}\,, \tag{20.122}$$

where $\mathbf{C}$ is a (fermion) ghost field and $\bar{\varepsilon}$ an anticommuting constant. The corresponding variation of $\ln \mathcal{Z}(\mathbf{A}_\mu)$ is

$$\delta \ln \mathcal{Z}(\mathbf{A}_\mu) = -\int \mathrm{d}^4 x \; \langle \mathbf{J}_\mu(x)\rangle \, \mathbf{D}_\mu \mathbf{C}(x)\bar{\varepsilon}\,. \tag{20.123}$$

If we write the anomaly equation:

$$\langle \mathbf{D}_\mu \mathbf{J}_\mu(x)\rangle = \mathcal{A}(\mathbf{A}_\mu, x), \tag{20.124}$$

equation (20.123), after an integration by parts, can be rewritten as

$$\delta \ln \mathcal{Z}(\mathbf{A}_\mu) = \int \mathrm{d}^4 x \, \mathcal{A}(\mathbf{A}_\mu, x)\mathbf{C}(x)\bar{\varepsilon}\,. \tag{20.125}$$

Since $\mathcal{A}\mathbf{C}$ is a BRS variation it is invariant under BRS transformation, where the gauge field transforms according to equation (19.50,20.122) and the fermion ghost $\mathbf{C}(x)$ as (equation (19.48)):

$$\delta \mathbf{C}(x) = \bar{\varepsilon}\, \mathbf{C}^2(x). \tag{20.126}$$

By expressing that $\mathcal{A}C$ is BRS invariant one obtains a constraint on the possible form of anomalies. One verifies that this condition determines the term cubic in $\mathbf{A}$ in the r.h.s. of equation (20.119).

20.5 Physical Applications

Weak–electromagnetic interactions. The condition of anomaly cancellation discussed in Section 20.4.2 constrains the model of w.e.m. interactions. In the Standard Model, for example, the anomalous contributions of leptons cancels the quark contributions. This cancellation occurs within each generation, as we now show, provided that for each flavour quarks exist in three states. In the w.e.m. 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. Actually one discovers that both conditions are equivalent. If we consider two $SU(2)$ and one $U(1)$ gauge fields, only $SU(2)$ doublets contribute and equation (20.119) leads to the condition:

$$\sum_{\text{all doublets}} Y_{\mathrm{L}} \operatorname{tr} \tau^\alpha \tau^\beta = 0\,,$$

in which Y_{L} is the $U(1)$ charge (see Section 20.1). This condition reduces to

$$\sum_{\text{all doublets}} Y_{\mathrm{L}} = 0\,. \tag{20.127}$$

The vertex with three $U(1)$ gauge fields yields the condition:

$$\sum_{\text{left-handed parts}} Y_{\mathrm{L}}^3 - \sum_{\text{right-handed parts}} Y_{\mathrm{R}}^3 = 0\,,$$

because the contributions to the anomaly of right-handed and left-handed couplings have opposite signs. In the Standard Model the left and right charges are related. Summing the charges of one doublet and the corresponding singlets, we obtain

$$\sum_{\text{all doublets}} (Y_{\mathrm{L}}+1)^3 + (Y_{\mathrm{L}}-1)^3 - 2Y_{\mathrm{L}}^3 = 0\,,$$

a condition which reduces to equation (20.127).

In one generation the lepton doublet has $Y_{\mathrm{L}} = -1$ and the quark $Y_{\mathrm{L}} = 1/3$. Therefore, a cancellation requires that the quarks exist in three states. These states are provided by the colour quantum number.

Electromagnetic π_0 decay. In an effective low energy theory for Strong Interactions based on a linearly broken $SU(2) \times SU(2)$ symmetry, where hadrons are considered as elementary fields, the non-conservation of the axial current $\mathbf{J}^5_\mu$ is at leading order expressed by the equation:

$$\partial_\mu \mathbf{J}^5_\mu = m_\pi^2 f_\pi \boldsymbol{\pi}\,. \tag{20.128}$$

We concentrate here on the third component $[J^5_\mu]_3$ of the current which corresponds in the r.h.s. to the neutral pion π_0 field. After introduction of electromagnetic interactions in the model, the relation between the divergence of the axial current and the π_0 field allows to calculate the electromagnetic decay rate of the neutral pion when the four-momentum $\mathbf{k}$ of the pion goes to zero. In the absence of anomalies, the expectation value of relation (20.128) multiplied by two photon fields implies that the decay rate vanishes for $\mathbf{k} = 0$ in contradiction with reasonable smoothness assumptions and experimental results. Taking instead into account anomaly equation (20.98) one finds:

$$\partial_\mu [J^5_\mu]_3 = m_\pi^2 f_\pi \pi_0 - i\frac{\alpha}{8\pi}\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\,. \tag{20.129}$$

Multiplying the equation by two photon fields, taking the expectation value and going to the limit $\mathbf{k} = 0$ to eliminate the l.h.s., one now obtains a non-vanishing decay amplitude for an unphysical π_0 at zero total momentum. In the σ-model at leading order one can extrapolate to $k^2 = -m_\pi^2$. The result is in excellent agreement with experiment. The theoretical rate Γ is given by

$$\Gamma = \frac{\alpha^2 m_\pi^3}{64\pi^3 f_\pi^2} = 7.6\ \mathrm{eV}\,,$$

while $\Gamma^{\exp} = (7.37 \pm 1.5)$ eV. A similar estimate was first derived by Steinberger from direct Feynman graph calculation, before the relation to anomalies had been discovered.

Note that a similar theoretical estimate is obtained in the quark model with massless quarks, for three colours.

The solution of the $U(1)$ problem. In a theory of Strong Interactions in which the quarks are massless and interact through a colour gauge group, the action has a chiral $U(N_{\mathrm{F}}) \times U(N_{\mathrm{F}})$ symmetry, in which N_{F} is the number of flavours. The spontaneous breaking of the chiral group to its diagonal subgroup $U(N_{\mathrm{F}})$ 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. Of course 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, non-abelian gauge theories admit instanton solutions which give a periodic structure to the vacuum (for details see Section 41.6). 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 (20.112) to the space integral of the anomaly (20.105). This indicates (but no 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 resolves a long standing puzzle since experimentally no corresponding light pseudoscalar boson is found for $N_F = 2, 3$.

Bibliographical Notes

The Standard Model was proposed in its present form by

S. Weinberg *Phys. Rev. Lett.* 19 (1967) 1264; A. Salam, *Elementary Particle Theory*, N. Svartholm ed. (Almqvist and Wiksell, Stockholm 1968).

Earlier attempts include

S.L. Glashow, *Nucl. Phys.* 22 (1961) 579.

To explain the suppression of strangeness changing neutral currents, a fourth quark (**c** for charm) was conjectured in

S.L. Glashow, J. Iliopoulos and L. Maiani, *Phys. Rev.* D2 (1970) 1285.

A selection of papers is collected in

Gauge Theory of Weak and Electromagnetic Interactions, C.H. Lai ed. (World Scientific, Singapore 1981).

Most references given in Chapter 19 are still relevant here. See also

E.S. Abers and B.W. Lee, *Phys. Rep.* 9C (1973) 1; J.C. Taylor, *Gauge Theories of Weak Interactions* (Cambridge University Press, Cambridge 1976).

For the KM matrix see

M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* 49 (1973) 652.

The problem of anomalies is discussed in

J.S. Bell and R. Jackiw, *Nuovo Cimento* A60 (1969) 47; S.L. Adler, *Phys. Rev.* 177 (1969) 2426; D.J. Gross and R. Jackiw, *Phys. Rev.* D6 (1972) 477; H. Georgi and S.L. Glashow, *Phys. Rev.* D6 (1972) 429; C. Bouchiat, J. Iliopoulos and Ph. Meyer, *Phys. Lett.* 38B (1972) 519; E. Witten, *Phys. Lett.* 117B (1982) 324.

It is related to the problem of regularizing the fermion measure in

K. Fujikawa, *Phys. Rev.* D21 (1980) 2848; D22 (1980) 1499(E).

See also the lectures

M. Peskin, *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, R. Stora and J.-B. Zuber eds. (North-Holland, Amsterdam 1984),

and the volume

S.B. Treiman, R. Jackiw, B. Zumino and E. Witten, *Current Algebra and Anomalies* (World Scientific, Singapore 1985) and references therein.

21 GAUGE THEORIES: MASTER EQUATION AND RENORMALIZATION

To discuss the renormalization of gauge theories in the non-abelian case in its full generality, it is necessary to use a rather abstract formalism, which allows one to understand the algebraic structure of the renormalization procedure without being overwhelmed by the notational complexity. There is, however, a price to pay: the translation of the general identities which then appear into usual and more concrete notation becomes a non-trivial exercise.

Our strategy is the following. We first quantize the theory in the temporal gauge. Using a simple identity, we show the equivalence with a quantization in a general class of gauges. This identity automatically implies a BRS symmetry, and, therefore, a set of WT identities for correlation functions. We show that WT identities are also direct consequences of a quadratic *master equation* satisfied the quantized action, equation in which gauge and BRS symmetries are no longer explicit. We show that in the case of renormalizable gauges the master equation is stable under renormalization. We solve it to determine the structure of the renormalized action. We verify that the master equation encodes in a subtle way the gauge properties of the quantized action.

Physical observables in the theory should be independent of the dynamics of the gauge group degrees of freedom. Following the analysis of previous chapters, we first prove that the vacuum amplitude does not depend on this dynamics, that is, is gauge-independent. We argue that this property remains true if we add to the original action sources for gauge invariant operators of canonical dimension low enough, so that the total action remains renormalizable. As a consequence, correlation functions of these gauge invariant operators are independent of the gauge fixing procedure, and have, therefore, a physical meaning. A similar property holds for all gauge invariant operators but the discussion is more involved and will not be given. Finally, when a S-matrix can be defined, S-matrix elements also are gauge-independent.

21.1 Notation and Geometric Structure

Notation. In what follows, we restrict ourselves to boson fields, the generalization to fermions being straightforward. All gauge fields and other scalar boson fields are combined 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:

$$A^i \equiv \left\{A^a_\mu(x), \varphi_b(x)\right\}.$$

A gauge transformation corresponding to an orthogonal representation of a compact Lie group G is written as

$$\delta A^i = D^i_\alpha(A)\omega^\alpha\,, \tag{21.1}$$

in which ω^α are the infinitesimal parameters of the group transformation. As in the example of equation (21.1), we reserve greek indices for the adjoint representation of the Lie algebra. Summation over repeated indices is always meant. It includes integration over space variables. The index α also includes group indices and space variables. In more concrete form, equation (21.1) stands for

$$\begin{aligned}\delta A^a_\mu(x) &= \int \mathrm{d}y\, \left[\partial^x_\mu \delta(x-y)\delta^a_b + f^a_{bc}A^c_\mu(x)\delta(x-y)\right]\omega^b(y),\\ \delta\varphi^{a'}(x) &= \int \mathrm{d}y\, \delta(x-y)[t_{b'}]^{a'}_{c'}\varphi^{c'}(x)\omega^{b'}(y),\end{aligned} \tag{21.2}$$

where the matrices $t_{b'}$ form a (in general reducible) representation of the Lie algebra of G and f^a_{bc} are the corresponding structure constants.

Finally, $\mathcal{S}(A)$ is the gauge invariant action and thus satisfies

$$\frac{\delta S}{\delta A^i} D^i_\alpha(A) = 0\,. \tag{21.3}$$

A group identity. We have explained in Section 15.3 that, as a consequence of the group structure, quite generally the functionals $D^i_\alpha(A)$ satisfy equations which can be regarded as compatibility conditions for the system (21.3) considered as a set of differential equations for $\mathcal{S}(A)$:

$$\frac{\delta D^i_\alpha}{\delta A^j} D^j_\beta - \frac{\delta D^i_\beta}{\delta A^j} D^j_\alpha = f^\gamma_{\alpha\beta} D^i_\gamma\,. \tag{21.4}$$

These equations are formally identical to equations (15.43) of Section 15.3 and (16.41) of Section 16.4.

Since here the functional $D^i_\alpha(A)$ is only affine, it is not difficult to verify the identities (21.4) by a direct calculation, using the commutation relations of the Lie algebra of G. Note, however, that because $D^i_\alpha(A)$ is a differential operator, it is necessary to carefully keep track of the δ-functions. In particular, since the indices α and i include space coordinates, the structure constants $f^\gamma_{\alpha\beta}$ which are proportional to the numerical structure constants f^a_{bc} of the Lie algebra, have a non-trivial dependence on space variables:

$$f^\gamma_{\alpha\beta} \equiv f^c_{ab}\delta(x-y)\delta(y-z). \tag{21.5}$$

As a consequence, it is actually convenient to rewrite the identity (21.4) with the help of two functions $\omega^\alpha_1 \equiv \omega^a_1(x)$ and $\omega^\alpha_2 \equiv \omega^a_2(x)$ and the operator

$$\Delta(\omega) = \omega^\alpha D^i_\alpha(A)\frac{\delta}{\delta A^i}\,. \tag{21.6}$$

The identity (21.4) takes the form

$$[\Delta(\omega_1), \Delta(\omega_2)] = \Delta(\omega_{12}), \quad \omega^\gamma_{12} = f^\gamma_{\alpha\beta}\omega^\alpha_1\omega^\beta_2\,. \tag{21.7}$$

Let us for instance indicate here how this identity can be recovered for the gauge field part. We write $\Delta(\omega)$ in explicit notation:

$$\Delta(\omega) = \int \mathrm{d}x\, \left[\partial_\mu\omega^a(x) + f^a_{bc}\omega^b(x)A^c_\mu(x)\right]\frac{\delta}{\delta A^a_\mu(x)}\,.$$

We then calculate the commutator explicitly

$$[\Delta(\omega_1), \Delta(\omega_2)] = \int \mathrm{d}x\, \left[\partial_\mu\omega^a_1(x) + f^a_{bc}\omega^b(x)A^c_\mu(x)\right]\omega^e_2(x) f^d_{ea}\frac{\delta}{\delta A^d_\mu(x)} - \{1 \leftrightarrow 2\}\,.$$

Using the antisymmetry of f^a_{bc} and the Jacobi identity one indeed obtains equation (21.7) in explicit form, with $\omega^a_{12}(x) = f^a_{cb}\omega^b_1(x)\omega^c_2(x)$.

21.2 Quantization

In Chapter 19, we have first quantized in the temporal gauge and then shown how to pass from the temporal gauge to a covariant gauge. This is a strategy we now generalize and thus begin with a gauge invariant action $\mathcal{S}(A)$ quantized in the temporal gauge.

We thus introduce an equation for the space-dependent group elements g:

$$E_\alpha(g) \equiv F_\alpha(A^g) - \nu_\alpha = 0\,, \tag{21.8}$$

where A^g is the gauge transform of A by g, 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 g (at least for A small so that the equation can be solved perturbatively).

We then calculate the variation of the equation in an infinitesimal gauge transformation of parameters ω^α:

$$\delta E_\alpha(g) = M_{\alpha\beta}(A^g)\omega^\beta, \tag{21.9}$$

$$M_{\alpha\beta}(A) = \frac{\delta F_\alpha(A)}{\delta A^i} D^i_\beta(A). \tag{21.10}$$

From the results derived in Sections 16.2–16.6, we infer the identity

$$1 = \int [\mathrm{d}g\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda] \exp\left[-\mathcal{S}_{\mathrm{gauge}}(A^g, C, \bar{C}, \lambda, \nu)\right] \tag{21.11}$$

with

$$\mathcal{S}_{\mathrm{gauge}}(A, C, \bar{C}, \lambda, \nu) = \lambda^\alpha \bigl(F_\alpha(A) - \nu_\alpha\bigr) - C^\alpha M_{\alpha\beta}(A)\bar{C}^\beta\,. \tag{21.12}$$

The fields $\bar{C}$ and C are spinless fermion fields ("ghost" fields) introduced to represent the determinant of M, and transforming under the adjoint representation of the group G.

We then insert the identity (21.11) into the representation of the partition function in the temporal gauge and change variables $A^g \mapsto A$. This change of variables has the form of a gauge transformation and thus the gauge invariant action $\mathcal{S}(A)$ is unchanged. The dependence of $\mathcal{S}_{\mathrm{gauge}}$ in g disappears, and only the temporal gauge condition remains g-dependent. We have shown in Section 19.3.2 that the integration over g then yields a constant.

In the new gauge, the partition function then has the functional representation

$$\mathcal{Z} = \int [\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda] \exp\left[-\mathcal{S}(A) - \mathcal{S}_{\mathrm{gauge}}(A, C, \bar{C}, \lambda, \nu)\right]. \tag{21.13}$$

Since the result does not depend on the field ν (the noise in the terminology of Section 16.6) we integrate over ν with a measure $\mathrm{d}\rho(\nu)$, 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,

$$\int [\mathrm{d}\rho(\nu)] \exp \lambda^\alpha \nu_\alpha = \exp w(\lambda), \tag{21.14}$$

we finally obtain

$$\mathcal{Z} = \int [\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda] \exp\left[-\mathcal{S}(A,C,\bar{C},\lambda)\right] \tag{21.15}$$

with

$$\mathcal{S}(A,C,\bar{C},\lambda) = \mathcal{S}(A) - w(\lambda) + \lambda^{\alpha} F_{\alpha}(A) - C^{\alpha} M_{\alpha\beta}(A)\bar{C}^{\beta}\,, \quad M_{\alpha\beta} = \frac{\delta F_{\alpha}}{\delta A^{i}} D^{i}_{\beta}\,. \tag{21.16}$$

The interpretation of this 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 functional integral in the continuum, it is necessary to provide one. For the field and group element g, we thus have introduced a stochastic dynamics in the sense of Section 16.6. This dynamics is somewhat arbitrary and we shall have eventually to prove that "physical results" (we shall explain later what we mean by physical results) do not depend on its choice.

21.3 BRS Symmetry

We first recall that in the representation (21.15) the functional integral is invariant under gauge transformations which translate ν_{α} (see Section 16.1):

$$\delta A^{i} = D^{i}_{\alpha}[M^{-1}]^{\alpha\beta}\mu_{\beta}\,, \tag{21.17}$$

and thus

$$\delta\left[F_{\alpha}(A)\right] = \frac{\delta F_{\alpha}}{\delta A^{i}} D^{i}_{\beta}[M^{-1}]^{\beta\gamma}\mu_{\gamma} = \mu_{\alpha}\,. \tag{21.18}$$

This Slavnov–Taylor symmetry has played an essential role in the initial proof of the renormalizability of non-abelian gauge theories. However, the non-local character of this transformation explains the complexity of the proof.

It follows from the general analysis of Chapter 16 that the final action (21.16) has a BRS symmetry. Because the field that has a stochastic dynamics belongs to the gauge group, the BRS transformation for the field A^{i} has the form of a gauge transformation (equation (17.39)). We have verified in Chapter 19 that, since we have provided a dynamics to a group element, the BRS symmetry is similar in form to the symmetry of the dynamic action of chiral models as described in Sections 16.4,17.3 (equations (16.46,16.47) and (17.37–17.39)). Also, in these more abstract notations, the basic equation (21.4) is formally identical to the commutation relations in the case of non-linear representation of groups of Section 15.3, and to the relations (16.41) of Section 16.4. The corresponding BRS transformations for the A and C fields are thus given by equations (16.39). As we know from the general analysis of Chapter 16, the transformations of λ and C are independent of the dynamics. We conclude that the BRS transformations have the form

$$\delta A^{i} = \varepsilon D^{i}_{\alpha}(A)\bar{C}^{\alpha}\,, \quad \delta\bar{C}^{\alpha} = -\tfrac{1}{2}\varepsilon f^{\alpha}_{\beta\gamma}\bar{C}^{\beta}\bar{C}^{\gamma}\,, \tag{21.19a}$$

$$\delta C^{\alpha} = \varepsilon\lambda^{\alpha}\,, \qquad \delta\lambda^{\alpha} = 0\,. \tag{21.19b}$$

Using a more explicit notation one verifies that these transformations are identical to the BRS transformations derived in Section 19.3.2. It is convenient to set

$$D^{i} = D^{i}_{\alpha}(A)\bar{C}^{\alpha}\,. \tag{21.20}$$

we finally obtain

$$\mathcal{Z} = \int [\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda] \exp\left[-\mathcal{S}(A, C, \bar{C}, \lambda)\right] \tag{21.15}$$

with

$$\mathcal{S}(A, C, \bar{C}, \lambda) = \mathcal{S}(A) - w(\lambda) + \lambda^\alpha F_\alpha(A) - C^\alpha M_{\alpha\beta}(A)\bar{C}^\beta\,, \quad M_{\alpha\beta} = \frac{\delta F_\alpha}{\delta A^i} D^i_\beta\,. \tag{21.16}$$

The interpretation of this 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 functional integral in the continuum, it is necessary to provide one. For the field and group element g, we thus have introduced a stochastic dynamics in the sense of Section 16.6. This dynamics is somewhat arbitrary and we shall have eventually to prove that "physical results" (we shall explain later what we mean by physical results) do not depend on its choice.

21.3 BRS Symmetry

We first recall that in the representation (21.15) the functional integral is invariant under gauge transformations which translate ν_α (see Section 16.1):

$$\delta A^i = D^i_\alpha [M^{-1}]^{\alpha\beta} \mu_\beta\,, \tag{21.17}$$

and thus

$$\delta\left[F_\alpha(A)\right] = \frac{\delta F_\alpha}{\delta A^i} D^i_\beta [M^{-1}]^{\beta\gamma} \mu_\gamma = \mu_\alpha\,. \tag{21.18}$$

This Slavnov–Taylor symmetry has played an essential role in the initial proof of the renormalizability of non-abelian gauge theories. However, the non-local character of this transformation explains the complexity of the proof.

It follows from the general analysis of Chapter 16 that the final action (21.16) has a BRS symmetry. Because the field that has a stochastic dynamics belongs to the gauge group, the BRS transformation for the field A^i has the form of a gauge transformation (equation (17.39)). We have verified in Chapter 19 that, since we have provided a dynamics to a group element, the BRS symmetry is similar in form to the symmetry of the dynamic action of chiral models as described in Sections 16.4,17.3 (equations (16.46,16.47) and (17.37–17.39)). Also, in these more abstract notations, the basic equation (21.4) is formally identical to the commutation relations in the case of non-linear representation of groups of Section 15.3, and to the relations (16.41) of Section 16.4. The corresponding BRS transformations for the A and C fields are thus given by equations (16.39). As we know from the general analysis of Chapter 16, the transformations of λ and C are independent of the dynamics. We conclude that the BRS transformations have the form

$$\delta A^i = \varepsilon D^i_\alpha(A)\bar{C}^\alpha\,, \quad \delta\bar{C}^\alpha = -\tfrac{1}{2}\varepsilon f^\alpha_{\beta\gamma}\bar{C}^\beta\bar{C}^\gamma\,, \tag{21.19a}$$

$$\delta C^\alpha = \varepsilon\lambda^\alpha\,, \qquad \delta\lambda^\alpha = 0\,. \tag{21.19b}$$

Using a more explicit notation one verifies that these transformations are identical to the BRS transformations derived in Section 19.3.2. It is convenient to set

$$D^i = D^i_\alpha(A)\bar{C}^\alpha\,. \tag{21.20}$$

The BRS operator. To express the BRS symmetry, it is also useful to introduce the anticommuting differential operator

$$\mathcal{D}_0 = D^i \frac{\delta}{\delta A^i} - \frac{1}{2} f^{\alpha}_{\beta\gamma} \bar{C}^{\beta} \bar{C}^{\gamma} \frac{\delta}{\delta C^{\alpha}} + \lambda^{\alpha} \frac{\delta}{C^{\alpha}} . \tag{21.21}$$

The essential property of the BRS operator is that it has a vanishing square:

$$[\mathcal{D}_0]^2 = 0 , \tag{21.22}$$

as we know from general arguments and as one can again verify explicitly here. Hence, apart from contributions which are gauge invariant, BRS symmetric terms include contributions which are BRS exact, that is, of the form $\mathcal{D}_0 \Phi(A, C, \bar{C}, \lambda)$. We verify that the action (21.16) has exactly such a decomposition:

$$\begin{aligned} \mathcal{S} &= \mathcal{S}(A) + \mathcal{D}_0 \Phi , \\ \Phi &= C^{\alpha} \left[F_{\alpha}(A) - \frac{\partial}{\partial \lambda^{\alpha}} \int_0^1 \mathrm{d}s \, w(s\lambda) \right] . \end{aligned} \tag{21.23}$$

This is not surprising because the property has been proven quite generally in Section 16.5 (see equation (16.59)). The quantized action is, therefore, BRS symmetric:

$$\mathcal{D}_0 \mathcal{S}(A, C, \bar{C}, \lambda) = 0 . \tag{21.24}$$

Remark. In all examples we consider here, the function $w(\lambda)$ defined by equation (21.14) is quadratic in λ and thus the corresponding gaussian integral can be performed. After integration the new action is still BRS symmetric, the variation of $\bar{C}^{\alpha}$ now takes the form

$$\delta C^{\alpha} = \varepsilon \tilde{a}^{\alpha\beta} F_{\beta}(A),$$

$\tilde{\mathbf{a}}$ being a constant matrix. The BRS operator is still nilpotent but its square only vanishes when the equation $F(A) = 0$ is satisfied. Therefore, the property (21.22) is not shared by all realizations of BRS transformations, and may require the introduction of additional auxiliary variables. The results which follow can be proven in all formulations.

21.4 WT Identities and Master Equation

Although we are interested only in A field correlation functions, to study the consequence of the BRS symmetry (21.19), it is necessary to introduce sources for all fields and all operators generated by BRS transformations:

$$\begin{aligned} \mathcal{Z}(J, \eta, \bar{\eta}, l, K, L) = \int & [\mathrm{d}A \mathrm{d}\bar{C} \mathrm{d}C \mathrm{d}\lambda] \left[\mathrm{d}\bar{C} \mathrm{d}C \right] \exp \left[-\mathcal{S}(A, \bar{C}, C, \lambda, K, L) \right. \\ & \left. + J_i A^i + \bar{\eta}_{\alpha} \bar{C}^{\alpha} + C^{\alpha} \eta_{\alpha} + l_{\alpha} \lambda^{\alpha} \right] \end{aligned} \tag{21.25}$$

with

$$\begin{aligned} \mathcal{S}(A, \bar{C}, C, \lambda, K, L) = \mathcal{S}(A) - w(\lambda) + \lambda^{\alpha} F_{\alpha}(A) - C^{\alpha} M_{\alpha\beta} \bar{C}^{\beta} \\ - K_i D^i - \tfrac{1}{2} L_{\alpha} f^{\alpha}_{\beta\gamma} \bar{C}^{\beta} \bar{C}^{\gamma} . \end{aligned} \tag{21.26}$$

Under a BRS transformation, only the source terms for A, $\bar{C}$ and C vary:

$$\delta\left[J_i A^i + \bar{\eta}_\alpha \bar{C}^\alpha + C^\alpha \eta_\alpha\right] = \bar{\varepsilon}\left(J_i D^i + \tfrac{1}{2}\bar{\eta}_\alpha f^\alpha_{\beta\gamma}\bar{C}^\beta \bar{C}^\gamma + \lambda^\alpha \eta_\alpha\right). \tag{21.27}$$

This implies an equation for $\mathcal{Z}$:

$$\left(J_i \frac{\delta}{\delta K_i} + \bar{\eta}_\alpha \frac{\delta}{\delta L_\alpha} + \eta_\alpha \frac{\delta}{\delta l_\alpha}\right)\mathcal{Z} = 0\,. \tag{21.28}$$

The generating functional $\mathcal{W}$ of connected correlation functions satisfies the same equation:

$$\left(J_i \frac{\delta}{\delta K_i} + \bar{\eta}_\alpha \frac{\delta}{\delta L_\alpha} + \eta_\alpha \frac{\delta}{\delta l_\alpha}\right)\mathcal{W} = 0\,. \tag{21.29}$$

We then perform a Legendre transformation:

$$\Gamma(A, \bar{C}, C, \lambda, K, L) + W(J, \eta, \bar{\eta}, l, K, L) = J_i A^i + \bar{\eta}_\alpha \bar{C}^\alpha + C^\alpha \eta_\alpha + l_\alpha \lambda^\alpha \tag{21.30}$$

with

$$\begin{aligned} J_i &= \frac{\delta \Gamma}{\delta A^i}, & \bar{\eta}_\alpha &= \frac{\delta \Gamma}{\delta \bar{C}^\alpha}, \\ \eta_\alpha &= \frac{\delta \Gamma}{\delta C^\alpha}, & l_\alpha &= \frac{\delta \Gamma}{\delta \lambda^\alpha}. \end{aligned} \tag{21.31}$$

Equations (21.30,21.31) imply as usual for the sources coupled to composite fields:

$$\frac{\delta \Gamma}{\delta K_i} + \frac{\delta \mathcal{W}}{\delta K_i} = 0\,, \qquad \frac{\delta \Gamma}{\delta L_\alpha} + \frac{\delta \mathcal{W}}{\delta L_\alpha} = 0\,. \tag{21.32}$$

Therefore, the generating functional of proper vertices Γ satisfies

$$\frac{\delta \Gamma}{\delta A^i}\frac{\delta \Gamma}{\delta K_i} + \frac{\delta \Gamma}{\delta \bar{C}^\alpha}\frac{\delta \Gamma}{\delta L_\alpha} - \lambda^\alpha \frac{\delta \Gamma}{\delta C^\alpha} = 0\,. \tag{21.33}$$

As in the case of the non-linear σ-model, because the transformations (21.19) involve composite operators, the WT identity satisfied by Γ takes the form of a quadratic equation. We thus expect that the transformations (21.19) will in general be renormalized.

Master equation. In the classical limit, the 1PI functional Γ reduces to the action. We infer that the action also satisfies equation (21.33), a property that can be easily verified by a direct calculation:

$$\frac{\delta \mathcal{S}}{\delta A^i}\frac{\delta \mathcal{S}}{\delta K_i} + \frac{\delta \mathcal{S}}{\delta \bar{C}^\alpha}\frac{\delta \mathcal{S}}{\delta L_\alpha} - \lambda^\alpha \frac{\delta \mathcal{S}}{\delta C^\alpha} = 0\,. \tag{21.34}$$

This *master equation* is the basic equation for the discussion of the renormalization of quantized gauge theories.

First, it implies without additional assumptions equation (21.33) as we now show. Equation (21.34) can be used in the form

$$\begin{aligned} \int [\mathrm{d}A]\,[\mathrm{d}\lambda]\,[\mathrm{d}\bar{C}\mathrm{d}C] &\left[\frac{\delta \mathcal{S}}{\delta A^i}\frac{\delta \mathcal{S}}{\delta K_i} + \frac{\delta \mathcal{S}}{\delta \bar{C}^\alpha}\frac{\delta \mathcal{S}}{\delta L_\alpha} - \lambda^\alpha \frac{\delta \mathcal{S}}{\delta C^\alpha}\right] \\ &\times \exp\left(-\mathcal{S} + J_i A^i + \bar{\eta}_\alpha \bar{C}^\alpha + C^\alpha \eta_\alpha + l_\alpha \lambda^\alpha\right) = 0\,. \end{aligned} \tag{21.35}$$

We integrate by parts over A^i, $\bar{C}^\alpha$ and C^α using

$$\frac{\delta \mathcal{S}}{\delta A^i}\,\mathrm{e}^{-\mathcal{S}}=-\frac{\delta}{\delta A^i}\,\mathrm{e}^{-\mathcal{S}},\quad \frac{\delta \mathcal{S}}{\delta \bar{C}^\alpha}\,\mathrm{e}^{-\mathcal{S}}=-\frac{\delta}{\delta \bar{C}^\alpha}\,\mathrm{e}^{-\mathcal{S}},\quad \frac{\delta \mathcal{S}}{\delta C^\alpha}\,\mathrm{e}^{-\mathcal{S}}=-\frac{\delta}{\delta C^\alpha}\,\mathrm{e}^{-\mathcal{S}}.$$

Equation (21.35) then becomes

$$\int [\mathrm{d}A]\,[\mathrm{d}\bar{C}\mathrm{d}C]\,[\mathrm{d}\lambda]\left(\frac{\delta^2\mathcal{S}}{\delta A^i\delta K_i}+\frac{\delta^2\mathcal{S}}{\delta\bar{C}^\alpha\delta L_\alpha}+J_i\frac{\delta\mathcal{S}}{\delta K_i}+\bar{\eta}_\alpha\frac{\delta\mathcal{S}}{\delta L_\alpha}-\eta_\alpha\lambda^\alpha\right)$$
$$\times\exp\left(-\mathcal{S}+\text{sources}\right)=0\,. \tag{21.36}$$

In the regularized theory $\delta^2\mathcal{S}/\delta A^i\delta K_i$ and $\delta^2\mathcal{S}/\delta\bar{C}^\alpha\delta L_\alpha$, which are proportional to traces of matrices belonging to an orthogonal representation of the Lie algebra, vanish.

Note, however, that in the case of fermions and chiral gauge transformations the corresponding property is not be necessarily true and this may be the source of anomalies (Section 20.3).

Then, equation (21.36) can be rewritten as

$$\int [\mathrm{d}A]\,[\mathrm{d}\bar{C}\mathrm{d}C]\,[\mathrm{d}\lambda]\left(J_i\frac{\delta}{\delta K_i}+\bar{\eta}_\alpha\frac{\delta}{\delta L_\alpha}+\eta_\alpha\lambda^\alpha\right)\exp\left(-\mathcal{S}+\text{sources}\right)=0\,. \tag{21.37}$$

This equation directly leads to the WT identities (21.33).

The quantities $\delta^2\mathcal{S}/\delta A^i\delta K_i$ and $\delta^2\mathcal{S}/\delta\bar{C}^\alpha\delta L_\alpha$ still vanish after renormalization if some group structure is preserved or quite generally if dimensional regularization has been used.

We now prove that the master equation (21.34), unlike the structure (21.26), is stable under renormalization.

21.5 Renormalization: General Considerations

We now prove the stability of the master equation following a method already explained in the example of the non-linear σ-model (see Chapter 14). We assume that the local action (21.26) has been regularized in a way compatible with gauge invariance. Perturbation theory then exhibits UV divergences which have to be removed by adding counter-terms to the action. The identities (21.33) imply relations among divergences. We use them to prove that equation (21.34) is stable under renormalization. In the next section, we then solve equation (21.34) to find the most general form of the renormalized action.

21.5.1 Counter-terms and master equation

As usual our analysis is based on a loop expansion of the regularized functional Γ, the first term being the unrenormalized action $\mathcal{S}$:

$$\Gamma=\mathcal{S}+\sum_{\ell=1}^{\infty}\Gamma_\ell\,. \tag{21.38}$$

To simplify notations, we introduce an auxiliary field μ_α and add to the action (21.16) and to the 1PI functional Γ the combination $-\mu_\alpha\lambda^\alpha$. The new functional then satisfies

$$\frac{\delta\Gamma}{\delta\mu_\alpha}=-\lambda^\alpha\,, \tag{21.39}$$

and equations (21.33) and (21.34) become homogeneous quadratic equations which can be written symbolically as

$$\mathcal{S} * \mathcal{S} = 0\,, \tag{21.40}$$

$$\Gamma * \Gamma = 0\,. \tag{21.41}$$

The general arguments presented in Section 14.6 now prove that these equations are stable under renormalization. We recall here only the main steps.

Assuming as an induction hypothesis that we have been able to construct a renormalized action $\mathcal{S}_{\ell-1}$ which satisfies equation (21.40) and renders Γ finite at $\ell-1$ loop order, we write the consequences of equation (21.41) at loop order ℓ:

$$\mathcal{S} * \Gamma_\ell + \Gamma_\ell * \mathcal{S} = -\sum_{m=1}^{\ell-1} \Gamma_m * \Gamma_{\ell-m}\,. \tag{21.42}$$

The r.h.s. of the equation is finite by induction. We define Γ_ℓ^{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 then satisfies

$$\mathcal{S} * \Gamma_\ell^{\text{div}} + \Gamma_\ell^{\text{div}} * \mathcal{S} = 0\,. \tag{21.43}$$

Equation (21.43) shows that by defining the ℓ-loop renormalized action $\mathcal{S}_\ell$ by

$$\mathcal{S}_\ell = \mathcal{S}_{\ell-1} - \Gamma_\ell^{\text{div}} + \text{higher orders}\,, \tag{21.44}$$

it is possible to render Γ ℓ-loop finite with a renormalized action still satisfying equation (21.40).

21.5.2 Solution of WT identities: general considerations

It now remains to solve equations (21.43) and (21.40) using power counting arguments to find the general form of the counter-terms and of the renormalized action. It is, however, useful to first exhibit several properties of these equations.

It is convenient to introduce some notation. We denote below by θ_i the set of all anticommuting fields K_i, $\bar{C}^\alpha$, C^α and x_i all commuting fields A^i, L_α, μ_α. As we see on the explicit expression (21.34), the master equation for the action $\mathcal{S}$ (and thus Γ) then takes the form

$$\frac{\partial \mathcal{S}}{\partial x_i}\frac{\partial \mathcal{S}}{\partial \theta_i} = 0\,. \tag{21.45}$$

Equation (21.43) reads

$$\tilde{\mathcal{D}}\Gamma_l^{\text{div}} = 0\,, \tag{21.46}$$

where $\tilde{\mathcal{D}}$ is the differential operator:

$$\tilde{\mathcal{D}} = \frac{\partial \mathcal{S}}{\partial \theta_i}\frac{\partial}{\partial x_i} + \frac{\partial \mathcal{S}}{\partial x_i}\frac{\partial}{\partial \theta_i}, \tag{21.47}$$

or in explicit form

$$\tilde{\mathcal{D}} = \frac{\delta \mathcal{S}}{\delta A^i}\frac{\delta}{\delta K_i} + \frac{\delta \mathcal{S}}{\delta K_i}\frac{\delta}{\delta A^i} + \frac{\delta \mathcal{S}}{\delta \bar{C}^\alpha}\frac{\delta}{\delta L_\alpha} + \frac{\delta \mathcal{S}}{\delta L_\alpha}\frac{\delta}{\delta \bar{C}^\alpha} + \frac{\delta \mathcal{S}}{\delta \mu_\alpha}\frac{\delta}{\delta C^\alpha}\,. \tag{21.48}$$

The nilpotency of the operator $\tilde{\mathcal{D}}$. Quite remarkably the master equation (21.45) implies that $\tilde{\mathcal{D}}^2$ vanishes as we now show. Since $\tilde{\mathcal{D}}$ 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 $\tilde{\mathcal{D}}^2$:

$$\tilde{\mathcal{D}}^2 = \left(\frac{\partial\mathcal{S}}{\partial\theta_i}\frac{\partial^2\mathcal{S}}{\partial x_i\partial\theta_j} + \frac{\partial\mathcal{S}}{\partial x_i}\frac{\partial^2\mathcal{S}}{\partial\theta_i\partial\theta_j}\right)\frac{\partial}{\partial x_j} + \left(\frac{\partial\mathcal{S}}{\partial\theta_i}\frac{\partial^2\mathcal{S}}{\partial x_i\partial x_j} + \frac{\partial\mathcal{S}}{\partial x_i}\frac{\partial^2\mathcal{S}}{\partial\theta_i\partial x_j}\right)\frac{\partial}{\partial\theta_j}. \tag{21.49}$$

The expression can be rewritten as

$$\tilde{\mathcal{D}}^2 = \left[-\frac{\partial}{\partial\theta_j}\left(\frac{\partial\mathcal{S}}{\partial\theta_i}\frac{\partial\mathcal{S}}{\partial x_i}\right)\right]\frac{\partial}{\partial x_j} + \left[\frac{\partial}{\partial x_j}\left(\frac{\partial\mathcal{S}}{\partial\theta_i}\frac{\partial\mathcal{S}}{\partial x_i}\right)\right]\frac{\partial}{\partial\theta_j} = 0\,, \tag{21.50}$$

as a consequence of equation (21.45). The operator $\tilde{\mathcal{D}}$, which plays an essential role in solving WT identities, again is a BRS operator with vanishing square.

Canonical invariance of the master equation. Equation (21.45) has properties analogous to the symplectic form $\mathrm{d}p\wedge\mathrm{d}q$ of classical mechanics, it is invariant under canonical transformations. Let us change variables $(\theta, x)\mapsto(\theta', x')$:

$$x_i = \frac{\partial\varphi}{\partial\theta_i}(\theta, x'), \tag{21.51}$$

$$\theta'_i = \frac{\partial\varphi}{\partial x'_i}(\theta, x'), \tag{21.52}$$

in which $\varphi(x', \theta)$ is an anticommuting type function. We first eliminate x_i in equation (21.45) using equation (21.51):

$$\frac{\partial\mathcal{S}}{\partial\theta_i}\left[\frac{\partial\varphi}{\partial\theta_i\partial x'_j}\right]^{(-1)}\frac{\partial\mathcal{S}}{\partial x'_j} = 0\,. \tag{21.53}$$

We then eliminate θ_i using (21.52). We verify that we recover equation (21.45) in the new variables:

$$\frac{\partial\mathcal{S}}{\partial\theta'_i}\frac{\partial\mathcal{S}}{\partial x'_i} = 0\,. \tag{21.54}$$

Infinitesimal transformations. We consider the infinitesimal form of canonical transformations, that is, expand the function φ to first order in a parameter ε:

$$\varphi = \theta_i x'_i + \varepsilon\psi(\theta, x'). \tag{21.55}$$

Then,

$$x'_i = x_i - \varepsilon\frac{\partial\psi}{\partial\theta_i}(\theta, x) + O\left(\varepsilon^2\right), \qquad \theta'_i = \theta_i + \varepsilon\frac{\partial\psi}{\partial x_i}(\theta, x) + O\left(\varepsilon^2\right). \tag{21.56}$$

We now calculate $\mathcal{S}(\theta', x')$:

$$\mathcal{S}(\theta', x') - \mathcal{S}(\theta, x) = -\varepsilon\frac{\partial\psi}{\partial x'_i}\frac{\partial\mathcal{S}}{\partial\theta_i} - \varepsilon\frac{\partial\psi}{\partial\theta_i}\frac{\partial\mathcal{S}}{\partial x_i} + O\left(\varepsilon^2\right) = -\varepsilon\tilde{\mathcal{D}}\psi + O\left(\varepsilon^2\right), \tag{21.57}$$

where the definition (21.47) has been used.

We thus find that any infinitesimal addition to $\mathcal{S}$ of a BRS exact term can be obtained by a canonical transformation acting on $\mathcal{S}$.

The effect of the quantization procedure can be understood as such a transformation. In our original problem 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 (21.51,21.52):

$$\varphi(A,\bar{C},C,\lambda,K,L,\mu)=C^\alpha\mu_\alpha+\tilde{\varphi}(A,\bar{C},C,\lambda,K,L).$$

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

$$\mathcal{S}\mapsto\mathcal{S}+\lambda^\alpha\frac{\delta\tilde{\varphi}}{\delta C^\alpha}. \tag{21.58}$$

One verifies that the gauge invariant action with the K and L source terms, and the renormalized quantized action (21.76) are related by such a transformation with

$$\tilde{\varphi}(A,\bar{C},C,\lambda,K,L)=C^\alpha\left[F_\alpha(A)-\tfrac{1}{2}a_{\alpha\beta}\lambda^\beta\right]+A^iK_i+L_\alpha\bar{C}^\alpha+\tfrac{1}{2}g_{\alpha\beta\gamma}C^\alpha C^\beta\bar{C}^\gamma. \tag{21.59}$$

21.6 The Renormalized Action

We have shown that the renormalized action satisfies the master equation (21.34). To find the form of the renormalized action it is thus necessary to find the most general solution of equation (21.34) local in the fields and sources, and consistent with power counting. We work in four dimensions and assume a renormalizable gauge .

21.6.1 General gauges

We have to solve equation (21.34) taking into account the power counting, symmetries and locality. First, we note that in the original action only the product $C\bar{C}$ appears. This leads to 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_α -2.

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, that is, 1 and $F(A)$ dimension 2 (we have exhibited such gauges in Chapters 18 and 19),

$$[A]=1\,,\qquad [F(A)]=2\,.$$

We impose $[\lambda]=2$, by choosing $w(\lambda)$ such λ has a constant propagator:

$$w(\lambda)=\tfrac{1}{2}a_{\alpha\beta}\lambda^\alpha\lambda^\beta,\qquad a_{\alpha\beta}\ \text{constants}\,. \tag{21.60}$$

The other dimensions follow. Since $D^i_\alpha(A)$ 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,

$$[A]=1\,,\quad [\bar{C}]=0\,,\quad [C]=2\,,\quad [\lambda]=2\,,\quad [K]=3\,,\quad [L]=4\,. \tag{21.61}$$

Since K and L have dimensions 3 and 4, respectively, the renormalized action can contain at most terms linear in K and L. Similarly, since λ has dimension 2, the renormalized action is at most a polynomial of second degree in λ, and the coefficient of $\lambda^\alpha\lambda^\beta$ is a constant matrix.

The solution. To solve equation (21.34), we now parametrize the solution in a way reminiscent of the initial action (21.26), but it should be kept in mind that the parameters which appear are renormalized and in general different from those parametrizing the action (21.26). The subscript "renormalized" is always implied, and is omitted only for notational simplicity.

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

$$\mathcal{S}(A,\bar{C},C,\lambda,K,L)=-K_iD^i_\alpha(A)\bar{C}^\alpha-\tfrac{1}{2}L_\alpha f^\alpha_{\beta\gamma}\bar{C}^\beta\bar{C}^\gamma+\mathcal{S}(A,\bar{C},C,\lambda).$$

Power counting tells us in addition that $f^\alpha_{\beta\gamma}$ has dimension 0 and is thus field independent and D^i_α has dimension 1 and can, therefore, only be an affine function of A.

The terms linear in L and K in equation (21.34) yield, respectively:

$$f^\beta_{\alpha\gamma}f^\alpha_{\delta\varepsilon}\bar{C}^\gamma\bar{C}^\delta\bar{C}^\varepsilon=0\,, \tag{21.62}$$

$$\left(\frac{\delta D^i_\alpha}{\delta A^j}D^j_\beta-\frac{1}{2}f^\gamma_{\alpha\beta}D^i_\gamma\right)\bar{C}^\alpha\bar{C}^\beta=0\,. \tag{21.63}$$

The first equation implies that the constants $f^\alpha_{\beta\gamma}$ satisfy a Jacobi identity since the product $\bar{C}^\gamma\bar{C}^\delta\bar{C}^\varepsilon$ is antisymmetric in $(\gamma,\delta,\varepsilon)$. They are, therefore, 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 gauge independence of physical observables as derived in Section 21.7. The structure constants are thus linearly related to the structure constants appearing in the initial action (21.26).

Equation (21.62) is also an integrability condition for equation (21.63) which implies the commutation relations:

$$\frac{\delta D^i_\alpha}{\delta A^j}D^j_\beta-\frac{\delta D^i_\beta}{\delta A^j}D^j_\alpha=f^\gamma_{\alpha\beta}D^i_\gamma\,. \tag{21.64}$$

Therefore, the whole group structure is recovered. In terms of the BRS operator

$$\mathcal{D}_+\equiv D^i_\alpha(A)\bar{C}^\alpha\frac{\delta}{\delta A_i}-\frac{1}{2}f^\alpha_{\beta\gamma}\bar{C}^\beta\bar{C}^\gamma\frac{\delta}{\delta\bar{C}^\alpha}\,, \tag{21.65}$$

the two equations can be combined into a unique equation $\mathcal{D}_+^2=0$. The master equation then reduces to a condition of BRS symmetry. Defining also

$$\mathcal{D}_-\equiv\lambda_\alpha\frac{\delta}{\delta C^\alpha}\,, \tag{21.66}$$

we can write it as

$$\mathcal{D}\mathcal{S}(A,\bar{C},C,\lambda) \equiv (\mathcal{D}_+ + \mathcal{D}_-)\,\mathcal{S}(A,\bar{C},C,\lambda) = 0 \tag{21.67}$$

with

$$\mathcal{D}_+^2 = \mathcal{D}_-^2 = 0\,, \quad \mathcal{D}_+\mathcal{D}_- + \mathcal{D}_-\mathcal{D}_+ = 0\,. \tag{21.68}$$

Because $\mathcal{D}_+$ and $\mathcal{D}_-$ have a different power of $\bar{C}$ it is natural to expand $\mathcal{S}(A,\bar{C},C,\lambda)$ in powers of $\bar{C}$, a situation we have already met in Section 16.5.2. Since the product $\bar{C}C$ has dimension 2, $\mathcal{S}$ is a polynomial of degree 2 in $\bar{C}C$:

$$\mathcal{S}(A,\bar{C},C,\lambda) = \mathcal{S}^{(0)}(A,\lambda) + \mathcal{S}^{(1)}(A,\bar{C},C,\lambda) + \mathcal{S}^{(2)}(\bar{C},C)\,.$$

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

$$\mathcal{S}^{(0)}(A,\lambda) = -\tfrac{1}{2}a_{\alpha\beta}\lambda^\alpha\lambda^\beta + \lambda^\alpha F_\alpha(A) + \mathcal{S}(A). \tag{21.69}$$

The equation $\mathcal{D}_-\mathcal{S}^{(0)} = 0$ is automatically satisfied, and we note that $\mathcal{S}^{(0)}$ can be written as

$$\mathcal{S}^{(0)}(A,\lambda) = \mathcal{D}_-\Phi^{(0)} + \mathcal{S}(A) \tag{21.70}$$

with

$$\Phi^{(0)} = -\tfrac{1}{2}C^\alpha a_{\alpha\beta}\lambda^\beta + C^\alpha F_\alpha(A). \tag{21.71}$$

More generally one verifies 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

$$\bar{C}^\alpha D^i_\alpha(A)\frac{\delta\mathcal{S}(A)}{\delta A^i} = 0\,. \tag{21.72}$$

This equation implies that $\mathcal{S}(A)$ is gauge invariant.

Then, from equation (21.70),

$$\mathcal{D}_-\mathcal{S}^{(1)} = -\mathcal{D}_+\mathcal{D}_-\Phi^{(0)} = \mathcal{D}_-\mathcal{D}_+\Phi^{(0)}\,.$$

The general solution is the sum of a special solution, and a $\mathcal{D}_-$ exact term

$$\mathcal{S}^{(1)} = \mathcal{D}_+\Phi^{(0)} + \mathcal{D}_-\Phi^{(1)}, \tag{21.73}$$

where $\Phi^{(1)}$ is proportional to $C^2\bar{C}$:

$$\Phi^{(1)} = \tfrac{1}{2}g_{\alpha\beta\gamma}C^\alpha C^\beta\bar{C}^\gamma\,, \quad g_{\alpha\beta\gamma} = -g_{\beta\alpha\gamma}\,. \tag{21.74}$$

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

$$\mathcal{S}^{(2)} = \mathcal{D}_+\Phi^{(1)}, \tag{21.75}$$

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

$$\mathcal{S}(A,\bar{C},C,\lambda) = \mathcal{S}(A) + \mathcal{D}\Phi(A,\bar{C},C,\lambda), \tag{21.76}$$

where $\mathcal{S}(A)$ is gauge invariant and

$$\Phi(A,\bar{C},C,\lambda) = \Phi^{(0)} + \Phi^{(1)} = -\tfrac{1}{2}C^\alpha a_{\alpha\beta}\lambda^\beta + C^\alpha F_\alpha(A) + \tfrac{1}{2}g_{\alpha\beta\gamma}C^\alpha C^\beta \bar{C}^\gamma. \tag{21.77}$$

The renormalized action has a form similar to the bare action (21.26) except for an additional BRS exact term

$$\mathcal{S}_4(\lambda,\bar{C},C) = \mathcal{D}\Phi^{(1)} = g_{\alpha\beta\gamma}\left(-\tfrac{1}{4}C^\alpha C^\beta f^\gamma_{\delta\varepsilon}\bar{C}^\delta\bar{C}^\varepsilon + \lambda^\alpha C^\beta \bar{C}^\gamma\right). \tag{21.78}$$

This term has exactly the form given in equation (21.57) and can be associated with a shift $L_\alpha \mapsto L_\alpha + \frac{1}{2}g_{\beta\gamma\alpha}C^\beta C^\gamma$.

The quartic ghost term. A comment now is in order: since in general the renormalized action is quartic in the ghost terms, 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 to the gauge function F_α a term linear in an auxiliary field transforming non-trivially under the gauge group, then the integration over this auxiliary field with an appropriate gaussian weight generates the quartic ghost terms in their most general form.

This property is expected from the general analysis of Chapter 16.

Renormalization of gauge invariant operators. 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 (21.34) is a non-trivial problem and requires more sophisticated cohomology techniques. In the case of compact Lie groups with semi-simple Lie algebras, the most general solution of equation (21.46) is the sum of gauge invariant terms and BRS exact contributions, that is, of the form $\mathcal{D}\Phi$. This result first conjectured has now 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 averages of products of gauge invariant operators, or the matrix elements between physical states, as we show in Section 21.7, are gauge-independent.

21.6.2 Linear gauges

For a special class of gauges, the preceding analysis can be simplified. This class is characterized by the property that the gauge fixing function $F_\alpha(A)$ is linear in the field A, rather than quadratic as in the most general renormalizable case:

$$F_\alpha(A) = F_{\alpha i} A^i \,. \tag{21.79}$$

In this case, 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 (21.79), we use the λ-field equation of motion. We again explicitly parametrize $w(\lambda)$ as

$$w(\lambda) = \tfrac{1}{2} a_{\alpha\beta} \lambda^\alpha \lambda^\beta \,. \tag{21.80}$$

Then, the λ-field equation of motion, which relies on the identity

$$\int [\mathrm{d}\lambda] \frac{\delta}{\delta \lambda^\alpha} \left[\exp\left(-\mathcal{S}(A, C, \bar{C}, \lambda, K, L) + \text{sources}\right)\right] = 0$$

with $\mathcal{S}$ given by equation (21.26), reads

$$\left(-a_{\alpha\beta} \frac{\delta}{\delta l_\beta} + F_{\alpha i} \frac{\delta}{\delta J_i} - l_\alpha\right) \mathcal{Z} = 0 \,, \tag{21.81}$$

or in terms of the functional $\mathcal{W} = \ln \mathcal{Z}$:

$$\left(a_{\alpha\beta} \frac{\delta}{\delta l_\beta} + F_{\alpha i} \frac{\delta}{\delta J_i}\right) \mathcal{W} = l_\alpha \,. \tag{21.82}$$

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 proper vertices Γ is simple:

$$\frac{\delta \Gamma}{\delta \lambda^\alpha} = -a_{\alpha\beta} \lambda^\beta + F_{\alpha i} A^i \,. \tag{21.83}$$

Equation (21.83) 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 $g_{\alpha\beta\gamma} C^\beta \bar{C}^\gamma$ can be generated. The action remains quadratic in the ghost fields. The renormalized action takes the simple form

$$\mathcal{S}(A, \bar{C}, C, \lambda, K, L) = \mathcal{S}(A) + \mathcal{D}\left[C^\alpha F_{\alpha i} A^i - \tfrac{1}{2} a_{\alpha\beta} C^\alpha \lambda^\beta + K_i A_i + L_\alpha \bar{C}^\alpha\right], \tag{21.84}$$

($\mathcal{D}$ being defined by equation (21.67)) where in addition as stated above $a_{\alpha\beta}$ and $F_{\alpha i}$ are unrenormalized.

Remark. In the case of linear gauges another equation can be used to show that the gauge function is unrenormalized, the C ghost equation of motion. The equation

$$\frac{\delta \mathcal{S}}{\delta C^\alpha} = F_{\alpha i} D^i_\beta \bar{C}^\beta = F_{\alpha i} \frac{\delta \mathcal{S}}{\delta K_i} \tag{21.85}$$

implies

$$\left(F_{\alpha i}\frac{\delta}{\delta K_i}+\eta_\alpha\right)\mathcal{Z}=0\,, \tag{21.86}$$

and thus, after Legendre transformation,

$$\left(F_{\alpha i}\frac{\delta}{\delta K_i}-\frac{\delta}{\delta C^\alpha}\right)\Gamma=0\,. \tag{21.87}$$

Equation (21.85) is thus stable under renormalization.

We mention this property here for the following reasons: in the case of linear gauges, the introduction of the λ-field is not always useful, except for strict gauge conditions (generalized Landau gauges). If we integrate over λ^α and set l_α to zero, then equation (21.83) disappears, while equation (21.87) remains and can be used to show that the gauge fixing term $F(A)$ is not renormalized.

21.7 Gauge Independence

General correlation functions are gauge-dependent and, therefore, cannot be associated with physical observables. Using arguments similar to those given in Section 16.5.2, we show here that expectation values of gauge invariant operators 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 (21.3)) and the effective action BRS symmetric. Assuming a gauge invariant regularization, we examine how the vacuum amplitude is affected by an infinitesimal change of gauge δF_α, before renormalization.

Since the non-gauge invariant part of the action has the general form $\mathcal{D}_0\Phi$ in which $\mathcal{D}_0$ has been defined by equation (21.21) (see equation (21.23)), the variation $\delta\mathcal{S}$ of the action takes the form $\delta\mathcal{S}=\mathcal{D}_0(\delta\Phi)$, and thus the variation of the vacuum amplitude is

$$\delta\mathcal{Z}=-\int[\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda]\mathcal{D}_0\left(\delta\Phi\right)\mathrm{e}^{-\mathcal{S}}\,. \tag{21.88}$$

The operator $\mathcal{D}_0$ is a differential operator and, therefore, we can integrate by parts. Using again the property that the traces $f^\alpha_{\alpha\beta}$ and $\delta D^i_\alpha\,/\delta A^i$ vanish we obtain

$$\delta\mathcal{Z}=\int[\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda]\delta\Phi\left(\mathcal{D}_0\mathcal{S}\right)\mathrm{e}^{-\mathcal{S}}, \tag{21.89}$$

and this expression vanishes as a consequence of the BRS symmetry.

We have, therefore, shown the bare correlation functions of gauge invariant operators are gauge-independent, at least within a class of gauges which can be continuously connected. There exists, therefore, a renormalization procedure 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.

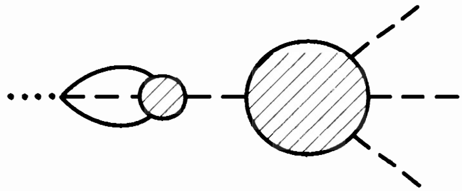


Fig. 21.1

S-matrix elements. We now want to study 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 above. 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 BRS operator) considered in Section 21.6.1. We can still integrate by parts, but the resulting integrand does not vanish identically because we have introduced sources for non-gauge invariant fields:

$$\delta\mathcal{Z}(J) = -\int [\mathrm{d}A\mathrm{d}\bar{C}\mathrm{d}C\mathrm{d}\lambda]\delta\Phi J_i D^i_\alpha(A)\bar{C}^\alpha \exp\left(-\mathcal{S} + J_i A_i\right). \tag{21.90}$$

When we calculate correlation functions we obtain a sum of contributions in which one A_i field has been replaced by $\delta\Phi J_i D^i_\alpha(A)\bar{C}^\alpha$, which is a linear combination of composite operators. When we go to the mass-shell, after amputation, we get a non-vanishing contribution 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. Then, on the mass-shell, we get a contribution proportional to the matrix element of the field itself (see figure 21.1). This argument was presented for the first time in Appendix A7.2. The final result is that an infinitesimal change of gauge renormalizes multiplicatively the S-matrix elements. This corresponds to a field amplitude renormalization. Therefore, the S-matrix, properly normalized, is gauge-independent.

Bibliographical Notes

Renormalization of gauge theories is discussed in

G. 't Hooft, *Nucl. Phys.* B33 (1971) 173; *ibidem* B35 (1971) 167; A.A. Slavnov, *Theor. Math. Phys.* 10 (1972) 99; J.C. Taylor, *Nucl. Phys.* B33 (1971) 436; B.W. Lee and J. Zinn-Justin, *Phys. Rev.* D5 (1972) 3121, 3137, 3155; *ibidem* D7 (1973) 1049; G. 't Hooft and M. Veltman, *Nucl. Phys.* B50 (1972) 318; D.A. Ross and J.C. Taylor, *Nucl. Phys.* B51 (1973) 125; B.W. Lee, *Phys. Lett.* 46B (1973) 214; *Phys. Rev.* D9 (1974) 933.

The anticommuting type symmetry of the quantized action is exhibited in

C. Becchi, A. Rouet and R. Stora, *Comm. Math. Phys.* 42 (1975) 127,

where it is used to renormalize the abelian Higgs model. In

H. Kluberg-Stern and J.-B. Zuber, *Phys. Rev.* D12 (1975) 467,

quadratic WT identities generated by BRS symmetry appear and and the renormalization of a class of gauge invariant operators is discussed.

The first derivation of the master equation for the quantized action and the proof of renormalizability in an arbitrary gauge presented here, are found in the proceedings of the Bonn summer school 1974,

J. Zinn-Justin in *Trends in Elementary Particle Physics (Lectures Notes in Physics* 37), H. Rollnik and K. Dietz eds. (Springer-Verlag, Berlin 1975).

For more details, see also

J. Zinn-Justin in *Proc. of the 12th School of Theoretical Physics, Karpacz 1975*, Acta Universitatis Wratislaviensis 368; B.W. Lee in *Methods in Field Theory*, Les Houches School 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976); J. Zinn-Justin, *Mod. Phys. Lett.* A19 (1999) 1227.

An alternative proof based on BRS symmetry and the BPHZ formalism can be found in

C. Becchi, A. Rouet and R. Stora, *Ann. Phys. (NY)* 98 (1976) 287.

Renormalization of background field gauges is discussed in

H. Kluberg-Stern and J.-B. Zuber, *Phys. Rev.* D12 (1975) 482, 3159.

A general study of gauge invariant operators can be found in

S.D. Joglekar and B.W. Lee, *Ann. Phys. (NY)* 97 (1976) 160.

The complete proof of the conjectured result is found in

G. Barnich, and M. Henneaux, *Phys. Rev. Lett.* 72 (1994) 1588; G. Barnich, F. Brandt and M. Henneaux, *Comm. Math. Phys.* 174 (1995) 93.

Attempts to quantize in quadratic gauges include

G. 't Hooft and M. Veltman, *Nucl. Phys.* B50 (1972) 318; S.D. Joglekar, *Phys. Rev.* D10 (1974) 4095.

Consequences of BRS symmetry in the operator formalism are exhibited in

T. Kugo and I. Ojima, *Phys. Lett.* B73 (1978) 459, *Nucl. Phys.* B144 (1978) 234.

The problem of non-linear gauges has been discussed in its full generality in

J. Zinn-Justin, *Nucl. Phys.* B246 (1984) 246.

22 CLASSICAL AND QUANTUM GRAVITY. RIEMANNIAN MANIFOLDS AND TENSORS

This chapter has two purposes; to present the few elements of differential geometry which are required in different places in this volume and to provide, for completeness, a short introduction to the problem of quantization of gravity.

We first briefly recall a few concepts related to reparametrization (more accurately diffeomorphism) of Riemannian manifolds. We introduce the notions of parallel transport, affine connection, curvature, in analogy with gauge theories as discussed in Chapters 19–21. To define fermions on Riemannian manifolds additional mathematical objects are required, the vielbein and the spin connection. We then construct Einstein's action for classical gravity (General Relativity) and derive the equation of motion. In the last section, we finally study the formal aspects of the quantization of the theory of gravity, following the lines of the quantization of non-abelian gauge theories of Chapter 19.

Because the theory of quantum gravity is not renormalizable in four dimensions (even in its extended forms like supersymmetric gravity), the general prejudice at present time is that gravity is the low energy, large distance, remnant of a more complete theory which probably no longer has the form of a quantum field theory (strings, non-commutative geometry?): in the language of critical phenomena, gravity belongs to the class of irrelevant interactions (due to the presence of the massless graviton the situation can be compared with the interaction of Goldstone modes at very low temperature in the ordered phase). The scale of this new physics is indicated by Planck's mass: it is of the order of $\sqrt{\hbar c/G_N} \sim 10^{19}$ GeV, where G_N is Newton's gravitational constant.

The appendix is devoted to a short introduction to 2D quantum gravity, approached from the point of view of matrix models in the large N limit.

We will not be concerned with mathematical rigour and the notation will be old-fashioned. For instance, we shall write most expressions in terms of *local coordinates*, ignoring, 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 language of fibre bundles will be avoided. The reader interested in more details is referred to the literature.

The convention of summation over repeated lower and upper indices will always be used, except when the metric is explicitly euclidean.

22.1 Change of Coordinates. Tensors

Let φ^i, $i = 1, \ldots, N$, be a set of local coordinates which parametrize a manifold $\mathfrak{M}$. A (locally) non-singular change of coordinates or reparametrization $\varphi \mapsto \varphi'$ is defined by a set of differentiable functions $\varphi^i(\varphi')$,

$$\varphi^i = \varphi^i(\varphi'), \tag{22.1}$$

and the mapping $\varphi' \mapsto \varphi$ is locally invertible (the proper extension to the complete manifold is a diffeomorphism).

If we define

$$\mathrm{d}\varphi^i = T^i_j(\varphi')\mathrm{d}\varphi'^j \;\Leftrightarrow\; T^i_j(\varphi') = \frac{\partial \varphi^i}{\partial \varphi'^j} \equiv \partial_j \varphi^i, \tag{22.2}$$

the matrix T^i_j thus is an element of the defining representation of the general linear group $GL(N)$ (general invertible matrices).

We now consider a set of fields defined on $\mathfrak{M}$ and classify them according to their transformation properties in a reparametrization of $\mathfrak{M}$.

Fields $S(\varphi)$ which transform by a simple substitution

$$S'(\varphi') = S\bigl(\varphi(\varphi')\bigr) \tag{22.3}$$

are called scalars.

Quantities like $\mathrm{d}\varphi^i$, which belong to the space tangent to the manifold at point φ^i, define the transformation properties of vector fields $V^i(\varphi)$ (called also contravariant vectors):

$$V^i(\varphi) = T^i_j V'^j(\varphi'). \tag{22.4}$$

Vector fields $W_i(\varphi)$ of the dual space are defined by the property that the scalar product $V^i(\varphi)W_i(\varphi)$ is a scalar (they are also called covariant vectors). This implies the transformation

$$W_i(\varphi) = \left(T^{-1}\right)^j_i W'_j(\varphi'), \tag{22.5}$$

in which $\left(T^{-1}\right)^j_i$ is the inverse of the transposed of the matrix T^i_j (and, therefore, belongs to another representation of $GL(N)$). An example is provided by partial derivatives of a scalar. Indeed differentiating equation (22.3), we find

$$\partial_i S = \frac{\partial\varphi'^j}{\partial\varphi^i} = \left(T^{-1}\right)^j_i \partial_j S', \tag{22.6}$$

which shows that $\partial_i S$ transforms like a dual vector.

Transformations of vectors are thus defined in terms of the substitution $\varphi \mapsto \varphi'$ and a linear transformation. More generally, we can classify all quantities with respect to their transformation properties under the linear group. General n-tensors $V^{i_1\ldots i_p}_{i_{p+1}\ldots i_n}$ transform like the tensor product of p vectors and $n-p$ dual vectors:

$$V^{i_1\ldots i_p}_{i_{p+1}\ldots i_n}(\varphi) = T^{i_1}_{j_1}\ldots T^{i_p}_{j_p}\left(T^{-1}\right)^{j_{p+1}}_{i_{p+1}}\ldots\left(T^{-1}\right)^{j_n}_{i_n} V'^{\,j_1\ldots j_p}_{j_{p+1}\ldots j_n}(\varphi'). \tag{22.7}$$

For what follows it is useful to introduce a notation for the element $\mathbf{T}$ of the abstract linear group $GL(N)$. Then,

$$\mathbf{T}V'^{\,i_1\ldots i_p}_{i_{p+1}\ldots i_n}(\varphi') = T^{i_1}_{j_1}\ldots T^{i_p}_{j_p}\left(T^{-1}\right)^{j_{p+1}}_{i_{p+1}}\ldots\left(T^{-1}\right)^{j_n}_{i_n} V'^{\,j_1\ldots j_p}_{j_{p+1}\ldots j_n}(\varphi').$$

This representation can in general be reduced according to the irreducible representations of the permutation group acting on lower or upper indices. Moreover, summing a n-tensor over a pair of upper and lower indices yields a $(n-2)$-tensor. We call this operation taking the covariant trace. Then using the property that the tensor δ^j_i, where δ is the Kronecker δ, is an invariant tensor we can render tensors traceless.

Infinitesimal change of coordinates. Often in this work we consider infinitesimal transformations. An infinitesimal change of coordinates can be written as

$$\varphi^i = \varphi'^i + \epsilon^i(\varphi'), \tag{22.8}$$

and all quantities are expanded in ϵ. We denote by $\delta_\epsilon V \sim \mathbf{T}V' - V$, the variation at first order in ϵ of all tensors. We find

$$\delta_\epsilon V^{i_1 \dots i_p}_{i_{p+1} \dots i_n}(\varphi) = \epsilon^j \partial_j V^{i_1 \dots i_p}_{i_{p+1} \dots i_n}(\varphi) - \sum_{\ell=1}^{p} \partial_j \epsilon^{i_\ell} V^{i_1 \dots j \dots i_p}_{i_{p+1} \dots i_n}(\varphi)$$
$$+ \sum_{\ell=p+1}^{n} \partial_{i_\ell} \epsilon^j V^{i_1 \dots i_p}_{i_{p+1} \dots j \dots i_n}(\varphi). \tag{22.9}$$

By construction $\delta_\epsilon V$, being the difference between two tensors, is a tensor. Applied to a tensor product δ_ϵ satisfies

$$\delta_\epsilon (V \otimes W) = (\delta_\epsilon V) \otimes W + V \otimes (\delta_\epsilon W).$$

Finally, note the commutation relation

$$[\delta_{\epsilon_1}, \delta_{\epsilon_2}] = \delta_{\epsilon_3} \quad \text{with} \quad \epsilon_3^i = \epsilon_1^j \partial_j \epsilon_2^i - \epsilon_2^j \partial_j \epsilon_1^i \,. \tag{22.10}$$

Differential forms. A class of tensors is often encountered which have special properties: totally antisymmetric tensors or forms. They can be conveniently represented by contracting indices with the generators θ^i of a Grassmann algebra (see also Section 1.4):

$$\Omega = \theta^{i_1} \theta^{i_2} \dots \theta^{i_n} \Omega_{i_1 i_2 \dots i_n} \,.$$

Examples are provided by gauge theories where the vector potential A_μ can be considered as a one-form and the curvature $F_{\mu\nu}$ (the electromagnetic tensor in QED) as a two-form (see Chapters 18,19).

The form differentiation, denoted below by d, is defined by

$$\mathrm{d} \equiv \theta^j \frac{\partial}{\partial \varphi_j} \,. \tag{22.11}$$

It is clear that d is nilpotent because the θ^i's anticommute:

$$\mathrm{d}^2 = \theta^j \partial_j \theta^k \partial_k = 0 \,.$$

With this notation, the transformation (22.7) acting on Ω can be written as

$$\Omega(\varphi) = \mathrm{d}\varphi'^{i_1} \mathrm{d}\varphi'^{i_2} \dots \mathrm{d}\varphi'^{i_n} \Omega'_{i_1 i_2 \dots i_n}(\varphi').$$

The operator d has an important property: acting on an n-form it yields an $(n+1)$-form, that is, an antisymmetric tensor with $n+1$ indices

$$\mathrm{d}\Omega = \theta^j \partial_j \, \theta^{i_1} \theta^{i_2} \dots \theta^{i_n} \Omega_{i_1 i_2 \dots i_n} = \tfrac{1}{n+1} \theta^{i_1} \theta^{i_2} \dots \theta^{i_{n+1}} \left[\partial_{i_{n+1}} \Omega_{i_1 i_2 \dots i_n} \right]_{\text{antisym.}} \,. \tag{22.12}$$

Let us verify that the antisymmetric quantity in the r.h.s. is indeed a tensor. Using $\mathrm{d}^2 = 0$, we find

$$\begin{aligned} \mathrm{d}\Omega(\varphi) &= \mathrm{d} \left(\mathrm{d}\varphi'^{i_1} \mathrm{d}\varphi'^{i_2} \dots \mathrm{d}\varphi'^{i_n} \Omega'_{i_1 i_2 \dots i_n}(\varphi') \right) \\ &= \mathrm{d}\varphi^{i_1} \mathrm{d}\varphi^{i_2} \dots \mathrm{d}\varphi^{i_n} \mathrm{d}\varphi'^{j} \partial_j \Omega'_{i_1 i_2 \dots i_n}(\varphi'), \end{aligned}$$

which proves the tensor property.

In this notation, 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

$$\mathbf{F} = 2(\mathrm{d}\mathbf{A} + \mathbf{A}^2).$$

A form Ω which satisfies $\mathrm{d}\Omega = 0$ is called a *closed form.* A form Ω which can be written as $\Omega = \mathrm{d}\Omega'$ is *exact.* Since $\mathrm{d}^2 = 0$ any exact form is closed.

Conversely, a typical problem in differential forms is to know whether a closed form is exact. The corresponding partial differential equations can always be integrated locally but depending on the topology of the φ-space sometimes no global solution can be found. Such a situation has been encountered in the quantization of spin variables (Appendix A3.2.1).

Finally, the property that by differentiation one generates new tensors is not general as we discuss in the next section.

22.2 Parallel Transport: Connection, Covariant Derivative

We have seen that the derivatives $\partial_i S$ of a scalar form a tensor. However, the derivative of a vector is not a tensor as one easily verifies. What is needed 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 a new concept, parallel transport. Covariant derivatives will then appear in the limit of infinitesimally close points. A similar problem has already been encountered in the context of gauge theories in Chapters 18–21.

22.2.1 Parallel transport

We consider two points, $\varphi_{(1)}, \varphi_{(2)}$ on the manifold and an oriented, continuous, piecewise differentiable curve C joining them. To this curve, we associate a linear mapping $\mathbf{U}(C)$ from the tangent space at $\varphi_{(1)}$ to the tangent space at $\varphi_{(2)}$, represented by a matrix with elements $U^i_j(C)$. We call this operation parallel transport from $\varphi_{(1)}$ to $\varphi_{(2)}$. Let $V^i(\varphi_{(1)})$ be a vector belonging to the space tangent at point $\varphi_{(1)}$. It is transformed into

$$V^i_U = U^i_j(C) V^j_{(1)}(\varphi_{(1)}), \tag{22.13}$$

which now belongs to the space tangent at the point $\varphi_{(2)}$.

We impose to the mapping $\mathbf{U}(C)$ to satisfy several conditions:

$\mathbf{U}(C)$ is a differentiable functional of the curve C.

If a curve C_1 goes from $\varphi_{(1)}$ to $\varphi_{(2)}$ and a curve C_2 from $\varphi_{(2)}$ to $\varphi_{(3)}$ then,

$$\mathbf{U}(C_1 \cup C_2) = \mathbf{U}(C_2)\mathbf{U}(C_1). \tag{22.14}$$

The definition (22.14) implies that to a curve reduced to a point corresponds the group identity:

$$\mathbf{U}(C \equiv 1 \text{ point}) = \mathbf{1} \ \Leftrightarrow \ U^i_j(C \equiv 1 \text{ point}) = \delta^i_j\,. \tag{22.15}$$

If C is a curve going from $\varphi_{(1)}$ to $\varphi_{(2)}$, we denote by C^{-1} the same curve but oriented from $\varphi_{(2)}$ to $\varphi_{(1)}$. Then,

$$\mathbf{U}(C^{-1}) = \mathbf{U}^{-1}(C). \tag{22.16}$$

The transformation properties of other tensors follow from the following rules: a scalar is invariant in a parallel transport. Demanding the invariance of the scalar product then determines the form of parallel transport for dual vectors:

$$[V_U]_i = \left[U^{-1}\right]_i^j V_j \,.$$

If we denote by $\mathbf{U}(C)$ the extension of the linear transformation to all tensors then the tensor product $V \otimes W$ transform like

$$\mathbf{U}(C)(V \otimes W) = \mathbf{U}(C)V \otimes \mathbf{U}(C)W \,. \tag{22.17}$$

We can, therefore, define parallel transport for all tensors, by imposing that a tensor transforms as a tensor product of vectors.

Diffeomorphisms. It is easy to verify that the transformation (22.4) of tensors under reparametrization induces a transformation for the parallel transporter $U(C)$ which is

$$\mathbf{U}(C) = \mathbf{T}(\varphi'_{(2)})\mathbf{U}'(C)\mathbf{T}^{-1}(\varphi'_{(1)}), \tag{22.18}$$

in which we have assumed that the curve C goes from the point $\varphi'_{(1)}$ to the point $\varphi'_{(2)}$. This equation shows that $U_j^i(C)$ is a (non-local) tensor.

The affine connection. We now explain the relation between parallel transport and the notions of connection and covariant derivative. We have assumed that the matrix $\mathbf{U}(C)$ as a functional of C is differentiable. Due to the composition law (22.14), $\mathbf{U}(C)$ is then entirely determined by its value for infinitesimal curves. We thus consider an infinitesimal differentiable curve, that is, a straight line C connecting two close points φ and $\tilde{\varphi} = \varphi + \delta\varphi$ and set

$$\mathbf{U} = \mathbf{1} - \mathbf{\Gamma}_k \delta\varphi^k + o\,\|\delta\varphi\| \,, \tag{22.19}$$

where $\mathbf{\Gamma}_k$ is called the affine *connection* on the manifold. The connection is entirely characterized by its action on vectors. We denote by Γ^j_{ik}, the Christoffel symbol, the matrix elements of the connection

$$U_j^i(C) = \delta_j^i - \Gamma^i_{jk}(\varphi)\delta\varphi^k + o\,\|\delta\varphi\| \,. \tag{22.20}$$

Conversely, the connection field completely characterizes parallel transport. Indeed, once a curve C is parametrized in terms of a parameter t, the corresponding parallel-transporter $U(C)$ and $\Gamma^i_{jk}(\varphi)\mathrm{d}\varphi^k/\mathrm{d}t$ play the roles respectively of the evolution operator and the hamiltonian of quantum mechanics. Therefore, as shown in Appendix A6.2, $U(C)$ can be written as a path-ordered integral along the curve C:

$$\mathbf{U} = \mathrm{P}\,\exp\left[-\oint_C \mathbf{\Gamma}_k \mathrm{d}\varphi^k\right], \tag{22.21}$$

where P stands for path-ordered.

Using equation (22.18), it is possible to determine how the affine connection transforms under reparametrization. It is convenient to express it in terms of the one-form $\mathbf{\Gamma} = \mathbf{\Gamma}_k \mathrm{d}\varphi^k$ associated with the matrix $\mathbf{\Gamma}_k$. One then finds

$$\mathbf{\Gamma} = \mathbf{T}\mathbf{\Gamma}'\mathbf{T}^{-1} - \mathrm{d}\mathbf{T}\mathbf{T}^{-1} = \mathbf{T}\mathbf{\Gamma}'\mathbf{T}^{-1} + \mathbf{T}\mathrm{d}\mathbf{T}^{-1}, \tag{22.22}$$

where $\mathbf{dT}$ is the one-form $d\varphi'^i\partial_i\mathbf{T}$. In component form, in the defining representation, the equation becomes

$$\Gamma'^{\,i}_{jk} = \left(T^{-1}\right)^i_l \partial_k T^l_j + \left(T^{-1}\right)^i_l \Gamma^l_{mn} T^n_k T^m_j \,. \tag{22.23}$$

We note that the connection does not transform like a tensor, because the transformation is not linear but only affine. Let us, however, decompose the tensor Γ^i_{jk} into symmetric and antisymmetric parts, $\mathcal{G}^i_{jk}$ and $\mathcal{T}^i_{jk}$, respectively, in the exchange $j \leftrightarrow k$:

$$\Gamma^i_{jk} = \mathcal{G}^i_{jk} + \mathcal{T}^i_{jk}\,, \qquad \mathcal{G}^i_{jk} = \mathcal{G}^i_{kj} \quad \mathcal{T}^i_{jk} = -\mathcal{T}^i_{kj}\,. \tag{22.24}$$

Because T^i_j has the form (22.2), the inhomogeneous term in (22.23) is symmetric in $j \leftrightarrow k$. Hence, both quantities transform independently under (22.23). It follows that the antisymmetric part $\mathcal{T}^i_{jk}$ is a tensor. Moreover, the restriction to connections Γ^i_{jk} such that $\mathcal{T}^i_{jk} = 0$, that is, symmetric in $j \leftrightarrow k$, is consistent with the transformation law (22.23) and characterizes the special class of parallel transports without *torsion.*

22.2.2 The covariant derivative

It is now possible to compare the vector field at point $\varphi + \delta\varphi$ with the same vector field parallel-transported from the point φ to the point $\varphi + \delta\varphi$. Using the definition (22.19) and equation (22.13) we obtain

$$\mathbf{U}(C)V^i(\varphi) - V^i(\varphi + \delta\varphi) = -\left[\partial_k V^i(\varphi) + \Gamma^i_{jk}(\varphi)V^j(\varphi)\right]\delta\varphi^k + o\left\|\delta\varphi\right\|. \tag{22.25}$$

The difference in the l.h.s. is a vector belonging to the tangent space at $\varphi + \delta\varphi$. In terms of the *covariant derivative* ∇_k whose action on vectors is

$$\nabla_k V^i = \partial_k V^i + \Gamma^i_{jk} V^j, \tag{22.26}$$

the equation can be written at leading order in $\delta\varphi$ as

$$\mathbf{U}(C)V^i(\varphi + \delta\varphi) - V^i(\varphi + \delta\varphi) = -\delta\varphi^k \nabla_k V^i(\varphi) + o\left\|\delta\varphi\right\|. \tag{22.27}$$

The covariant derivative characterizes infinitesimal parallel transport and allows to construct new local covariant quantities, that is, tensors, from derivatives of tensors.

Parallel transport of general tensors yields the form of the corresponding covariant derivative:

$$\nabla_i = \partial_i \mathbf{1} + \mathbf{\Gamma}_i\,, \tag{22.28}$$

to which we can associate the covariant form differentiation

$$\nabla = \mathrm{d} + \mathbf{\Gamma} \equiv \mathrm{d}\varphi^i\left(\mathbf{1}\partial_i + \mathbf{\Gamma}_i\right). \tag{22.29}$$

The explicit representation of the covariant derivative thus 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 dual vector is

$$\nabla_i V_j = \partial_i V_j - \Gamma^k_{ji} V_k\,. \tag{22.30}$$

From equation (22.17), 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 $V \otimes W$ we find

$$\nabla(V \otimes W) = \nabla(V) \otimes W + V \otimes \nabla(W).$$

The general form of the covariant derivative of a tensor with n indices follows:

$$\nabla_i V^{j_1\ldots j_p}_{j_{p+1}\ldots j_n} = \partial_i V^{j_1\ldots j_p}_{j_{p+1}\ldots j_n} + \sum_{\ell=1}^{p} \Gamma^{j_\ell}_{k_\ell i} V^{j_1\ldots k_\ell \ldots j_p}_{j_{p+1}\ldots j_n} - \sum_{\ell=p+1}^{n} \Gamma^{k_\ell}_{j_\ell i} V^{j_1\ldots j_p}_{j_{p+1}\ldots k_\ell \ldots j_n}, \tag{22.31}$$

which is a tensor with $n+1$ indices.

Remarks.

(i) For the class of symmetric Christoffel connections Γ^i_{jk} (torsion-free transport), one has identity between the derivative of a form and its covariant derivative. In the notation (22.12)

$$\mathrm{d} \equiv \theta^i \partial_i = \theta^i \nabla_i ,$$

as one easily verifies. In particular, the ordinary curl of a vector is a two-form and thus the covariant curl and the ordinary curl coincide:

$$\partial_i V_j - \partial_j V_i = \nabla_i V_j - \nabla_j V_i . \tag{22.32}$$

(ii) The infinitesimal form of the transformation (22.23) of the Christoffel connection corresponding to (22.8), which is not homogeneous, can be written as

$$\delta_\epsilon \Gamma^i_{jk} = \partial_j \partial_k \epsilon^i + \epsilon^l \partial_l \Gamma^i_{jk} - \partial_l \epsilon^i \Gamma^l_{jk} + \partial_k \epsilon^l \Gamma^i_{jl} + \partial_j \epsilon^l \Gamma^i_{lk}. \tag{22.33}$$

Although this is not obvious from this expression, $\delta_\epsilon \Gamma^i_{jk}$ is a tensor.

22.3 The Metric Tensor

We now introduce the concept of distance in $\mathfrak{M}$. It is characterized by the line element $\mathrm{d}s$, distance from the point φ^i to the point $\varphi^i + \mathrm{d}\varphi^i$:

$$(\mathrm{d}s)^2 = g_{ij}(\varphi)\mathrm{d}\varphi^i \mathrm{d}\varphi^j ,$$

where 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 $g_{ij}(\varphi)$ in a reparametrization:

$$g_{ij}(\varphi) = \left(T^{-1}\right)^k_i \left(T^{-1}\right)^l_j g'_{kl}(\varphi'). \tag{22.34}$$

It is consistent with the notation (22.7) to write the inverse (in the sense of matrices) of the metric tensor as $g^{ij}(\varphi)$:

$$g_{ik} g^{kj} = \delta^j_i . \tag{22.35}$$

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,

$$V_i = g_{ij} V^j . \tag{22.36}$$

Compatibility of parallel transport with the metric. We say that parallel transport is compatible with the metric if it leaves invariant the scalar product of two vectors:

$$V^i(\varphi) g_{ij}(\varphi) W^j(\varphi) = \tilde{V}^i(\tilde{\varphi}) g_{ij}(\tilde{\varphi}) \widetilde{W}^j(\tilde{\varphi}), \tag{22.37}$$

in which $\tilde{V}^i$ and $\widetilde{W}^i$ are the parallel-transported of V^i and W^i from φ to $\tilde{\varphi}$. Using the definition (22.13), we can rewrite the equation

$$g_{ij}(\varphi) = {}^T U_i^k(C) g_{kl}(\tilde{\varphi}) U_j^l(C), \tag{22.38}$$

where we denote by ${}^T\mathbf{U}$ the transposed of the matrix $\mathbf{U}$.

The interpretation of this equation is simple: the metric transported from φ to $\tilde{\varphi}$ is identical to the metric at point $\tilde{\varphi}$. Taking then the limit of an infinitesimal curve, we conclude

$$\nabla_m g_{ij} = 0 . \tag{22.39}$$

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 (22.39). Using the explicit form (22.31), it can be written as

$$\partial_m g_{ij} - g_{il} \Gamma^l_{jm} - \Gamma^k_{im} g_{kj} = 0 . \tag{22.40}$$

If, eliminating torsion in parallel transport, we assume Γ^i_{jk} symmetric we can solve the compatibility condition and express Γ^i_{jk} explicitly in terms of the metric tensor. We find

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{il} \left(\partial_j g_{lk} + \partial_k g_{lj} - \partial_l g_{jk} \right) . \tag{22.41}$$

Compatibility with the metric *uniquely* determines torsion-free parallel transport.

22.4 The Curvature (Riemann) Tensor

Let us consider an infinitesimal parallelogram C in the manifold, joining the points: $\varphi, \varphi+\epsilon_1, \varphi+\epsilon_1+\epsilon_2, \varphi+\epsilon_2$, and back to φ (see figure 22.1), and calculate the corresponding parallel transporter:

$$\begin{aligned} \mathbf{U}(C) = {} & \mathbf{U}^{-1}(\varphi+\epsilon_2, \varphi) \mathbf{U}^{-1}(\varphi+\epsilon_1+\epsilon_2, \varphi+\epsilon_2) \\ & \times \mathbf{U}(\varphi+\epsilon_1+\epsilon_2, \varphi+\epsilon_1) \mathbf{U}(\varphi+\epsilon_1, \varphi). \end{aligned} \tag{22.42}$$

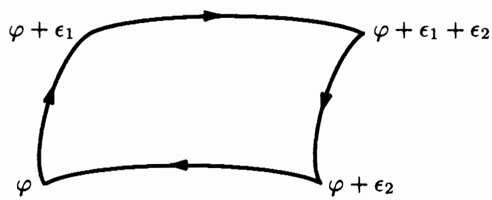


Fig. 22.1 The loop C.

We parametrize the expansion of $\ln \mathbf{U}(C)$ up to second order in ϵ_1 and ϵ_2 as

$$\ln \mathbf{U}(\varphi+\epsilon,\varphi) = -\boldsymbol{\Gamma}_i\epsilon^i + \tfrac{1}{2}\boldsymbol{\Gamma}_{ij}\epsilon^i\epsilon^j + o\left(\epsilon^2\right), \tag{22.43}$$

in which $\boldsymbol{\Gamma}_i$ is the connection.

To evaluate the r.h.s. of equation (22.42), we use repeatedly Baker–Hausdorf's formula (the calculation follows the lines of the derivation of the commutation relations of generators in a Lie algebra, see also Section 34.2.2):

$$\mathrm{e}^A\,\mathrm{e}^B = \mathrm{e}^{(A+B+[A,B]/2+\cdots)}\,.$$

The first order in ϵ in the exponential of the r.h.s. vanishes as a consequence of equation (22.16). At second order, we find

$$\mathbf{U}(C) = \mathbf{1} - \epsilon_1^i\epsilon_2^j\mathbf{R}_{ij} + O\left(\epsilon^3\right),$$

where the antisymmetric tensor $\mathbf{R}_{ij}$ is defined by

$$\mathbf{R}_{ij} = [\nabla_i,\nabla_j] = \partial_i\boldsymbol{\Gamma}_j - \partial_j\boldsymbol{\Gamma}_i + [\boldsymbol{\Gamma}_i,\boldsymbol{\Gamma}_j]\,. \tag{22.44}$$

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 (22.21) to perform the calculation for an arbitrary closed curve. We fix one point φ on the curve and write a generic point $\varphi+\epsilon$. Expanding the path-ordered integral up to second order, we find

$$\mathbf{U}(C) = \mathbf{1} - \oint \boldsymbol{\Gamma}_k(\varphi+\epsilon)\mathrm{d}\epsilon^k + \frac{1}{2}\oint\!\!\oint \mathrm{d}\epsilon^k\mathrm{d}\epsilon^l\,\mathrm{P}[\boldsymbol{\Gamma}_k(\varphi+\epsilon)\boldsymbol{\Gamma}_l(\varphi+\epsilon)] + O\left(\epsilon^3\right).$$

We then expand $\boldsymbol{\Gamma}_k$ for ϵ small. In the first integral, the term proportional to $\boldsymbol{\Gamma}_k(\varphi)$ 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 $\mathbf{R}_{ij}\mathrm{d}\varphi^i\wedge\mathrm{d}\varphi^j$ (see equation (22.29)):

$$\mathbf{R} = 2\nabla^2 = 2\left(\mathrm{d}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^2\right),$$

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:

$$\mathbf{R}_{ij}V^k \equiv R^k_{lij}V^l,$$

Equation (22.44) in component form reads

$$R^k_{lij} = \partial_i\Gamma^k_{lj} - \partial_j\Gamma^k_{li} + \Gamma^k_{mi}\Gamma^m_{lj} - \Gamma^k_{mj}\Gamma^m_{li}\,. \tag{22.45}$$

A general tensor transforms like

$$\mathbf{R}_{ij} V^{k_1 k_2 \dots k_p}_{l_1 l_2 \dots l_q} = \sum_{r=1}^{p} R^{k_r}_{mij} V^{k_1 k_2 \dots m \dots k_p}_{l_1 l_2 \dots l_q} - \sum_{r=1}^{q} R^{m}_{l_r ij} V^{k_1 k_2 \dots k_p}_{l_1 l_2 \dots m \dots l_q} . \tag{22.46}$$

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, $R_{klij} = g_{km} R^m_{lij}$, the expression

$$\begin{aligned} R_{klij} = {} & \tfrac{1}{2} \left(\partial_i \partial_l g_{kj} - \partial_i \partial_k g_{jl} \right) - \tfrac{1}{4} \left(\partial_i g_{mk} + \partial_k g_{mi} - \partial_m g_{ik} \right) \\ & \times g^{mn} \left(\partial_j g_{nl} + \partial_l g_{nj} - \partial_n g_{jl} \right) - (i \leftrightarrow j) . \end{aligned} \tag{22.47}$$

A few properties. Being defined as a commutator, R^k_{lij} satisfies

$$R^k_{lij} = -R^k_{lji} . \tag{22.48}$$

For the same reason, the curvature tensor satisfies the consequence of a Jacobi identity:

$$[\nabla_i, [\nabla_j, \nabla_k]] + \text{cyclic permutations } (i,j,k) = 0 ,$$

which in terms of R^k_{lij} reads

$$\nabla_i R^l_{mjk} + \nabla_j R^l_{mki} + \nabla_k R^l_{mij} = 0 , \tag{22.49}$$

and in the framework of Riemannian geometry is called the *Bianchi identity.*

Note that the compatibility condition (22.39) implies

$$[\nabla_i, \nabla_j] g_{kl} = 0 = R^m_{kij} g_{ml} + R^m_{lij} g_{km} .$$

This equation implies that the tensor R_{klij} is antisymmetric in $k \leftrightarrow l$:

$$R_{klij} = -R_{lkij} . \tag{22.50}$$

a property that can be verified on the explicit form (22.47).

We also note that R_{klij} is antisymmetric in $(k \leftrightarrow l)$ and symmetric in the exchange $(kl) \leftrightarrow (ij)$:

$$R_{klij} = R_{ijkl} . \tag{22.51}$$

It is easy to verify a cyclic identity:

$$R_{ijkl} + R_{kijl} + R_{jkil} = 0 . \tag{22.52}$$

The covariant trace of the curvature tensor R_{ij} is called the *Ricci tensor*:

$$R_{ij} = R^k_{ijk} , \quad R_{ij} = R_{ji} . \tag{22.53}$$

It is a symmetric tensor and thus has, therefore, properties similar to the metric. Taking again the covariant trace one obtains the *scalar curvature*

$$R = R^i_i \equiv R^{ij} g_{ji} \equiv R^{kl}_{kl} . \tag{22.54}$$

Remarks.

(i) An important problem is to classify local tensors which are functionals of the metric tensor. The compatibility condition (22.39) implies that the covariant derivative of g_{ij} is not a new tensor.

However, we observe 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 (22.33) can be rewritten in a simple way that exhibit its tensor character:

$$\delta_\epsilon \Gamma^i_{jk} = \nabla_k \nabla_j \epsilon^i - R^i_{jkl} \epsilon^l . \tag{22.55}$$

The symmetry of the r.h.s. in the exchange $j \leftrightarrow k$ relies on the cyclic identity (22.52).

22.4.1 Holonomy variables, holonomy group

The parallel-transporters $\mathbf{U}(C)$ associated with closed curves C are often called *holonomy variables.* They in particular probe some topological properties of the manifold. This is specially 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 $\mathbf{U}(\varphi; C)$ originating from a common fixed point φ, called the *holonomy group.*

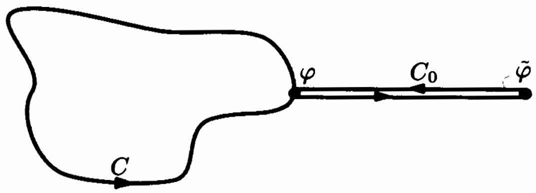


Fig. 22.2

Independence of the holonomy group of the initial point. Let us show that the holonomy groups associated with different initial points are isomorphic. If φ and $\tilde{\varphi}$ are two different initial points, we introduce a fixed curve C_0 joining them. To each curve C passing through φ, we can now associate a curve $\tilde{C}$ passing through $\tilde{\varphi}$ (see figure 22.2):

$$\tilde{C} = C_0^{-1} \cup C \cup C_0 \,. \tag{22.56}$$

In this way, we have constructed a one-to-one mapping between curves passing through φ and curves passing through $\tilde{\varphi}$. The corresponding relation between parallel transporters is

$$\mathbf{U}(\tilde{C}) = \mathbf{U}(C_0)\mathbf{U}(C)\mathbf{U}^{-1}(C_0). \tag{22.57}$$

It establishes an isomorphism between the holonomy groups associated with the points φ and $\tilde{\varphi}$. Therefore, the abstract holonomy group is independent of the initial point and intrinsic to the set of curves on the manifold equipped with the equivalence relation (22.56).

Holonomy and metric. For a general parallel transport, the compatibility condition (22.38) for a closed curve reads

$$g_{ij}(\varphi) = {}^T U_i^k(C) g_{kl}(\varphi) U_j^l(C). \tag{22.58}$$

If the metric is positive, equation (22.58) shows that the matrices $\mathbf{U}(\varphi; C)$ corresponding to closed curves passing through a point φ belong to a subgroup of the $O(N)$ orthogonal group (N is the dimension of the manifold $\mathfrak{M}$). Let us now consider the curve C formed by the infinitesimal parallelogram of figure 22.1. If we expand equation (22.58) up to second order in ϵ, we obtain a condition which, in terms of the curvature tensor, can be written as

$$\epsilon_1^k \epsilon_2^l \left(R_{ijkl} + R_{jikl}\right) = 0 \,. \tag{22.59}$$

This result provides a geometric interpretation for the antisymmetry (22.50) of the curvature tensor: it is the antisymmetry of the generators of the orthogonal group.

22.5 Covariant Volume Element

To construct, for example, a classical action we need a volume element invariant under reparametrization. The euclidean measure transforms like

$$\prod_i \mathrm{d}\varphi^i = \det T_l^k \prod_i \mathrm{d}\varphi'^i . \tag{22.60}$$

Let us denote by g the determinant of the metric g_{ij}:

$$g = \det \mathbf{g} . \tag{22.61}$$

It transforms like

$$g' = \det \mathbf{g}' = \left(\det T_j^i\right)^2 g . \tag{22.62}$$

The volume element

$$\mathrm{d}\rho(\varphi) = \sqrt{g} \prod_i \mathrm{d}\varphi^i , \tag{22.63}$$

is thus invariant. Of course, the volume element $\mathrm{d}\rho(\varphi)$ multiplied by a scalar (like a function of the scalar curvature) is invariant too.

Remark. Differentiating the identity

$$\ln g \equiv \ln \det \mathbf{g} = \operatorname{tr} \ln \mathbf{g} , \tag{22.64}$$

one finds

$$g^{-1} \partial_k g = g^{ij} \partial_k g_{ji} . \tag{22.65}$$

Comparing the r.h.s. with the expression (22.41) for the connection, one infers the simple relation

$$\Gamma_{ki}^k = \frac{1}{\sqrt{g}} \partial_i \sqrt{g} . \tag{22.66}$$

This, in particular, implies for the covariant divergence of a vector

$$\nabla_i V^i = \frac{1}{\sqrt{g}} \partial_i \left(\sqrt{g}\, V^i\right) = \partial_i V^i + (\partial_i \ln \sqrt{g}) V^i , \tag{22.67}$$

and, therefore,

$$\int \left(\prod_i \mathrm{d}\varphi^i\right) \sqrt{g}\, \nabla_j V^j = \int \left(\prod_i \mathrm{d}\varphi^i\right) \partial_j \left(\sqrt{g}\, V^j\right) . \tag{22.68}$$

22.6 Fermions, Vielbein, Spin Connection

We now briefly explain how one can construct spin 1/2 fermions living in Riemannian manifolds, because the construction is non-trivial. 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 $e^{ai}(\varphi)$ which form the local basis is called the *vielbein*. We assume that the metric is positive. We can then choose vectors orthogonal with respect to the metric $g_{ij}(\varphi)$, of length 1:

$$e^{ai}(\varphi)g_{ij}(\varphi)e^{bj}(\varphi) = \delta_{ab}\,. \tag{22.69}$$

We do not distinguish upper and lower internal indices a, b because as equation (22.69) shows the internal metric is euclidean. Introducing the vectors e^a_i, obtained as usual from the e^{ai} by lowering the index with the metric tensor, we can rewrite equation (22.69) as

$$e^{ai}e^b_i = e^a_i e^{bi} = \delta_{ab}\,, \tag{22.70}$$

which shows that the matrix e^{ai} is the inverse of the matrix e^b_i. Finally, combining equations (22.70,22.69) one finds

$$g_{ij}(\varphi) = e^a_i(\varphi)e^a_j(\varphi). \tag{22.71}$$

This equation expresses the metric tensor in terms of the vielbein. As we have done above, in what follows we use the first letters of the alphabet $a, b, c, d, \ldots$ to represent indices corresponding to the euclidean metric and the letters $i, j, k, \ldots$ to represent tensor indices.

The metric is invariant under orthogonal transformations acting on the local frame:

$$e^a_i(\varphi) = O_{ab}(\varphi)\,(e')^b_i\,(\varphi), \quad O_{ab}(\varphi)O_{cb}(\varphi) = \delta_{ac}. \tag{22.72}$$

Spinors then transform under the spin group Spin(N) associated with this local $O(N)$ group (N is the dimension of $\mathfrak{M}$). As described in Section 8.2.2, this implies that the spinors ψ and $\bar\psi$ transform like

$$\begin{aligned}(\psi)_\alpha\,(\varphi) &= \Lambda^{-1}_{\alpha\beta}\psi'_\beta(\varphi),\\ (\bar\psi)_\alpha\,(\varphi) &= \bar\psi'_\beta(\varphi)\Lambda_{\beta\alpha},\end{aligned} \tag{22.73}$$

in which the matrix O_{ab} is given in terms of Λ by

$$\Lambda\gamma_b O_{ab}\Lambda^{-1} = \gamma_a\,. \tag{22.74}$$

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

$$\Lambda = \exp\left(\tfrac{i}{4}\theta_{ab}\sigma_{ab}\right), \tag{22.75}$$

in which θ_{ab} is an antisymmetric real matrix, then,

$$O_{ab} = \left(\mathrm{e}^{\theta}\right)_{ab}\,. \tag{22.76}$$

22.6 Fermions, Vielbein, Spin Connection

We now briefly explain how one can construct spin 1/2 fermions living in Riemannian manifolds, because the construction is non-trivial. 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 $e^{ai}(\varphi)$ which form the local basis is called the *vielbein*. We assume that the metric is positive. We can then choose vectors orthogonal with respect to the metric $g_{ij}(\varphi)$, of length 1:

$$e^{ai}(\varphi)g_{ij}(\varphi)e^{bj}(\varphi) = \delta_{ab}\,. \tag{22.69}$$

We do not distinguish upper and lower internal indices a, b because as equation (22.69) 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 (22.69) as

$$e^{ai}e_i^b = e_i^a e^{bi} = \delta_{ab}\,, \tag{22.70}$$

which shows that the matrix e^{ai} is the inverse of the matrix e_i^b. Finally, combining equations (22.70,22.69) one finds

$$g_{ij}(\varphi) = e_i^a(\varphi)e_j^a(\varphi). \tag{22.71}$$

This equation expresses the metric tensor in terms of the vielbein. As we have done above, in what follows we use the first letters of the alphabet $a, b, c, d, \ldots$ to represent indices corresponding to the euclidean metric and the letters $i, j, k, \ldots$ to represent tensor indices.

The metric is invariant under orthogonal transformations acting on the local frame:

$$e_i^a(\varphi) = O_{ab}(\varphi)\,(e')_i^b\,(\varphi), \quad O_{ab}(\varphi)O_{cb}(\varphi) = \delta_{ac}. \tag{22.72}$$

Spinors then transform under the spin group Spin(N) associated with this local $O(N)$ group (N is the dimension of $\mathfrak{M}$). As described in Section 8.2.2, this implies that the spinors ψ and $\bar\psi$ transform like

$$\begin{aligned}(\psi)_\alpha\,(\varphi) &= \Lambda^{-1}_{\alpha\beta}\psi'_\beta(\varphi),\\ \left(\bar\psi\right)_\alpha(\varphi) &= \bar\psi'_\beta(\varphi)\Lambda_{\beta\alpha},\end{aligned} \tag{22.73}$$

in which the matrix O_{ab} is given in terms of Λ by

$$\Lambda\gamma_b O_{ab}\Lambda^{-1} = \gamma_a\,. \tag{22.74}$$

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

$$\Lambda = \exp\left(\tfrac{i}{4}\theta_{ab}\sigma_{ab}\right), \tag{22.75}$$

in which θ_{ab} is an antisymmetric real matrix, then,

$$O_{ab} = \left(\mathrm{e}^{\theta}\right)_{ab}\,. \tag{22.76}$$

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 (22.72). To enforce this condition, one must introduce 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

$$\mathbf{O}'(C) = \mathbf{O}(\tilde{\varphi})\mathbf{O}(C)\mathbf{O}^{-1}(\varphi), \tag{22.77}$$

in which the curve C goes from φ to $\tilde{\varphi}$. Such transformations are called *gauge* transformations (for details see Chapter 19). 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 $O(N)$:

$$\mathbf{O}(C) = \mathbf{1} - \boldsymbol{\omega}_i(\varphi)\mathrm{d}\varphi^i + o\left(\left\|\mathrm{d}\varphi^i\right\|\right). \tag{22.78}$$

In the case of the gauge theories discussed in Chapter 19, the gauge field is a new independent mathematical quantity. However, here, we notice that if we consider the matrix (which is a scalar under reparametrization)

$$U^{ab}(C) = e_i^a(\tilde{\varphi})U_j^i(C)e^{jb}(\varphi), \tag{22.79}$$

then, as a consequence of the compatibility condition (22.38) and equation (22.71), it satisfies

$${}^T U^{ac}(C)U^{cb}(C) = \delta_{ab}\,, \tag{22.80}$$

and is, therefore, orthogonal. Equation (22.72) implies that it also satisfies equation (22.77). The matrix $U^{ab}(C)$, therefore, has the properties required from $O^{ab}(C)$ and one identifies $O^{ab}(C) \equiv U^{ab}(C)$. Equation (22.79) in the limit of an infinitesimal curve yields a relation between the spin connection ω_i^{ab} and the connection

$$\Gamma_{ik}^j = e^{aj}\partial_k e_i^a + e^{aj}e_i^b\omega_k^{ab}. \tag{22.81}$$

Equation (22.41) 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

$$\omega_i^{ab} = \tfrac{1}{2}e^{aj}\left(\partial_i e_j^b - \partial_j e_i^b\right) + \tfrac{1}{4}e^{aj}e^{bk}\left(\partial_k e_j^c - \partial_j e_k^c\right)e_i^c - (a \leftrightarrow b). \tag{22.82}$$

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). To the connection ω_i^{ab} is associated a covariant derivative ∇_i which on a vector of components

$$V^a = V^i e_i^a\,, \tag{22.83}$$

acts like

$$\nabla_i V^a = \partial_i V^a + \omega_i^{ab}V^b. \tag{22.84}$$

Finally, one defines a general covariant derivative whose action on euclidean indices is given by equation (22.84) and its action on tensor indices by equation (22.31). Let us then calculate, for example,

$$\nabla_i e_j^a = \partial_i e_j^a + \omega_i^{ab}e_j^b - \Gamma_{ij}^k e_k^a\,. \tag{22.85}$$

With this definition, equation (22.81) takes the simple form

$$\nabla_i e_j^a = 0\,, \tag{22.86}$$

and, therefore,

$$\nabla_i V^a = e_j^a \nabla_i V^j. \tag{22.87}$$

This last equation directly follows from the definition of the covariant derivative and equation (22.79).

A last remark: when C is a closed curve, equation (22.79) 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

$$R_{kl}^{ab} = R_{ijkl} e_i^a e_j^b = -R_{kl}^{ba}\,. \tag{22.88}$$

Covariant derivative and fermions. In the case of fermions the covariant derivative $\not{\nabla}$ takes a form which can be inferred from the considerations of Chapter 19, the transformation properties of spinors and the expression of matrices of the spinor representation (22.75):

$$\not{\nabla} = e^{ai}\gamma^a \left(\partial_i + \tfrac{i}{4}\sigma_{bc}\omega_i^{bc}\right). \tag{22.89}$$

22.7 Classical Gravity. Equations of Motion

In the theory of gravity known under the name of General Relativity, the metric tensor field $g_{ij}(\varphi)$ becomes a dynamical variable. Using previous considerations, one can construct a classical action for a metric tensor coupled to matter which is local and purely geometrical, that is, independent of the parametrization of the manifold.

22.7.1 The classical action

The pure gravity action $\mathcal{S}(\mathbf{g})$ 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 action

$$\mathcal{S}(\mathbf{g}) = \int \mathrm{d}^N\varphi\, g^{1/2}(\varphi) R(\mathbf{g}(\varphi)). \tag{22.90}$$

To Einstein's gravity action one can add a derivative-free term, called a cosmological term,

$$\mathcal{S}_{\text{cos.}} = \int \mathrm{d}^N\varphi\, g^{1/2}(\varphi). \tag{22.91}$$

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. Finally, the simplest covariant action for a scalar matter field $\phi(\varphi)$ coupled to gravity takes the form

$$\mathcal{S}_{\text{scalar}} = \int \mathrm{d}^N\varphi\, g^{1/2}(\varphi)\left[\tfrac{1}{2}g^{ij}(\varphi)\partial_i\phi\partial_j\phi + \mathcal{V}(\phi)\right]. \tag{22.92}$$

In the case of matter in the form of a vector field $\mathbf{A}_i$ associated to some external gauge symmetry of the type discussed in Chapters 18,19, we note that the gauge field curvature

$$\mathbf{F}_{ij} = \partial_i\mathbf{A}_j - \partial_j\mathbf{A}_i + [\mathbf{A}_i, \mathbf{A}_j],$$

is a tensor because it is a two-form (see (22.32)). Therefore, a possible action is

$$\mathcal{S}_{\text{gauge}} = \tfrac{1}{4}\int \mathrm{d}^N\varphi\, g^{1/2}(\varphi)\,\mathrm{tr}\,\mathbf{F}_{ij}(\varphi)\mathbf{F}^{ij}(\varphi). \tag{22.93}$$

Fermion matter coupled to gravity. The considerations of Section 22.6, in particular the definitions of the vielbein and the covariant derivative (22.89), also allow us to write an action for Dirac fermion matter coupled to gravity:

$$\mathcal{S}_{\text{fermion}}(\psi,\bar\psi) = -\int \mathrm{d}^N\varphi\, \det\mathbf{e}\,\bar\psi(\varphi)\,(\not{\nabla} + M)\,\psi(\varphi)\,. \tag{22.94}$$

22.7.2 Classical equation of motion

Let us derive the equation of motion for pure gravity, obtained from the variation of the action (22.90):

$$\mathcal{S}_{\text{gravity}} = \int \mathrm{d}^N\varphi\sqrt{g}R\,.$$

Note the identity (be careful about the signs)

$$R^k_{lij} = \nabla_i\Gamma^k_{lj} - \nabla_j\Gamma^k_{li} - \Gamma^k_{mi}\Gamma^m_{lj} + \Gamma^k_{mj}\Gamma^m_{li}\,. \tag{22.95}$$

Using this identity and integrating by parts, we can also write the action

$$\mathcal{S}_{\text{gravity}} = \int \mathrm{d}^N\varphi\sqrt{g}g^{il}\left(-\Gamma^j_{mi}\Gamma^m_{lj} + \Gamma^j_{mj}\Gamma^m_{li}\right)\,.$$

If we add a cosmological term (equation (22.91)), the action becomes

$$\mathcal{S}_{\text{cosmo.}} = \mathcal{S}_{\text{gravity}} + \Lambda\int \mathrm{d}^N\varphi\sqrt{g}\,. \tag{22.96}$$

Equation of motion. We have, therefore, to calculate the variation of various quantities when the metric g_{ij} varies:

$$\delta g_{ij} = h_{ij}\,. \tag{22.97}$$

Note that the variation of the connection is a tensor (the inhomogeneous part of the transformation cancels in the variation). Indeed, one finds after a short calculation,

$$\delta\Gamma^i_{jk} = \tfrac{1}{2}g^{il}\left(\nabla_j h_{lk} + \nabla_k h_{lj} - \nabla_l h_{jk}\right). \tag{22.98}$$

In the same way, one can calculate the variation of the curvature in terms of $\delta\Gamma$. Since $\delta\Gamma$ is a tensor one is not surprised to find

$$\delta R^k_{lij} = \nabla_i\delta\Gamma^k_{jl} - \nabla_j\delta\Gamma^k_{il}.$$

Substituting, we obtain

$$\delta R^k_{lij} = \tfrac{1}{2}g^{km}\left([\nabla_i,\nabla_j]h_{ml} + \nabla_i\nabla_l h_{mj} - \nabla_j\nabla_l h_{mi} - \nabla_i\nabla_m h_{lj} + \nabla_j\nabla_m h_{il}\right)\,. \tag{22.99}$$

Since

$$[\nabla_i, \nabla_j] h_{ml} = -R^n_{mij} h_{nl} - R^n_{lij} h_{nm} \,,$$

we find

$$\begin{aligned}\delta R = \delta\left(g^{li} R^j_{lij}\right) &= -R^{ij} h_{ij} + g^{li} \delta R^j_{lij} \\ &= -R^{ij} h_{ij} + \tfrac{1}{2}\left(g^{il} g^{jm} - g^{im} g^{jl}\right) \nabla_i \nabla_l h_{mj} \,.\end{aligned}$$

Moreover,

$$\begin{aligned}\delta\left(\sqrt{g} R\right) &= \frac{1}{2\sqrt{g}} g^{ij} h_{ij} R + \sqrt{g}\,\delta R \\ &= \frac{1}{2\sqrt{g}} g^{ij} h_{ij} R - \sqrt{g} R^{ij} h_{ij} + \tfrac{1}{2}\left(g^{il} g^{jm} - g^{im} g^{jl}\right) \nabla_i \nabla_l h_{mj} \,.\end{aligned}$$

Integrating by parts in the action and using $\nabla_i g_{jk} = 0$, we finally obtain

$$\delta \mathcal{S} = \int \mathrm{d}^N \varphi \sqrt{g}\, h_{ij} \left(\tfrac{1}{2} g^{ij} R - R^{ij}\right),$$

and thus the equation of motion

$$R g^{ij} - 2R^{ij} = 0 \,. \tag{22.100}$$

Note that in two dimensions, the curvature tensor has only one component and thus $R^{ij} \propto g^{ij}$. Equation (22.100) 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 (see Appendix A22).

The cosmological constant. The action (22.96) with a cosmological term leads to the equation of motion

$$(R + \Lambda) g^{ij} - 2R^{ij} = 0 \,. \tag{22.101}$$

Taking the covariant trace, one finds

$$(N-2)R + N\Lambda = 0 \,,$$

which shows that the cosmological constant induces a non-trivial curvature of space even in the absence of matter.

Matter coupled to gravity. When matter is coupled to gravity, for example, the action terms (22.92) and (22.94) are added to the gravity action, the equation of motion becomes

$$(R + \Lambda) g^{ij} - 2R^{ij} + \mathcal{T}^{ij} = 0 \,, \tag{22.102}$$

where the quantity

$$\mathcal{T}^{ij} = 2 \frac{\delta \mathcal{S}_{\text{matter}}}{\delta g_{ij}} ,$$

coincides in the limit $g_{ij} \to \delta_{ij}$ with the energy–momentum tensor as defined in Appendix A13.3.

22.8 Quantization in the Temporal Gauge: Pure Gravity

It is easy to verify that, due to the covariance of the equations of motion under reparametrization, 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 only generate constraints. This is a problem we have already encountered in gauge theories and we shall use the same strategy 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, we can reduce the metric to the form $g_{00} = 1$ and $g_{0i} = 0$ for $i \neq 0$. If we then specialize the action to such metrics we obtain, as equation of motions, the space components of equations (22.100). The remaining equations have to imposed as additional constraints. As a notation we below use the letters a, b, c, d for space indices.

The action in the temporal gauge. We first calculate the components of the curvature tensor:

$$\begin{aligned} R_{a00d} = R_{0ad0} &= \tfrac{1}{2}\partial_0^2 g_{ad} - \tfrac{1}{4}\partial_0 g_{am} g^{mn} \partial_0 g_{nd} \\ R_{abcd} &= R^{\text{sp.}}_{abcd} - \tfrac{1}{4}\partial_0 g_{ac}\partial_0 g_{bd} + \tfrac{1}{4}\partial_0 g_{bc}\partial_0 g_{ad}\,, \end{aligned}$$

where $R^{\text{sp.}}_{abcd}$ is the curvature tensor in the $N-1$ space. It follows

$$\begin{aligned} R_{00} &= \tfrac{1}{2} g^{ad}\partial_0^2 g_{ad} + \tfrac{1}{4}\partial_0 g^{ab}\partial_0 g_{ab}\,, \\ R &= R_{abcd} g^{ad} g^{bc} + 2R_{00} \\ &= R^{\text{sp.}} + \tfrac{3}{4}\partial_0 g^{ab}\partial_0 g_{ab} + \tfrac{1}{4}\left(g^{ab}\partial_0 g_{ab}\right)^2 + g^{ab}\partial_0^2 g_{ab}\,. \end{aligned}$$

We have to integrate by parts to eliminate the second derivatives.

$$\sqrt{g}\, g^{ab}\partial_0^2 g_{ab} = -\sqrt{g}\left[\partial_0 g_{ab}\partial_0 g^{ab} + \tfrac{1}{2}\left(g^{ab}\partial_0 g_{ab}\right)^2\right] + \text{ total derivatives}\,.$$

We finally obtain

$$\begin{aligned} \mathcal{S}_{\text{temp}} &= \int \mathrm{d}^N\varphi\, \mathcal{L} \\ \mathcal{L} &= \sqrt{g}\left[R^{\text{sp.}} - \tfrac{1}{4}\partial_0 g_{ab}\partial_0 g^{ab} - \tfrac{1}{4}\left(g^{ab}\partial_0 g_{ab}\right)^2\right]. \end{aligned} \tag{22.103}$$

The corresponding conjugate momenta are

$$\Pi^{ab} = -\tfrac{1}{2}\sqrt{g}\left(\partial_0 g^{ab} - g^{ab} g_{cd}\partial_0 g^{cd}\right) = -\frac{1}{2\sqrt{g}}\partial_0\left(g g^{ab}\right) \tag{22.104}$$

and, conversely,

$$\partial_0 g^{ab} = -\frac{2}{\sqrt{g}}\left(\Pi^{ab} - \frac{g^{ab}}{N-2}\Pi^c_c\right). \tag{22.105}$$

The hamiltonian density follows:

$$\begin{aligned} \mathcal{H} &= \Pi^{ab}\partial_0 g_{ab} - \mathcal{L} \\ &= \frac{1}{\sqrt{g}}\left[\Pi_{ab}\Pi^{ab} - \frac{1}{N-2}\left(\Pi^c_c\right)^2\right] - \sqrt{g} R^{\text{sp.}}. \end{aligned} \tag{22.106}$$

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:

$$\mathcal{Z} = \int [\mathrm{d}g_{ij} g^{N(N-5)/8}] \prod_{i=0,N-1} \delta(g_{0i}) \exp[-\mathcal{S}(\mathbf{g})].$$

It is then possible to use the same functional techniques as in gauge theories, and introduce covariant gauges which lead to a BRS symmetric quantized action.

Constraints. We must still study the constraints which are

$$R^{0c} = 0\,, \quad R - 2R^{00} = 0\,. \tag{22.107}$$

The first set is a simple generalization of Gauss's law in gauge theories:

$$R^{0c} = 0 = g^{ad} g^{bc} R_{ab0d} = -\frac{1}{\sqrt{g}} \nabla_a^{\text{sp.}} \Pi^{ac}, \tag{22.108}$$

where $\nabla^{\text{sp.}}$ is the covariant derivative in $N-1$ space dimension. Note that to prove this result one can use

$$\nabla_c^{\text{sp.}} g_{ab} = 0\,. \tag{22.109}$$

These constraints imply 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 (22.9))

$$x^i \mapsto x^i + \epsilon^i(x) \ \Rightarrow h_{ij} = g_{ik} \partial_j \epsilon^k + g_{jk} \partial_i \epsilon^k + \partial_k g_{ij} \epsilon^k = \nabla_i \epsilon_j + \nabla_j \epsilon_i\,. \tag{22.110}$$

Thus, the invariance of the wave functional is expressed by

$$\nabla_b^{\text{sp.}} \frac{\delta \Psi}{\delta g_{ab}} = 0\,, \tag{22.111}$$

which is exactly the constraint. These constraints commute with the hamiltonian since time-independent reparametrization remains a symmetry of the theory in the temporal gauge.

The last constraint is related to time-independent time reparametrization and thus has no equivalent in gauge theories:

$$R - 2R^{00} = R^{\text{sp.}} + \tfrac{1}{4} \partial_0 g^{ab} \partial_0 g_{ab} + \tfrac{1}{4} \left(g^{ab} \partial_0 g_{ab}\right)^2 = 0\,. \tag{22.112}$$

A short calculation shows that

$$R - 2R^{00} = -\frac{1}{\sqrt{g}} \mathcal{H} \ \Rightarrow \ \mathcal{H}\Psi = 0\,. \tag{22.113}$$

Therefore, the constraint, known as Wheeler–DeWitt equation, implies that the physical states corresponds to wave functionals which 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.

Remarks. 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 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, from which we list only a few.

While in general the simplest way to define a quantum field theory, 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 in the case of non-gravitational theories because a change in the signature of the metric is involved.

Lattice regularization of the euclidean theory by simplicial gravity, can be easily achieved in two dimensions (see Appendix A22) but remains a partially unsolved problem in four dimensions.

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 Minkowsky metric $\{+1, -1, -1, -1, \ldots\}$. How nature chooses this particular signature is a non-trivial interesting problem. In perturbation theory pure gravity then describes the self-interaction of a hypothetical spin-two massless particle, called the graviton. From the equations of motions, we see that it has $\frac{1}{2}N(N+1)/2 - 2N = \frac{1}{2}N(N-3)$ dynamical degrees of freedom. In particular, the dimension 3 is for gravity the analogous of dimension 2 for gauge theories: due to reparametrization invariance and constraints, the metric has no dynamical degrees of freedom.

Even 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 that pure gravity is 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 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 commonly believed that the theory, even in its supersymmetric form, remains non-renormalizable, a property which would indicate the breakdown of local quantum field theory at Planck's scale.

Note an additional outstanding problem: vacuum energy contributions of all particles automatically generate a cosmological term in quantum gravity. However, experimentally such a term is bound by an exceedingly small number when expressed in the natural scale of the Planck mass. It has been noticed that a supersymmetric theory has an exact vanishing cosmological constant, but, since experimentally supersymmetry, if it is a symmetry of nature, must be severely broken, this has not provided a solution to the problem.

Finally, the peculiarities of two dimensions are discussed in the appendix.

Bibliographical Notes

For more mathematical details see, for example,

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 University Press, Princeton 1967); S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces* (Academic Press, New York 1978); Y. Choquet-Bruhat, C. DeWitt-Morette and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam 1977).

For a discussion of two-dimensional models on Riemannian manifolds, see

D.H. Friedan, *Ann. Phys. (NY)* 163 (1985) 318.

Finally, for an introduction to the classical theory of gravitation see, for example,

S. Weinberg, *Gravitation and Cosmology* (Wiley, New York 1972).

An elementary introduction to quantum gravity is found in

M. Veltman, in *Methods in Field Theory*, Les Houches School 1975, R. Balian and J. Zinn-Justin eds. (North-Holland, Amsterdam 1976).

Various issues in quantum gravity are discussed by B.S. DeWitt and S.W. Hawking in their contributions to

Relativity, Groups and Topology, proceedings of the 1983 Les Houches school, B.S. DeWitt and R. Stora eds. (Elsevier Science, Amsterdam 1984),

and in the selected papers gathered in

Euclidean Quantum Gravity, G.W.Gibbons and S.W. Hawking eds. (World Scientific, Singapore 1993).

For a discussion of matrix models in the large N-limit and its application to two-dimensional quantum gravity or critical models on random surfaces see, for example,

G. 't Hooft, *Nucl. Phys.* B72 (1974) 461; E. Brézin, C. Itzykson, G. Parisi and J.-B. Zuber, *Comm. Math. Phys.* 59 (1978) 35; D. Bessis, C. Itzykson, and J.-B. Zuber, *Adv. Appl. Math.* 1 (1980) 109, and references therein; E. Brézin and V. Kazakov, *Phys. Lett.* B236 (1990) 144; M. Douglas and S. Shenker, *Nucl. Phys.* B235 (1990) 635; D.J. Gross and A.A. Migdal, *Phys. Rev. Lett.* 64 (1990) 127; M. R. Douglas, *Phys. Lett.* B238 (1990) 176; P. Ginsparg, M. Goulian, M. R. Plesser, and J. Zinn-Justin, *Nucl. Phys.* B342 (1990) 539; E. Brézin, V. Kazakov and Al. A. Zamolodchikov, *Nucl. Phys.* B338 (1990) 673; P. Ginsparg and J. Zinn-Justin, *Phys. Lett.* B240 (1990) 333; D.J. Gross and N. Miljkovic, *Phys. Lett.* B238 (1990) 217; G. Parisi, *Phys. Lett.* B238 (1990) 209.

See also

Random Surfaces and Quantum Gravity, proceedings of the Cargèse workshop 1990, O. Alvarez, E. Marinari and P. Windey eds. (Plenum Press B262, New York 1991),

and the review

P. Di Francesco, P. Ginsparg and J. Zinn-Justin, *Phys. Rep.* 254 (1995) 1.

APPENDIX A22
MATRIX MODELS IN THE LARGE N LIMIT AND 2D GRAVITY

Quantum gravity in four dimensions, beyond the formal level, is far from being well understood. It is, therefore, quite natural to investigate the properties of quantum gravity in lower dimensions. Two dimensions are of special interest because a speculative extension of quantum field theory, string theory, can be formulated as 2D gravity coupled to some matter. In recent years progress has been reported in the problem of 2D gravity coupled to discrete matter, in the euclidean formulation. The solution uses original techniques: space-discretized gravity is reformulated in terms of matrix integrals, and the continuum is recovered in a peculiar "double scaling limit" in which the size N of the matrices becomes large. Since most of these developments are rather technical we here briefly review only the solution of the one-matrix problem, which is relevant for two-dimensional pure quantum gravity. The interested reader is referred to the literature for details. Note that the problem of evaluating integrals over matrices of large size has also statistical applications for fluctuating surfaces, or random hamiltonians or transfer matrices.

A22.1 Quantum 2D Euclidean Gravity

A22.1.1 Classical 2D gravity

Einstein's action for pure gravity in the presence of a cosmological term reads (see equations (22.90,22.91))

$$\mathcal{S}(g)=\int \mathrm{d}^d\varphi\sqrt{g}(KR+\Lambda)\,,$$

in which g_{ij} is the metric tensor, $g=\det g_{ij}$, R the scalar curvature and K,Λ are two coupling constants. The parameter Λ, which multiplies the volume element, is called the cosmological constant. In $d=2$ dimensions classical gravity is trivial because the curvature term does not contribute to the equations of motion (see Section 22.7.2). As a consequence of the Gauss–Bonnet theorem, the total curvature has a topological interpretation: the total curvature of a euclidean closed surface is proportional to the genus of the surface,

$$\int \mathrm{d}^2\varphi\sqrt{g}R=4\pi\chi\,,$$

where χ is the Euler–Poincaré character of the surface, related to the genus, that is, the number of "handles" h, by $h=\frac{1}{2}(2-\chi)$.

2D quantum gravity, however, is somewhat less trivial because large quantum fluctuations may change the genus of the surface. To calculate the partition function it is thus necessary to sum over surfaces of all genera. This problem is difficult to solve starting from the continuum field theory, and surprisingly enough it can be more easily studied in a discretized version.

A22.1.2 Matrix representation of discretized 2D gravity

Following 't Hooft's analysis of the large N limit of $SU(N)$ gauge theories it has been recognized that if one considers the integral

$$\mathcal{Z}(\alpha_k,N)=\int \mathrm{d}M\,\mathrm{e}^{-N\,\mathrm{tr}\,V(M)}\,,$$

in which M is an $N \times N$ hermitian matrix and

$$V(\lambda) = \lambda^2 + \sum_{k \geq 3} \alpha_k \lambda^k ,$$

then, the expansion of $F \equiv \ln \mathcal{Z}$ for N large takes the form

$$F = \ln \mathcal{Z} = \sum N^{\chi} F_{\chi}(\alpha_k) , \qquad (A22.1)$$

where F_χ is the sum of all Feynman diagrams with Euler–Poincaré character $\chi = 2 - 2h$. It follows that the dual of a Feynman diagram contributing to F_χ can be represented as an orientable surface of fixed topology, the powers of $\alpha_3, \alpha_4, \ldots, \alpha_k$ counting the number of triangles, squares,..., k-gons of the surface (taking the logarithm $F \equiv \ln \mathcal{Z}$, yields the gravity model partition function which represents the sum only over connected surfaces). If by convention the area of each triangle, square, ..., is assumed to be 1, the power of α_k measures the total area of the surface. This gives a formulation of discrete 2D gravity in terms of a distribution of random matrices.

In what follows it will be convenient to change the normalization in the integrand and consider instead

$$\mathcal{Z}(g, \alpha_k, N) = \int \mathrm{d}M \; \mathrm{e}^{-(N/g) \operatorname{tr} V(M)} , \qquad (A22.2)$$

where the coupling constant g plays the role of the cosmological constant. Note that to describe pure gravity only one α_k is needed, for instance one can use only triangles. More general models correspond to the addition of some new degrees of freedom on the surface.

The continuum limit. The continuum limit is obtained by letting the total area of the surfaces tend towards infinity. Surfaces of large area are connected with the large order behaviour of the Taylor series expansion of F_χ in powers of g and, therefore, to the singularity of $F_\chi(g)$ closest to the origin. At g fixed, the large N limit selects the surfaces of the topology of the sphere. It will be shown below, however, that it is possible to take $N \to \infty$ while simultaneously taking g to the location of the leading singularity g_c, in a coherent way so that all surfaces of infinite area and arbitrary topology contribute.

Remark. In Chapters 39–42, we shall argue that quite generally integrals of the form (A22.2) are singular at $g = 0$. The result $g_c \neq 0$ reflects the property that the number of planar Feynman diagrams (selected by the large N limit) grows only geometrically instead of a factorial for generic diagrams.

A22.2 The One-Matrix Model

In contrast with vector models where the large N limit can be determined quite generally (see Chapter 30), the large N limit of matrix models has been obtained only in very simple cases, although the problem is of great potential interest (distribution of eigenvalues of random hamiltonians, QCD, critical models on random surfaces, two-dimensional quantum gravity, strings...). Problems for which solutions exist are simple integrals, already a non-trivial problem, over one matrix or several matrices with a one-dimensional nearest neighbour coupling and quantum mechanics of large hermitian matrices (only a partial solution). Most methods of solution are based on the possibility of first diagonalizing the matrices and being able to reduce the matrix integral to an integral over eigenvalues.

In contrast with vector problems even then the problem is not solved because a $N \times N$ matrix has N eigenvalues and the dependence on N is not yet explicit. At this point in the case of integrals two methods of solution have been used, steepest descent and orthogonal polynomials. In one dimension (quantum mechanics), in the sector without unitary excitations, the hamiltonian can be transformed into a hamiltonian of N independent fermions. In this work, we briefly review only the solution of the one-matrix problem, which is relevant for 2D pure quantum gravity.

A22.2.1 The one-matrix model

We want to evaluate for N large the integral over a $N \times N$ hermitian matrix M:

$$\mathcal{Z} = \int \mathrm{d}^{N^2} M \, \mathrm{e}^{-(N/g)\,\mathrm{tr}\, V(M)}, \tag{A22.3}$$

where $\mathrm{d}^{N^2} M$ means integration with a flat measure over the N^2 independent real variables $\mathrm{Re}\, M_{ij}, \mathrm{Im}\, M_{ij}$ and $V(M)$ is a general polynomial potential:

$$V(M) = M^2 + \sum_{k \geq 3} \alpha_k M^k .$$

Since the integrand depends only on the N eigenvalues λ_i of M, we factorize the measure of integration $\mathrm{d}M$ into the measure of the unitary group and a measure for eigenvalues. Integrating over the unitary group one finds

$$\mathcal{Z} = \int \prod_{i=1}^{N} \mathrm{d}\lambda_i \, \Delta^2(\lambda) \, \exp\left[-\sum_i (N/g) V(\lambda_i)\right], \tag{A22.4}$$

where

$$\Delta(\lambda) = \prod_{1 \leq i < j \leq N} (\lambda_j - \lambda_i). \tag{A22.5}$$

$\Delta(\lambda)$ can also be written as a Vandermonde determinant:

$$\Delta(\lambda) = \det \lambda_i^{j-1} ,$$

as one verifies using the antisymmetry of the determinant in the interchange of any two eigenvalues (the normalization is determined by comparing leading terms).

This first step is crucial for all methods used up to now. However, in contrast with the vector models, we still have to integrate over a large number of variables. Thus, additional technical steps are involved in the evaluation of the large N limit.

The eigenvalues measure. Equation (A22.4) may be derived via the usual Faddeev–Popov method. Let U_0 be the unitary matrix such that $M = U_0^\dagger \Lambda U_0$, where Λ is a diagonal matrix with eigenvalues λ_i. The r.h.s. of (A22.4) follows by substituting the definition $1 = \int \mathrm{d}U \, \delta(UMU^\dagger - \Lambda) \, \Delta^2(\lambda)$ (where $\int \mathrm{d}U \equiv 1$). We first perform the integration over M, and then U decouples due to the cyclic invariance of the trace so the integration over U is trivial, leaving only the integral over the eigenvalues λ_i of Λ with flat measure. To determine $\Delta(\lambda)$, we note that only the infinitesimal neighbourhood $U = (1 + T)U_0$ contributes to the U integration, so that

$$1 = \int \mathrm{d}U \, \delta(UMU^\dagger - \Lambda) \, \Delta^2(\lambda) = \int \mathrm{d}T \; \delta([T, \Lambda]) \Delta^2(\lambda) .$$

Now $[T, \Lambda]_{ij} = T_{ij}(\lambda_j - \lambda_i)$, and thus equation ($A$22.5) follows (up to a sign) since we integrate over both real and imaginary parts of the off diagonal T_{ij}'s.

A22.2.2 The large N limit: steepest descent

The partition function ($A22.4$) can be considered as the partition function of a gas of charges of positions λ_i interacting via a repulsive 2D Coulomb force generated by the Vandermonde determinant (see Section 33.3.1) which prevents the charges of accumulating at the minimum of the one-body potential $V(\lambda)$. The saddle point equations, obtained by varying the λ_i's, yield the configuration of minimum energy:

$$\frac{2}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j}=\frac{1}{g}V'(\lambda_i). \qquad (A22.6)$$

This equation can be solved by the following method: we introduce the trace of the resolvent $\omega(z)$ of the matrix M:

$$\omega(z)=\frac{1}{N}\operatorname{tr}\frac{1}{M-z}=\frac{1}{N}\sum_i\frac{1}{\lambda_i-z}. \qquad (A22.7)$$

Multiplying equation ($A22.6$) by $1/(\lambda_i-z)$ and summing over i, we find

$$\omega^2(z)-\frac{1}{N}\omega'(z)+\frac{1}{g}V'(z)\omega(z)=-\frac{1}{g}\sum_i\frac{V'(z)-V'(\lambda_i)}{z-\lambda_i}. \qquad (A22.8)$$

This equation is analogous to the Riccati form of the Schrödinger equation, the wave function ψ being related to ω by $N\omega(z)+NV'(z)/(2g)=\psi'/\psi$. The eigenvalues λ_i are the zeros of the wave function. In the large N limit, we can neglect $N^{-1}\omega'(z)$ (the WKB approximation). In this limit, the distribution of eigenvalues $\rho(\lambda)=\frac{1}{N}\sum_i\delta(\lambda-\lambda_i)$ becomes continuous, and

$$\omega(z)=\int\frac{\rho(\lambda)\mathrm{d}\lambda}{\lambda-z}. \qquad (A22.9)$$

Note that the normalization condition $\int\rho(\lambda')\mathrm{d}\lambda'=1$ is the analogue of the Bohr–Sommerfeld quantization condition. Conversely, the eigenvalue density $\rho(\lambda)$ is extracted from $\omega(z)$ via the relation

$$\rho(\lambda)=\frac{1}{2i\pi}\left(\omega(z+i0)-\omega(z-i0)\right). \qquad (A22.10)$$

In the continuum limit equation ($A22.8$) becomes

$$\omega^2(z)+\frac{1}{g}V'(z)\,\omega(z)+\frac{1}{4g^2}R(z)=0\,, \qquad (A22.11)$$

where

$$R(z)=4g\int\mathrm{d}\lambda\,\rho(\lambda)\frac{V'(z)-V'(\lambda)}{z-\lambda} \qquad (A22.12)$$

is a polynomial of degree $l-2$ when V is of degree l. Note that the coefficient of highest degree of R is fixed by the normalization of $\rho(\lambda)$ while the remaining coefficients depend explicitly on the eigenvalue distribution.

Finally, the free energy $\ln\mathcal{Z}$ is obtained from the saddle point value of the integrand ($A22.4$) in the continuum limit:

$$\ln\mathcal{Z}=N^2\left(\int\mathrm{d}\lambda\,\mathrm{d}\mu\,\rho(\lambda)\rho(\mu)\ln|\lambda-\mu|-\frac{1}{g}\int\mathrm{d}\lambda\,\rho(\lambda)V(\lambda)\right). \qquad (A22.13)$$

A22.2.2 The large N limit: steepest descent

The partition function ($A22.4$) can be considered as the partition function of a gas of charges of positions λ_i interacting via a repulsive 2D Coulomb force generated by the Vandermonde determinant (see Section 33.3.1) which prevents the charges of accumulating at the minimum of the one-body potential $V(\lambda)$. The saddle point equations, obtained by varying the λ_i's, yield the configuration of minimum energy:

$$\frac{2}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j}=\frac{1}{g}V'(\lambda_i). \tag{A22.6}$$

This equation can be solved by the following method: we introduce the trace of the resolvent $\omega(z)$ of the matrix M:

$$\omega(z)=\frac{1}{N}\operatorname{tr}\frac{1}{M-z}=\frac{1}{N}\sum_i\frac{1}{\lambda_i-z}. \tag{A22.7}$$

Multiplying equation ($A22.6$) by $1/(\lambda_i-z)$ and summing over i, we find

$$\omega^2(z)-\frac{1}{N}\omega'(z)+\frac{1}{g}V'(z)\omega(z)=-\frac{1}{g}\sum_i\frac{V'(z)-V'(\lambda_i)}{z-\lambda_i}. \tag{A22.8}$$

This equation is analogous to the Riccati form of the Schrödinger equation, the wave function ψ being related to ω by $N\omega(z)+NV'(z)/(2g)=\psi'/\psi$. The eigenvalues λ_i are the zeros of the wave function. In the large N limit, we can neglect $N^{-1}\omega'(z)$ (the WKB approximation). In this limit, the distribution of eigenvalues $\rho(\lambda)=\frac{1}{N}\sum_i\delta(\lambda-\lambda_i)$ becomes continuous, and

$$\omega(z)=\int\frac{\rho(\lambda)\mathrm{d}\lambda}{\lambda-z}. \tag{A22.9}$$

Note that the normalization condition $\int\rho(\lambda')\mathrm{d}\lambda'=1$ is the analogue of the Bohr–Sommerfeld quantization condition. Conversely, the eigenvalue density $\rho(\lambda)$ is extracted from $\omega(z)$ via the relation

$$\rho(\lambda)=\frac{1}{2i\pi}\left(\omega(z+i0)-\omega(z-i0)\right). \tag{A22.10}$$

In the continuum limit equation ($A22.8$) becomes

$$\omega^2(z)+\frac{1}{g}V'(z)\,\omega(z)+\frac{1}{4g^2}R(z)=0, \tag{A22.11}$$

where

$$R(z)=4g\int\mathrm{d}\lambda\,\rho(\lambda)\frac{V'(z)-V'(\lambda)}{z-\lambda} \tag{A22.12}$$

is a polynomial of degree $l-2$ when V is of degree l. Note that the coefficient of highest degree of R is fixed by the normalization of $\rho(\lambda)$ while the remaining coefficients depend explicitly on the eigenvalue distribution.

Finally, the free energy $\ln\mathcal{Z}$ is obtained from the saddle point value of the integrand ($A22.4$) in the continuum limit:

$$\ln\mathcal{Z}=N^2\left(\int\mathrm{d}\lambda\,\mathrm{d}\mu\,\rho(\lambda)\rho(\mu)\ln|\lambda-\mu|-\frac{1}{g}\int\mathrm{d}\lambda\,\rho(\lambda)V(\lambda)\right). \tag{A22.13}$$

The solution. The solution to equation ($A22.11$) is

$$\omega(z) = \frac{1}{2g}\left(-V'(z) + \sqrt{[V'(z)]^2 - R(z)}\right). \tag{$A22.14$}$$

Generically $\omega(z)$ has $2(l-1)$ branch points corresponding to the roots of the polynomial $V'^2 - R$. Therefore, the support of $\rho(\lambda)$ is formed of $l-1$ disconnected pieces. In the simplest case, when the potential has only one minimum, we expect a single connected support and thus only two branch points. It follows that the polynomial $V'^2 - R$ must have $l-2$ double roots and this yields $l-2$ conditions that fully determine R.

A22.3 The Method of Orthogonal Polynomials

The steepest descent method allows a general discussion of the large N limit. It is difficult, however, to calculate the subleading orders in the $1/N$ expansion and, therefore, to discuss perturbation theory to all orders. We now present another method that allows us to recover previous results and to extend them to all orders in $1/N$.

This alternative method for solving ($A22.4$) makes use of a set of polynomials $P_n(\lambda)$, orthogonal with respect to the measure

$$\int_{-\infty}^{\infty} \mathrm{d}\lambda\ \mathrm{e}^{-(N/g)V(\lambda)}\ P_n(\lambda)\, P_m(\lambda) = s_n\, \delta_{nm}\,. \tag{$A22.15$}$$

The normalization of the orthogonal polynomials P_n is given by having the leading term $P_n(\lambda) = \lambda^n + \cdots$, hence the constant s_n on the r.h.s. of ($A22.15$). Due to the relation

$$\Delta(\lambda) = \det\, \lambda_i^{j-1} = \det\, P_{j-1}(\lambda_i) \tag{$A22.16$}$$

(recall that arbitrary polynomials may be built up by adding linear combinations of preceding columns, a procedure that leaves the determinant unchanged), the polynomials P_n can be employed to solve ($A22.4$). We substitute the determinant $\det P_{j-1}(\lambda_i) = \sum(-1)^\pi \prod_k P_{i_k-1}(\lambda_k)$ for each of the $\Delta(\lambda)$'s in ($A22.4$) (where the sum is over permutations i_k and $(-1)^\pi$ is the signature of the permutation). The integrals over individual λ_i's factorize, and due to orthogonality the only contributions are from terms with all $P_i(\lambda_j)$'s paired. There are $N!$ such terms so ($A22.4$) reduces to

$$\begin{aligned}\mathcal{Z} &= \int \prod_\ell \mathrm{d}\lambda_\ell\ \mathrm{e}^{-(N/g)V(\lambda_\ell)} \sum_{\pi,\pi'} (-1)^\pi (-1)^{\pi'} \prod_k P_{i_k-1}(\lambda_k) \prod_j P_{i_j-1}(\lambda_j) \\ &= N! \prod_{i=0}^{N-1} s_i = N!\, s_0^N \prod_{k=1}^{N-1} f_k^{N-k}\,,\end{aligned} \tag{$A22.17$}$$

where we have defined $f_k \equiv s_k/s_{k-1}$. The solution of the original matrix integral is thus reduced to the problem of determining the normalizations s_k, or equivalently the ratios f_k.

In the naive large N limit, the rescaled index k/N becomes a continuous variable t that runs from 0 to 1, and f_k/N becomes a continuous function $f(t)$. In this limit, the free energy (up to an irrelevant additive constant) reduces to a simple one-dimensional integral:

$$\frac{1}{N^2}\ln \mathcal{Z} = \frac{1}{N}\sum_k (1-k/N)\ln f_k \sim \int_0^1 \mathrm{d}t\,(1-t)\ln f(t)\,. \tag{$A22.18$}$$

To derive the functional form for $f(t)$, we assume for simplicity that the potential $V(\lambda)$ in (A22.15) is even. Since the P_i's from a complete set of basis vectors in the space of polynomials, it is clear that $\lambda P_n(\lambda)$ must be expressible as a linear combination of lower P_i's, $\lambda P_n(\lambda) = \sum_{i=0}^{n+1} c_{ni} P_i(\lambda)$ (with $c_{ni} = s_i^{-1} \int \mathrm{e}^{-V} \lambda P_n P_i$). However, terms proportional to P_i for $i < n-1$ vanish since $\int \mathrm{e}^{-V} P_n \lambda P_i = 0$ (recall λP_i is a polynomial of order at most $i+1$ so is orthogonal to P_n for $i+1 < n$). Moreover, the term proportional to P_n also vanishes due to the assumption that the potential is even, $\int \mathrm{e}^{-V} \lambda P_n P_n = 0$. Thus, the orthogonal polynomials satisfy the simple recursion relation,

$$\lambda P_n = P_{n+1} + r_n P_{n-1} \tag{A22.19}$$

with r_n a scalar coefficient independent of λ.

By considering the quantity $P_n \lambda P_{n-1}$ with λ paired alternately with the preceding or succeeding polynomial, we derive

$$\int \mathrm{d}\lambda\, \mathrm{e}^{-(N/g)V} P_n \lambda P_{n-1} = r_n s_{n-1} = s_n .$$

This shows that the ratio $f_n = s_n/s_{n-1}$ for this simple case is identically the coefficient defined by equation (A22.19), $f_n = r_n$. Similarly, if we pair the λ in $P_n' \lambda P_n$ before and afterwards, an integration by parts yields

$$\begin{aligned} n s_n = \int \mathrm{d}\lambda\, \mathrm{e}^{-(N/g)V} P_n' \lambda P_n &= \int \mathrm{e}^{-(N/g)V} P_n' r_n P_{n-1} \\ = r_n \int \mathrm{d}\lambda\, \mathrm{e}^{-(N/g)V} (N/g) V' P_n P_{n-1} , \end{aligned} \tag{A22.20}$$

a relation that will allow us to determine recursively r_n.

A22.3.1 The large N limit

Our intent now is to find an expression for $f_n = r_n$ and substitute into (A22.18) to calculate a partition function. For definiteness, we take as an example the potential

$$V(\lambda) = \tfrac{1}{2}\lambda^2 + \tfrac{1}{2}\lambda^4 , \Rightarrow V'(\lambda) = \lambda + 2\lambda^3 . \tag{A22.21}$$

The calculation of the r.h.s. of equation (A22.20) involves then expanding $(\lambda + 2\lambda^3)P_{n-1}$ on the basis of polynomials P_i. These may be accomplished by using repeatedly equation (A22.19). For the potential (A22.21), this gives

$$g(n/N) = r_n + 2r_n(r_{n+1} + r_n + r_{n-1}) . \tag{A22.22}$$

As explained before equation (A22.18), in the large N limit n/N can be treated as a continuous variable $t = n/N$, and we have $r_n \mapsto r(t)$ and $r_{n\pm1} \mapsto r(t \pm \varepsilon)$, where $\varepsilon \equiv 1/N$. To leading order in $1/N$, equation (A22.22) reduces to

$$gt = r + 6r^2 = W(r). \tag{A22.23}$$

It can be verified, inserting the solution of equation (A22.23) into equation (A22.18), that the result obtained by the method of steepest descent is recovered at this order, however, the orthogonal polynomial method allows to calculate much more easily the higher order corrections.

We note that equation (A22.23) has no solution when g is smaller than the minimum value g_c of $W(r)$, where $g_c = -1/24$ for $t = 1$.

A22.3.2 Large N and double scaling limit

Let us construct surfaces in terms of squares. The corresponding potential is then the potential (A22.21). The large N limit, that is, the sum of all contributions of surfaces of genus zero (planar surfaces) can be directly deduced from the results of Section A22.2.2. In particular, we have noted the existence of a critical value $g_c = -1/24$. The limit $g \to g_c$ thus corresponds to the continuum limit. At this point, the second derivative of $\ln \mathcal{Z}$ has a square root singularity and thus the singular part of $\ln \mathcal{Z}$ behaves like $(g - g_c)^{5/2}$.

The all genus partition function. We now search for another solution to equation (A22.22) and its generalizations that describes the contribution of all genus surfaces to the partition function (A22.18). We shall retain higher order terms in $1/N$ in (A22.22) so that, for example, equation (A22.23) instead reads

$$\begin{aligned} gt &= W(r) + 2r(t)\big(r(t+\varepsilon) + r(t-\varepsilon) - 2r(t)\big) \\ &= g_c + \tfrac{1}{2}W''|_{r=r_c}\big(r(t) - r_c\big)^2 + 2r(t)\big(r(t+\varepsilon) + r(t-\varepsilon) - 2r(t)\big) + \cdots \end{aligned} \qquad (A22.24)$$

As suggested at the end of Section A22.3.1, we shall simultaneously let $N \to \infty$ and $g \to g_c$ in a particular way. Since $g - g_c$ has dimension [length]^2, it is convenient to introduce a parameter a with dimension length and let $g - g_c = \kappa^{-4/5}a^2$, with $a \to 0$. Our ansatz for a coherent large N limit will be to take $\varepsilon \equiv 1/N = a^{5/2}$ so that the quantity $\kappa^{-1} = (g - g_c)^{5/4}N$ remains finite as $g \to g_c$ and $N \to \infty$.

Moreover, since the integral (A22.18) is dominated by t near 1 in this limit, it is convenient to change variables from t to z, defined by $g_c - gt = a^2 z$. Our scaling ansatz in this region is $r(t) = r_c + au(z)$. If we substitute these definitions into equation (A22.23), the leading terms are of order a^2 and result in the relation $u^2 \sim z$. To include the higher derivative terms, we note that

$$r(t+\varepsilon) + r(t-\varepsilon) - 2r(t) \sim \varepsilon^2 \frac{\partial^2 r}{\partial t^2} = a \frac{\partial^2}{\partial z^2} au(z) \sim a^2 u'',$$

where we have used $\varepsilon(\partial/\partial t) = -ga^{1/2}(\partial/\partial z)$ (which follows from the above change of variables from t to z). Substituting into (A22.24), the vanishing of the coefficient of a^2 implies the differential equation

$$z = u^2 - \tfrac{1}{3}u'', \qquad (A22.25)$$

(after a suitable rescaling of u and z). The second derivative of the partition function (the "specific heat") has a leading singular behaviour given by $f(t)$ with $t = 1$, and thus by $u(z)$ for $z = (g - g_c)/a^2 = \kappa^{-4/5}$. The solution to equation (A22.25) characterizes the behaviour of the partition function of pure gravity to all orders in the genus expansion. (Notice that the leading term is $u \sim z^{1/2}$ so after two integrations the leading term in $\ln \mathcal{Z}$ is $z^{5/2} = \kappa^{-2}$.)

Equation (A22.25) is known in the mathematical literature as the Painlevé I equation. The perturbative solution in powers of $z^{-5/2} = \kappa^2$ takes the form $u = z^{1/2}(1 - \sum_{k=1} u_k z^{-5k/2})$, where the u_k are all positive. The solution of the Painlevé I equation seems to define 2D quantum gravity beyond the perturbative topological expansion. However, it can be verified that there exists no real solution to the Painlevé equation satisfying the proper boundary conditions. This property can be related to the non-Borel summability of the perturbative expansion. Its physical interpretation is that pure quantum gravity is unstable, because the number of surfaces increases too rapidly with the genus.

23 CRITICAL PHENOMENA: GENERAL CONSIDERATIONS

Chapters 23–38 are devoted to a large extent to the study of second order phase transitions and the determination of their universal properties through the use of renormalization group methods.

Most of the transitions we will consider have the following character: spins on a lattice or particles in the continuum interacts through short range forces. 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 Appendix A4 the system is connected). However, in the infinite volume limit, depending of the value of some control parameter which usually is the temperature, the system either remains ergodic, or experiences a breaking of ergodicity. In the latter case, phase space decomposes in 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 possible values of the spontaneous magnetization.

In classical lattice models with finite range interactions, a transfer matrix can be defined. As already discussed in Section 2.4, the thermodynamic limit is related to the largest eigenvalue of the matrix. In the "ergodic" phase the corresponding eigenvector is unique while in the non-ergodic phase it is degenerate.

Our goal is to analyse the behaviour of thermodynamic quantities in the neighbourhood of the phase transition, in particular, their singularities as function of the control parameter.

For the simplest systems, it is possible to find local observables whose values depend on the phase in the several phase region. One calls such an observable an *order parameter*. It is, for example, the spin in ferromagnetic systems.

In the one phase region, for systems with short range interactions (e.g. decreasing exponentially or faster) the connected correlation functions decrease exponentially with distance when two non-empty sets of points are separated (this is directly related to the cluster property of Section 7.4.2). We call *correlation length* the inverse of the smallest decay rate of correlation functions (the smallest physical mass in the particle physics sense). Because we are interested only in long distance properties, we study only these special phase transitions for which the correlation length, measured in units of the microscopic scale (range of forces, lattice spacing), diverges at the transition point. For such systems it will be shown that, near the transition point (e.g. the critical temperature), some properties of thermodynamic functions are *universal* and furthermore can be described by euclidean field theories and renormalization group equations.

Several chapters will be devoted to the derivation of universal properties of correlation functions, but let us here emphasize already the deep connection between non-mean field universality in phase transitions with divergent correlation length, and renormalization in local field theories. (Locality of field theory is the analogue, as we shall see later, of short range forces.) We have shown that renormalizable local field theories are short distance insensitive, in the sense that, although they are not finite in the limit in which the UV cut-off becomes much larger than all masses and momenta, after an infinite renormalization a unique (up to finite renormalizations) renormalized field is obtained. The translation into the phase transition language is simple: universality emerges in a regime in which the correlation length and all distances are much larger than the microscopic scale which plays the role of the inverse of the cut-off. However, this universality is non-mean field like (at least in low dimensions) because the microscopic scale cannot be completely eliminated. Degrees of freedom on all scales remain coupled, and universal correlations cannot be calculated from a gaussian measure.

In particular, we will show that the particular dimension in which the field theory is just renormalizable separates higher dimensions where phase transitions are quasi-gaussian from lower dimensions where the gaussian approximation is no longer valid.

Since renormalization group equations are a direct consequence of renormalization theory, as we have shown in Chapter 10, we shall not be too surprised to discover that they appear and play a central role in the theory of phase transitions.

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. This language is, therefore, in a sense, almost as abstract as the language of quantum field theory. Since we shall eventually show that all the universal problems in the theory of phase transitions can be formulated in the language of euclidean field theory, we could have stayed completely within the framework developed in the preceding chapters. However, since we also wish to introduce field theory methods to readers with a background in statistical mechanics, we shall often use the statistical and magnetic language. In preceding chapters, we have already called correlation functions what often is called imaginary time Green's functions or Schwinger's functions. A short glossary is then:

Euclidean field theory	**Classical magnetic systems**
Fields	*Spin variables*
Source	*Magnetic field*
Euclidean action	*Configuration energy, hamiltonian*
exp $(-\Delta t\, H)$	*Transfer matrix in the case of finite range interactions*
Functional integral	*Sum over spin configurations*
Field vacuum expectation value	*Magnetization*
Zero momentum two-point function	*Magnetic susceptibility*
Physical mass	*Inverse correlation length*
Massless theory	*Critical theory*
Generating functional $\mathcal{Z}$ *of correlation functions*	*Partition function in a space-dependent magnetic field*
Generating functional $\mathcal{W}$ *of connected correlation functions*	*Free energy in a magnetic field*
Generating functional Γ *of proper vertices*	*Thermodynamic potential, function of the magnetization*

We now briefly explain the organization of the chapters devoted to critical phenomena:

In Chapter 24, we study ferromagnetic systems within mean field theory and show how in dimension less than or equal to 4, summation of leading divergent corrections to mean field theory leads to the continuum ϕ^4 field theory. In Chapters 25–27, we then introduce renormalization group ideas, and show that, combined with field theory techniques, they lead to a complete description of universal properties of second order phase transitions in ferromagnetic systems near four dimensions. We also evaluate the leading corrections to the universal behaviour in the critical domain.

In Chapter 28, we show how the same ideas also apply to non-magnetic systems, and in Chapter 29, we compare RG numerical results with the information available coming from experiments, and numerical investigations of lattice models.

In Chapters 30 and 31, we expand the N-vector model for N large and at low temperature and verify RG predictions. The special roles of two dimensions and the abelian $O(2)$ model emerge. Chapters 32 and 33 are thus devoted to a technique specific to two dimensions, bosonization, which gives information about the properties of several models like the sine-Gordon, Coulomb gas, Thirring or Schwinger models, and its application to the famous Kosterlitz–Thouless phase transition.

In Chapters 34 and 35, we try to gain some non-perturbative insight in the properties of gauge theories through the use of lattice regularization and RG techniques. We derive the conditions for large momentum asymptotic freedom in four dimensions. In Chapter 36, we examine Critical Dynamics, that is, the time evolution of critical systems while Chapter 37 deals with finite size effects. Finally, in Chapter 38, we discuss finite temperature quantum field theory.

However, before we discuss phase transitions with more sophisticated techniques, we want to recall a few properties, from the point of view of phase transitions, of simple ferromagnetic lattice models. In systems with finite range interactions a transfer matrix can be defined. We first examine its properties in a finite volume. In the infinite volume limit, low and high temperature considerations provide convincing evidence of the existence of phase transitions in Ising-like systems. We relate the notion of order parameter to cluster properties in the low temperature broken phase. We show in a simple example that phase transitions indeed correspond to breaking of ergodicity. We finally extend the analysis to ferromagnetic systems with continuous symmetries.

The appendix contains a brief discussion of quenched disorder.

23.1 Phase Transitions and Transfer Matrix

In Section 2.4, we have already introduced the transfer matrix associated with a class of one-dimensional classical systems with nearest neighbour interactions, relating it to the quantum statistical operator $\mathrm{e}^{-\varepsilon H}$ for small euclidean time steps ε. This has enabled us to write the partition function $\mathcal{Z}$ with periodic boundary conditions as

$$\mathcal{Z} = \operatorname{tr} \mathbf{T}^{l}, \tag{23.1}$$

where l is a measure of the lattice size in the euclidean time direction and $\mathbf{T}$ the transfer matrix.

Another simple one-dimensional example is provided by the Ising model with nearest neighbour (n.n.) interactions, at temperature $1/\beta$ and in a magnetic field h. The partition function

$$\mathcal{Z} = \sum_{\{S_i=\pm 1\}} \exp\left[\beta \sum_{i=1}^{l} (J S_i S_{i+1} + h S_i)\right] \tag{23.2}$$

can be expressed in terms of the transfer matrix which is a 2×2 matrix:

$$\langle S' | \mathbf{T} | S \rangle = \exp\left[\beta \left(J S S' + \frac{h}{2}(S + S')\right)\right], \tag{23.3}$$

where the bra–ket notation will be used in this chapter to represent matrix elements.

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 corresponds to euclidean time.

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).

Note the difference between the roles played by the inverse temperature β 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 often quantum fluctuations are irrelevant. One exception is provided by zero temperature quantum statistical systems in $d-1$ dimensions which share many properties with classical statistical systems in d dimensions. Moreover, from the discussion of Section 2.4, we understand that the infinite β quantum correlation functions 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.

Simple 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 and adapting the arguments of Appendix A4, we can derive an important property of its spectrum. We give 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 symmetric, its eigenvalues are real. Let λ_0 be the eigenvalue with largest modulus and $|0\rangle$ the corresponding eigenvector:

$$\mathbf{T}\,|0\rangle = \lambda_0\,|0\rangle\,. \tag{23.4}$$

We call T_{ab} the matrix elements of the transfer matrix and v_a^0 the components of the eigenvector $|0\rangle$ and rewrite equation (23.4) in component form:

$$T_{ab}v_b^0 = \lambda_0 v_a^0\,. \tag{23.5}$$

Taking the scalar product with the vector $|0\rangle$, we derive

$$\lambda_0 = \frac{v_a^0 T_{ab} v_b^0}{v_a^0 v_a^0}\,. \tag{23.6}$$

The denominator of the r.h.s. is a sum of positive terms and the matrix elements of $\mathbf{T}$ are positive. Taking the modulus of equation (23.6) we obtain the inequality:

$$|\lambda_0| \le \frac{\left|v_a^0\right| T_{ab} \left|v_b^0\right|}{\left|v_a^0\right| \left|v_a^0\right|}. \tag{23.7}$$

The r.h.s. of the inequality is the average of $\mathbf{T}$ taken with the vector $\left|v_a^0\right|$. 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 non-negative. Then, according to equation (23.5), λ_0 is positive and all components of $|0\rangle$ are strictly positive. It follows that 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

$$\mathcal{W} = \ln \mathcal{Z} \sim l \ln \lambda_0 \,. \tag{23.8}$$

Note that in classical systems, we generally omit the temperature factor in front of the free energy, except when discussing some low temperature properties.

Since no crossing of levels can occur, λ_0 is a regular function of the inverse temperature β. For example, for the Ising model, one finds

$$\lambda_0 = e^{\beta J} \cosh \beta h + \left(e^{2\beta J} \sinh^2 \beta h + e^{-2\beta J}\right)^{1/2} . \tag{23.9}$$

The connected two-point spin correlation function behaves at large time separation as

$$\langle S_i S_j \rangle_c = \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \sim \left(\langle 0 \,|\mathbf{S}|\, 1\rangle\right)^2 e^{-|i-j|/\xi}, \tag{23.10}$$

in which $|1\rangle$ is the eigenvector corresponding to the second largest eigenvalue in modulus λ_1, which we have here assumed to be positive, and ξ is given by

$$\xi^{-1} = \ln(\lambda_0/\lambda_1). \tag{23.11}$$

Finally, the matrix $\mathbf{S}$ is a matrix diagonal in the spin representation which corresponds to a multiplication by the spin variable. We have assumed that its matrix element $\langle 0 \,|\mathbf{S}|\, 1\rangle$ does not vanish. Again in the example of the Ising model, one finds

$$\xi^{-1} = 2 \tanh^{-1} R \,, \quad R = \frac{\left(\sinh^2 \beta h + e^{-4\beta J}\right)^{1/2}}{\cosh \beta h} . \tag{23.12}$$

We conclude from this analysis that 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.

23.2 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 Ising model with nearest neighbour (n.n.) interactions. The partition function can be written as

$$\mathcal{Z} = \sum_{\{S_\mathbf{r} = \pm 1\}} \exp\Bigl(\beta J \sum_{\substack{\text{n.n.} \\ \mathbf{r},\mathbf{r}' \in \mathbb{Z}^d}} S_\mathbf{r} S_{\mathbf{r}'}\Bigr). \tag{23.13}$$

The matrix elements of the transfer matrix now take the form

$$\langle\{S'_\rho\}|\mathbf{T}|\{S_\rho\}\rangle = \exp\left[\beta J\left(\sum_{\rho\in\mathbb{Z}^{d-1}} S_\rho S'_\rho + \frac{1}{2}\sum_{\substack{\text{n.n.}\\ \rho,\rho'\in\mathbb{Z}^{d-1}}} (S_\rho S_{\rho'} + S'_\rho S'_{\rho'})\right)\right]. \tag{23.14}$$

The Ising model is characterized by a discrete symmetry corresponding to spin reversal. We call $\mathbf{P}$ the corresponding matrix

$$\mathbf{P}\,|\{S_\rho\}\rangle = |\{-S_\rho\}\rangle\,. \tag{23.15}$$

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 (23.15) 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 $+1$,

$$\mathbf{P}\,|0\rangle = |0\rangle\,. \tag{23.16}$$

We now examine the infinite L limit at low and high temperature, that is, at high and low β.

High temperature. At high temperature ($\beta \to 0$), 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 23.1, 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 23.1 no longer apply. Note that $|+\rangle$ and $|-\rangle$ are not eigenvectors of $\mathbf{P}$ since

$$\mathbf{P}\,|+\rangle = |-\rangle\,. \tag{23.17}$$

At low but finite temperature, the eigenvectors can no longer be exactly $|+\rangle$ and $|-\rangle$ but have also necessarily components on configurations in which a finite number of spins have been reversed, as low temperature perturbation theory shows. However, this does not forbid 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 δ:

$$\delta = \frac{\langle +|\mathbf{T}|-\rangle}{\langle +|\mathbf{T}|+\rangle}. \tag{23.18}$$

At low temperature, this ratio is related to the difference of the minimal energy of the configurations corresponding to the two different boundary conditions (b.c.). The cost in energy of imposing "twisted" b.c., that is, spins up on one side and spins down on the other side, is proportional to the area of the surface on which the spins are reversed. Therefore, δ behaves like

$$\delta \propto \mathrm{e}^{-\beta J L^{d-1}} . \qquad (23.19)$$

The eigenvectors and eigenvalues are qualitatively given by diagonalizing a 2×2 matrix τ in the $\{|+\rangle, |-\rangle\}$ subspace:

$$\tau = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} . \qquad (23.20)$$

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

$$\xi_L \propto \mathrm{e}^{\beta J L^{d-1}} . \qquad (23.21)$$

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. We argue in Section 23.4 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 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 will be studied in detail in Chapter 37.

(ii) We have analysed here a lattice model, but we will show in Section 41.1 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,

$$\beta \longmapsto \hbar^{-1}, \quad J \longmapsto \text{instanton action},$$

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:

$$r = \lim_{l \to \infty} \frac{\operatorname{tr} \mathbf{P}\, \mathbf{T}^l}{\operatorname{tr} \mathbf{T}^l} , \qquad (23.22)$$

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.

On the contrary, 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,

$$r \propto \mathrm{e}^{-\beta J L^{d-1}} .$$

the other side, is proportional to the area of the surface on which the spins are reversed. Therefore, δ behaves like

$$\delta \propto \mathrm{e}^{-\beta J L^{d-1}} . \tag{23.19}$$

The eigenvectors and eigenvalues are qualitatively given by diagonalizing a 2×2 matrix τ in the $\{|+\rangle, |-\rangle\}$ subspace:

$$\tau = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} . \tag{23.20}$$

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

$$\xi_L \propto \mathrm{e}^{\beta J L^{d-1}} . \tag{23.21}$$

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. We argue in Section 23.4 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 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 will be studied in detail in Chapter 37.

(ii) We have analysed here a lattice model, but we will show in Section 41.1 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,

$$\beta \longmapsto \hbar^{-1}, \quad J \longmapsto \text{instanton action},$$

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:

$$r = \lim_{l \to \infty} \frac{\operatorname{tr} \mathbf{P}\, \mathbf{T}^l}{\operatorname{tr} \mathbf{T}^l} , \tag{23.22}$$

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.

On the contrary, 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,

$$r \propto \mathrm{e}^{-\beta J L^{d-1}} .$$

One advantage of the criterion involving the ratio (23.22) is that 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 b.c.: the denominator corresponds to periodic b.c. in the time direction, the numerator corresponds to twisted b.c., that is, the configurations on the two boundaries differ by a transformation of the symmetry group. In the case of the Ising model, twisted b.c. are antiperiodic b.c.. 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 23.4, 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.

23.3 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. This sensitivity to boundary conditions is another characteristic of the several phase region.

It is useful in this situation to examine the cluster properties of correlation functions.

We consider the two phase region of an Ising-like system and denote by $|+\rangle$, $|-\rangle$ the two ground states which are exchanged by $\mathbf{P}$ (equation (23.17)) and orthogonal. Any linear combination $|\alpha\rangle$ of the form

$$|\alpha\rangle = \cos\alpha\,|+\rangle + \sin\alpha\,|-\rangle \tag{23.23}$$

is also an eigenvector of the transfer matrix with the same eigenvalue. Let us assume that the boundary conditions select the vector $|\alpha\rangle$. Correlation functions are then obtained by calculating averages of matrices in the state $|\alpha\rangle$.

It is intuitive that the spin S_σ at site σ is, according to the definition given in the introduction, an order parameter. The corresponding matrix $\mathbf{S}$ (we omit the index σ), whose matrix elements between two spin configurations are

$$\langle\{S'_\rho\}|\,\mathbf{S}\,|\{S_\rho\}\rangle = S_\sigma \prod_\rho \delta_{S_\rho S'_\rho}$$

is odd under reflection:

$$\mathbf{P}^{-1}\mathbf{S}\mathbf{P} = -\mathbf{S}\,. \tag{23.24}$$

In the high temperature phase, in which the vacuum state satisfies

$$\mathbf{P}\,|0\rangle = |0\rangle\,, \tag{23.25}$$

then,

$$\begin{aligned}\langle 0\,|\mathbf{S}|\,0\rangle &= -\,\langle 0\,|\mathbf{P}^{-1}\mathbf{S}\mathbf{P}|\,0\rangle \\ &= -\,\langle 0\,|\mathbf{S}|\,0\rangle = 0\,.\end{aligned} \tag{23.26}$$

In the low temperature phase, we define

$$\langle +|\,\mathbf{S}\,|+\rangle = m > 0\,, \tag{23.27}$$

in which m approaches 1 in the Ising model in the zero temperature limit. Then,

$$\langle - |\mathbf{S}| - \rangle = \langle + \left|\mathbf{P}^{-1}\mathbf{S}\mathbf{P}\right| + \rangle = -m\,. \tag{23.28}$$

We also have

$$\langle + |\mathbf{S}| - \rangle = \langle - \left|\mathbf{P}^{-1}\mathbf{S}\mathbf{P}\right| + \rangle = -\langle - |\mathbf{S}| + \rangle\,. \tag{23.29}$$

Since $\mathbf{S}$ is a symmetric matrix,

$$\langle + |\mathbf{S}| - \rangle = 0\,. \tag{23.30}$$

The essential property used below is that $\mathbf{S}$, restricted to the $\{|+\rangle\,, |-\rangle\}$ subspace, is diagonal in the $\{|+\rangle\,, |-\rangle\}$ basis.

It follows from equations (23.27–23.30), in particular, that

$$\langle \alpha |\mathbf{S}| \alpha \rangle = m \cos 2\alpha\,, \tag{23.31}$$

$$\langle \alpha |\mathbf{S}| \pi/2 + \alpha \rangle = m \sin 2\alpha\,. \tag{23.32}$$

Except for $\alpha = \pi/4 \pmod{\pi/2}$, the average of the spin does not vanish and this characterizes the several phase region.

Let us calculate what one would naively expect to be the connected two-point correlation function of the spin, that is, the two-point function of $\mathbf{S} - < \mathbf{S} >$, for two points separated by a distance l in time but at the same position in the transverse direction. It is given by

$$W^{(2)}(l) = \frac{\langle \alpha | \left(\mathbf{S} - m\cos 2\alpha\right) \mathbf{T}^{l} \left(\mathbf{S} - m\cos 2\alpha\right) |\alpha\rangle}{\langle \alpha |\mathbf{T}^{l}| \alpha \rangle}. \tag{23.33}$$

The transfer matrix $\mathbf{T}$ projects onto the ground states for large l:

$$\begin{aligned} \mathbf{T}^{l} &= \lambda_0^{l} \left[\left(|+\rangle\langle+| + |-\rangle\langle-|\right) + O\left(\mathrm{e}^{-l/\xi}\right)\right], \\ &= \lambda_0^{l} \left[|\alpha\rangle\langle\alpha| + |\pi/2+\alpha\rangle\langle\pi/2+\alpha| + O\left(\mathrm{e}^{-l/\xi}\right)\right]. \end{aligned} \tag{23.34}$$

It then follows that

$$W^{(2)}(l) \sim m^2 \sin^2 2\alpha + O\left(\mathrm{e}^{-l/\xi}\right). \tag{23.35}$$

It is only for $\alpha = n\pi$ that the correlation functions satisfy cluster properties. The corresponding eigenvectors are then $|+\rangle$ and $|-\rangle$ which are exchanged by $\mathbf{P}$. According to the definition of Chapter 13, the reflection symmetry is spontaneously broken.

Note that correlation functions calculated by summing over all configurations (the analogues of expressions (2.37,2.39)) correspond to average over the two ground states and do not satisfy cluster properties. This problem can be solved by adding to the configuration energy an infinitesimal interaction term coupled to the order parameter which favours one phase and lifts the degeneracy. For example, in a spin system, correlation functions that satisfy cluster properties in the infinite volume limit can be obtained by the following procedure: at finite volume one adds a constant magnetic field, one takes the infinite volume limit, and then the zero field limit. The ground state, $|+\rangle$ or $|-\rangle$ and low temperature, is chosen by the sign of the magnetic field.

In Section 7.4.2, we have added a small constant part to the sources which allowed us to define the generating functional of connected correlation functions even in the degenerate case.

One may wonder about the physical interpretation of this procedure: in the next section, we argue that a system in the spontaneously broken phase is no longer ergodic. Once prepared in one phase, it remains for ever. Therefore, in the two phase region of a spin system one should only average over configurations which fluctuate around the configurations with all spins up or all spins down.

23.4 Stochastic Processes and Phase Transitions

To construct a stochastic process which converges toward the equilibrium distribution of the Ising model, we can use detailed balance (see Appendix A4) and set for the transition probability p of a configuration $\{S'_{\mathbf{r}}\}$ towards a configuration $\{S'_{\mathbf{r}}\}$, $\mathbf{r}$ belonging now to a lattice Ω of volume L^d in $\mathbb{Z}^d$:

$$\begin{aligned} p(S_{\mathbf{r}}, S'_{\mathbf{r}}) &= e^{-\beta[E(S'_{\mathbf{r}})-E(S_{\mathbf{r}})]} && \text{for } E(S_{\mathbf{r}}) < E(S'_{\mathbf{r}}), \\ p(S_{\mathbf{r}}, S'_{\mathbf{r}}) &= 1\,, && \text{otherwise}, \end{aligned} \tag{23.36}$$

in which $E(S_{\mathbf{r}})$ is the configuration energy:

$$E(S) = \sum_{\substack{\text{n.n.}\\ \mathbf{r},\mathbf{r}'\in\Omega\subset\mathbb{Z}^d}} -JS_{\mathbf{r}}S'_{\mathbf{r}}\,. \tag{23.37}$$

For the argument which 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.

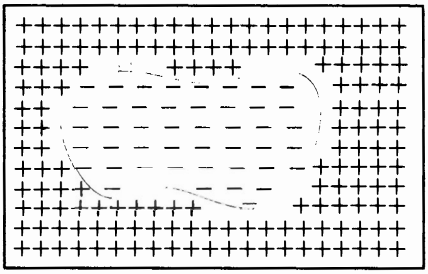


Fig. 23.1

Therefore, at low temperature, if we start from a configuration in which all spins are $+1$, for example, the probability of creating a bubble of minus spins is proportional to $e^{-\beta JA}$, in which A is the area of the bubble surface (figure 23.1). If a large fluctuation creates a bubble which fills half of the total volume, then of course 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 allows the existence of several phases at low temperature 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 Appendix A37.1.3).

23.5 Continuous Symmetries

We now briefly discuss a model which has a continuous symmetry to here exhibit some difference with the case of discrete symmetries. We consider an $O(N)$ symmetric classical spin system, where the spins $\mathbf{S_r}$ are N-component vectors interacting through a pair n.n. ferromagnetic interaction. The model then has a continuous symmetry. The partition function on a d-dimensional lattice is given by

$$\mathcal{Z} = \int \prod_{\mathbf{r}\in\mathbb{Z}^d} \mathrm{d}\mathbf{S_r}\, \delta(\mathbf{S_r}^2 - 1) \exp\left[\sum_{\mathbf{r},\mathbf{r}'\,\mathrm{n.n.}\in\mathbb{Z}^d} \beta\, \mathbf{S_r}\cdot\mathbf{S_{r'}}\right]. \tag{23.38}$$

Such a model has already been proposed as a regularization of the $O(N)$ non-linear σ-model in Chapter 14.

At high temperature ($\beta \to 0$), 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 equivalent of ratio (23.22). The symmetry operation here is a rotation $R(\alpha)$ of angle α. We thus define a partition function $\mathcal{Z}_L(\alpha)$ with twisted b.c. in the euclidean time direction:

$$\mathbf{S}_{t=L,\rho}\cdot\mathbf{S}_{t=0,\rho} = \cos\alpha\,, \qquad \rho\in\mathbb{Z}^{d-1}. \tag{23.39}$$

For convenience, we take periodic b.c. in all other dimensions.

We then examine the behaviour of the ratio

$$r_L = \frac{\mathcal{Z}_L(\alpha)}{\mathcal{Z}_L(0)}, \tag{23.40}$$

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 b.c. occurs between two sites.

The cost in energy ΔE due to the rotation is thus

$$\Delta E = L^{d-1}\times L\times\left[\cos\left(\alpha/L\right) - 1\right] \sim -\alpha^2 L^{d-2}/2\,, \tag{23.41}$$

and, therefore,

$$r_L \propto \exp\left(-\alpha^2\beta L^{d-2}/2\right). \tag{23.42}$$

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 $d \le 2$, 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).

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 $l \neq L$, we obtain instead

$$r \propto \exp\left(-\alpha^2 \beta L^{d-1}/2l\right). \tag{23.43}$$

Hence, using the results of Sections 3.3,3.4, we conclude that the finite size correlation length ξ_L behaves like

$$\xi_L \propto \beta L^{d-1}. \tag{23.44}$$

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

$$\langle S \rangle^2 \propto L^{-2d} \sum_{\mathbf{r},\mathbf{r}'} \langle \mathbf{S}_\mathbf{r} \cdot \mathbf{S}_{\mathbf{r}'} \rangle \propto L^{-2d} \sum_{\mathbf{r},\mathbf{r}'} \exp\left(-|\mathbf{r}-\mathbf{r}'|/\xi_L\right),$$

and, therefore,

$$\langle S \rangle \propto \left(\xi_L/L\right)^d. \tag{23.45}$$

The assumption implies that $\langle S \rangle$ goes to zero in the large volume limit. A phase transition with ordering is only possible if

$$\xi_L \underset{L\to\infty}{\geq} L \quad \Rightarrow d \geq 2\,.$$

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.

These high and low temperature analyses provide information about the existence of a phase transition, and the nature of the phases. However, nothing can be inferred about the phase transition itself, or the behaviour of thermodynamic quantities in the neighbourhood of the critical temperature. These problems will be examined in the coming chapters.

Bibliographical Notes

The proof of spontaneous symmetry breaking of systems with discrete symmetries is given in

R. Peierls, *Phys. Rev.* 54 (1938) 918.

The impossibility of spontaneous breaking of continuous symmetries in one and two dimensions has been proven by

N.D. Mermin and H. Wagner, *Phys. Rev. Lett.* 17 (1966) 1133; S. Coleman, *Commun. Math. Phys.* 31 (1973) 259.

Two relevant textbooks are

S.K. Ma, *Statistical Mechanics* (World Scientific, Singapore 1985);

G. Parisi, *Statistical Field Theory* (Addison-Wesley, New York 1988).

The presentation of this chapter follows

J. Zinn-Justin, *Basko Polje Summer School 1976*, Saclay preprint DPh-T/76-99 unpublished, and in *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, Appendix C, J.-B. Zuber and R. Stora eds. (North-Holland, Amsterdam 1984).

The replica trick has been introduced in

S.F. Edwards and P.W. Anderson, *J. Phys.* F5 (1975) 965.

For the spin glass phase transition see, for instance,

M. Mézard, G. Parisi and M. Virasoro, *Spin Glass and Beyond* (World Scientific, Singapore 1987).

APPENDIX A23
QUENCHED AVERAGES

In this work, we discuss only the so-called annealed averages, that is, averages over random configurations 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 which has to be averaged over disorder, but directly the physical observables, a procedure usually called *quenched average.*

We give here for completeness 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 a local free action:

$$\mathcal{S}(\sigma) = \mathcal{S}_0(\sigma) - \int \mathrm{d}^d x\, H(x)\sigma(x),$$

$$\mathcal{S}_0(\sigma) = \tfrac{1}{2}\int \mathrm{d}^d x \left[(\nabla\sigma(x))^2 + \mu^2\sigma^2(x)\right].$$

For each magnetic field distribution, the partition function is given by

$$\mathcal{Z} = \int [\mathrm{d}\sigma(x)] \exp\left[-\mathcal{S}(\sigma)\right].$$

We assume the magnetic fields at different points uncorrelated with a gaussian distribution characterized by the one- and two-point functions:

$$\overline{H(x)} = h\,, \qquad \overline{H(x)H(y)}_{\mathrm{c}} = w^2\delta^d(x-y),$$

h and w being two constants and $\overline{\bullet}$ meaning quenched average.

In this model, we can immediately calculate the expectation value of the spin variable, at fixed field distribution, at a point x:

$$\langle\sigma(x)\rangle = \int \mathrm{d}^d y\, \Delta(x-y)H(y),$$

where $\Delta(x)$ is the σ-field propagator. If we now average $\langle\sigma(x)\rangle$ over the position x in a large volume Ω, we find

$$\Omega^{-1}\int \mathrm{d}^d x\, \langle\sigma(x)\rangle = \int \mathrm{d}^d x\, \Delta(x)\Omega^{-1}\int \mathrm{d}^d y\, H(y).$$

Because the propagator in the absence of disorder is invariant under translation we now find in the r.h.s. the average of the magnetic field. The result is

$$\Omega^{-1}\int \mathrm{d}^d x\, \langle\sigma(x)\rangle = \frac{1}{\mu^2}\overline{H(x)} = \frac{h}{\mu^2}.$$

We can also calculate the average over the sample of

$$\Omega^{-1}\int \mathrm{d}^d y\,\langle\sigma(y)\rangle\,\langle\sigma(x+y)\rangle = \Omega^{-1}\int \mathrm{d}^d y\,\mathrm{d}^d z\,\mathrm{d}^d t\,\Delta(y-z)\Delta(x+y-t)H(z)H(t)$$
$$= \Omega^{-1}\int \mathrm{d}^d y\,\mathrm{d}^d z\,\Delta(y)\Delta(x+y+z)\int \mathrm{d}^d t\,H(z+t)H(t).$$

Again translation invariance, in the absence of field, reduces the space average to a field average:

$$\Omega^{-1}\int \mathrm{d}^d t\,H(z+t)H(t) = \overline{H(z+t)H(t)} = h^2 + w^2\delta^d(z).$$

It follows that

$$\Omega^{-1}\int \mathrm{d}^d y\,\langle\sigma(y)\rangle\,\langle\sigma(x+y)\rangle = \frac{h^2}{\mu^4} + \frac{w^2}{(2\pi)^d}\int \frac{\mathrm{e}^{ipx}\,\mathrm{d}^d p}{(p^2+\mu^2)^2}\,.$$

Note on the other hand that the σ-field two-point function is independent of the magnetic field:

$$\langle\sigma(y+x)\sigma(y)\rangle = \frac{1}{(2\pi)^d}\int \frac{\mathrm{e}^{ipx}\,\mathrm{d}^d p}{p^2+\mu^2}\,.$$

The replica trick. 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 general situation, of course, so an algebraic trick has been invented to overcome this difficulty. If we want to average correlation functions we have to average the free energy $\mathcal{W} = \ln \mathcal{Z}$ instead of the partition function. Then, the following formal trick is used:

$$\ln \mathcal{Z} = \lim_{n\to 0} \frac{\mathcal{Z}^n - 1}{n}\,.$$

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. In our example, we introduce a field σ_i with n components and find

$$\mathcal{Z}^n(J) = \int \prod_{i=1}^{n}[\mathrm{d}\sigma_i(x)]\exp\left[-\sum_i \mathcal{S}(\sigma_i) + \sum_i \int \mathrm{d}^d x\,J(x)\sigma_i(x)\right].$$

The gaussian disorder average is easy to perform:

$$\overline{\mathcal{Z}^n(J)} = \int [\mathrm{d}\sigma_i(x)]\exp\left[-\sum_i \mathcal{S}_0(\sigma_i)\right.$$
$$\left. +\frac{1}{2}w^2\int \mathrm{d}^d x \sum_{i,j}\sigma_i(x)\sigma_j(x) + \int \mathrm{d}^d x \sum_i \bigl(h+J(x)\bigr)\sigma_i(x)\right].$$

The subtle question is then to take the $n=0$ limit. For instance to calculate the field expectation value, we minimize the potential. We find the replica symmetric solution:

$$\langle\sigma_i(x)\rangle = \frac{h}{\mu^2 - n\omega^2}\,.$$

In this simple example, we set $n = 0$ and obtain the correct result:

$$\overline{\langle\sigma(x)\rangle} = \lim_{n\to 0}\frac{1}{n}\sum_i \sigma_i = \frac{h}{\mu^2}.$$

Notice, however, 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. Therefore, basically the domain of validity of the replica trick is unknown.

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 approximation we discuss in the next chapter, and the mean field solution relies on a breaking of the symmetry between field replica, as shown by Parisi.

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 obviously generalizes with $2n$ replicas.

24 MEAN FIELD THEORY FOR FERROMAGNETIC SYSTEMS

In Chapter 23, 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, we present in this chapter a simple approximation scheme, the mean field approximation, which allows a study of 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 Appendix A24), as the result of variational calculations, or as the leading order in a steepest descent calculation. We have adopted the latter point of view because it allows a systematic calculation of corrections to the mean field approximation and, therefore, a discussion of its validity. The role of space dimension four then emerges.

Moreover, in this framework, it becomes apparent that the mean field approximation is a gaussian or weakly perturbed, *quasi-gaussian*, approximation. In the appendix, we construct the mean field expansion for general lattice models.

Notation. In this chapter, all sums are explicitly indicated.

24.1 Ising-like Ferromagnetic Systems

We consider a classical spin system on a d-dimensional lattice. The spin variable on a lattice site i 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. In this way, the partition function $\mathcal{Z}(H)$ is also the generating functional of spin correlation functions.

We first discuss systems which, in zero field, have an Ising-like $S \mapsto -S$ symmetry. We call β the inverse temperature and $\mathrm{d}\rho(S)$ the normalized distribution of the spin configurations at each site.

The partition function can then be written as

$$\mathcal{Z}(H) = \int \left(\prod_i \mathrm{d}\rho(S_i)\right) \exp\left[-\beta\,\mathcal{H}(S)\right] \tag{24.1}$$

with

$$-\beta\mathcal{H}(S) = \sum_{ij} V_{ij} S_i S_j + \sum_i H_i S_i\,. \tag{24.2}$$

Expressions similar to (24.1) also appear in the construction of the Feynman–Kac representation of partition function in quantum mechanics (e.g. in Chapter 2) or, in quantum field theory, as lattice regularizations of the generating functional of euclidean correlation functions. Note, however, as we recall in Section 24.2, that on the lattice it is possible to calculate the partition function by a high temperature expansion: the site measure $\mathrm{d}\rho(S)$ is kept fixed and one expands in powers of $V_{ij}S_iS_j$, the term which connects different lattice sites. In the perturbative expansion of field theory instead the term $V_{ij}S_iS_j$ is the analogue of the kinetic term while the measure $\mathrm{d}\rho(S)$ contains the interactions. These remarks will be clarified later.

The pair potential. The potential $V_{ij} = V_{ji}$ is translation invariant, $V_{ij} \equiv V(\mathbf{r}_{ij})$, where $\mathbf{r}_{ij}$ is the vector joining the sites i and j. It is then natural to introduce its Fourier transform $\tilde{V}(\mathbf{k})$:

$$\tilde{V}(\mathbf{k}) = \sum_{\mathbf{r}} V(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}), \tag{24.3}$$

where the momentum $\mathbf{k}$ has components that can be restricted to the Brillouin zone $|k_\mu| \le \pi$. 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 $\mathbf{k} = \mathbf{0}$.

Since the two-spin interaction is ferromagnetic, V_{ij} is non-negative. The positivity of the pair potential implies

$$|\tilde{V}(\mathbf{k})| \le \tilde{V}(0) = \sum_{\mathbf{r}} V(\mathbf{r}) \equiv v\,, \tag{24.4}$$

where the parameter v is proportional to β, the inverse temperature. We assume, for simplicity, that the potential has the hypercubic symmetry of the lattice (this requirement is convenient but only reflection symmetry is essential since one can always make linear transformations on coordinates). Then, the expansion of $\tilde{V}(\mathbf{k})$ for $\mathbf{k}$ small takes the form

$$\tilde{V}(\mathbf{k}) = v\tilde{U}(\mathbf{k}), \quad \tilde{U}(\mathbf{k}) = 1 - a^2k^2 + O\left(k^4\right), \tag{24.5}$$

where k is the length of the vector $\mathbf{k}$, and a a constant microscopic length.

The one-site problem. The solution of the trivial one-site problem allows us to introduce and describe the properties of few useful functions.

The one-site partition function $z(h)$ is the Laplace transform of the spin distribution:

$$z(h) = \int \mathrm{d}\rho(S)\, \mathrm{e}^{Sh}\,. \tag{24.6}$$

The corresponding free energy $A(h)$ is then (note that we omit a factor β in the definition of the free energy which for finite temperature transitions plays no role)

$$A(h) = \ln z(h). \tag{24.7}$$

For the Ising model, where $S_i = \pm 1$, for example, $A(h) = \ln\cosh h$.

The symmetry of the distribution implies that the function $A(h)$ is even and increasing for $h > 0$. Moreover, $A''(h) > 0$ (Schwartz's inequality) and thus $A(h)$ is convex. Finally, we assume that the 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 which implies that $A(h)$ is regular on the real axis and in some neighbourhood of the origin. Moreover,

$$\lim_{|h|\to\infty} A(h)/h^2 = 0\,. \tag{24.8}$$

The magnetization is

$$m = \langle S\rangle = A'(h).$$

We also introduce the thermodynamic potential $B(m)$, the Legendre transform of $A(h)$,

$$B(m) + A(h) = hm\,, \quad m = A'(h)\,. \tag{24.9}$$

In the example of the Ising model, one finds

$$B(m) = \tfrac{1}{2}(1+m)\ln(1+m) + \tfrac{1}{2}(1-m)\ln(1-m). \tag{24.10}$$

The Legendre transformation implies the relation

$$B''(m) = 1/A''(h). \tag{24.11}$$

Since $A(h)$ is a convex function of h, $B(m)$ is also convex. It is analytic for m small and can thus be parametrized as

$$B(m) = \sum_{p=1} \frac{b_{2p}}{2p!}\, m^{2p}, \quad b_2 > 0\,. \tag{24.12}$$

In the Ising model, one finds

$$b_2 = 1\,, \qquad b_4 = 2\,.$$

24.2 High Temperature Expansion

The simplest method to calculate $\mathcal{Z}(H)$ is to expand expression (24.1) 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:

$$\mathcal{Z}_0(H) = \prod_i z(H_i). \tag{24.13}$$

Higher order terms in the high temperature expansion,

$$\mathcal{Z}(H)/\mathcal{Z}_0(H) = 1 + \sum_{ij} V_{ij}\,\langle S_i S_j\rangle_0 + \frac{1}{2!}\sum_{ijkl} V_{ij}V_{kl}\,\langle S_i S_j S_k S_l\rangle_0 + \cdots\,, \tag{24.14}$$

can then all be expressed in terms of the moments of the one-site spin distribution in a magnetic field:

$$\langle S_i^n\rangle_0 = z^{-1}(H_i)\left(\frac{\partial}{\partial H_i}\right)^n z(H_i). \tag{24.15}$$

Let, for example, calculate the corresponding Weiss free energy $\mathcal{W}(H) = \ln \mathcal{Z}(H)$ at order β with the assumption $V_{ii} = 0$:

$$\mathcal{W}(H) = \sum_i A(H_i) + \sum_{ij} V_{ij} \langle S_i \rangle_0 \langle S_j \rangle_0 + O(\beta^2)$$

with the definition (24.7). The local magnetization follows

$$M_i = \langle S_i \rangle = \frac{\partial \mathcal{W}}{\partial H_i} = A'(H_i) + 2\sum_j V_{ij} \left(\langle S_i^2 \rangle_0 - (\langle S_i \rangle_0)^2 \right) \langle S_j \rangle_0 + O(\beta^2). \qquad (24.16)$$

Finally, the thermodynamic potential $\Gamma(M)$, Legendre transform of $\mathcal{W}(H)$, can be written (using the definition (24.9) and the usual tricks) as

$$\Gamma(M) = \sum_i H_i M_i - \mathcal{W}(H) = \sum_i B(M_i) - \sum_{ij} V_{ij} M_i M_j + O(\beta^2). \qquad (24.17)$$

The high temperature expansion is useful mainly if the leading term, the infinite temperature result, already is qualitatively correct: the expansion gives information about the high temperature phase and necessarily diverges at the critical temperature. Therefore, in the next section, we introduce another approximation scheme, which allows to describe phase transitions. Diagrammatically, it 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 thermodynamical potential for uniform magnetization, as we have argued in Section 7.10.2. 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. This remark will be used in Section 24.5.

24.3 Mean Field Approximation

We now set up an algebraic formalism which allows to obtain mean field results as a leading order approximation and to systematically calculate corrections to mean field theory. It is technically convenient here to assume that V_{ii} (which is independent of the site i) vanishes (see Appendix A24).

24.3.1 Steepest descent method and mean field

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, 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

$$\mathcal{Z}(H) \propto \int \prod_i d\rho(S_i) d\sigma_i d\lambda_i \exp\left[-\beta\, \mathcal{H}(\sigma) + \sum_i \lambda_i (S_i - \sigma_i) \right] . \qquad (24.18)$$

Let, for example, calculate the corresponding Weiss free energy $\mathcal{W}(H) = \ln \mathcal{Z}(H)$ at order β with the assumption $V_{ii} = 0$:

$$\mathcal{W}(H) = \sum_i A(H_i) + \sum_{ij} V_{ij} \langle S_i \rangle_0 \langle S_j \rangle_0 + O(\beta^2)$$

with the definition (24.7). The local magnetization follows

$$M_i = \langle S_i \rangle = \frac{\partial \mathcal{W}}{\partial H_i} = A'(H_i) + 2\sum_j V_{ij} \left(\langle S_i^2 \rangle_0 - (\langle S_i \rangle_0)^2 \right) \langle S_j \rangle_0 + O(\beta^2). \qquad (24.16)$$

Finally, the thermodynamic potential $\Gamma(M)$, Legendre transform of $\mathcal{W}(H)$, can be written (using the definition (24.9) and the usual tricks) as

$$\Gamma(M) = \sum_i H_i M_i - \mathcal{W}(H) = \sum_i B(M_i) - \sum_{ij} V_{ij} M_i M_j + O(\beta^2). \qquad (24.17)$$

The high temperature expansion is useful mainly if the leading term, the infinite temperature result, already is qualitatively correct: the expansion gives information about the high temperature phase and necessarily diverges at the critical temperature. Therefore, in the next section, we introduce another approximation scheme, which allows to describe phase transitions. Diagrammatically, it 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 thermodynamical potential for uniform magnetization, as we have argued in Section 7.10.2. 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. This remark will be used in Section 24.5.

24.3 Mean Field Approximation

We now set up an algebraic formalism which allows to obtain mean field results as a leading order approximation and to systematically calculate corrections to mean field theory. It is technically convenient here to assume that V_{ii} (which is independent of the site i) vanishes (see Appendix A24).

24.3.1 Steepest descent method and mean field

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, 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

$$\mathcal{Z}(H) \propto \int \prod_i d\rho(S_i) d\sigma_i d\lambda_i \exp\left[-\beta\, \mathcal{H}(\sigma) + \sum_i \lambda_i (S_i - \sigma_i) \right]. \qquad (24.18)$$

Integration over the spin variables S_i is then immediate. Introducing the function $A(H)$ defined by (24.7), we find

$$\mathcal{Z}(H) = \int \prod_i \mathrm{d}\sigma_i \mathrm{d}\lambda_i \, \exp\left[-\beta\,\mathcal{H}(\sigma) + \sum_i \left(-\lambda_i\sigma_i + A(\lambda_i)\right)\right]. \tag{24.19}$$

We now evaluate expression (24.19) by the steepest descent method. The saddle point equations are obtained by differentiating the integrand with respect to λ_i and σ_i:

$$\sigma_i = A'(\lambda_i), \tag{24.20a}$$

$$\lambda_i = 2\sum_j V_{ij}\sigma_j + H_i\,. \tag{24.20b}$$

At leading order, we replace σ_i, λ_i in expression (24.19) by the solution of equations (24.20). The mean field free energy is then

$$\mathcal{W}(H) \equiv \ln \mathcal{Z}(H) = \sum_{ij} V_{ij}\sigma_i\sigma_j + \sum_i \left(H_i\sigma_i - \lambda_i\sigma_i + A(\lambda_i)\right) \tag{24.21}$$

$$= -\sum_{ij} V_{ij}\sigma_i\sigma_j + \sum_i A(\lambda_i). \tag{24.22}$$

Since the expression (24.21) is stationary with respect to variations of σ_i and λ_i, the local magnetization M_i in the mean field approximation is given by the explicit derivative with respect to H_i:

$$M_i = \frac{\partial \mathcal{W}}{\partial H_i} = \sigma_i\,. \tag{24.23}$$

Discussion. Comparing equation (24.16) for $V = 0$ with (24.20a), we see that λ_i has the meaning of an effective magnetic field. It is then determined by equation (24.20b) as the sum of the applied external field and a mean field representing the action of the other spins. The interaction which 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 sites on a given site can indeed be replaced by a mean magnetic field (a property reminiscent of the central limit theorem of probability).

24.3.2 Thermodynamic potential and phase transition

From equations (24.21,24.23), we derive the thermodynamic potential $\Gamma(M)$, Legendre transform of $\mathcal{W}(H)$:

$$\Gamma(M) = \sum_i M_i H_i - \mathcal{W}(H) = -\sum_{ij} M_i V_{ij} M_j + \sum_i B(M_i), \tag{24.24}$$

in which $B(M)$ is the function (24.9) (expression which coincides with (24.17) because it is linear in V).

In the case of translation invariant systems (and thus in uniform field), the magnetization is uniform: $M_i = M$. Let us call Ω the number of lattice sites, then,

$$\Omega^{-1}\Gamma(M) = -vM^2 + B(M), \tag{24.25}$$

in which v is the parameter (24.4).

The relation between field and magnetization becomes

$$H = \Omega^{-1} \frac{\partial \Gamma}{\partial M} = -2vM + B'(M). \tag{24.26}$$

In zero magnetic field, the magnetization, that is, the spin expectation value, is given by an extremum of the thermodynamic potential $\Gamma(M)$. Furthermore, since the partition function in zero field is $\exp[-\Gamma(M)]$, the leading saddle points correspond to minima of $\Gamma(M)$.

We now look for the minima of $\Gamma(M)$ when the temperature, and, therefore, also v, vary. The property (24.8) implies that for $|M|$ large enough, $\Gamma(M)$ in equation (24.25) is an increasing function. For v small (high temperature) vM^2 is negligible, and the r.h.s. of equation (24.25) is convex. The minimum of $\Gamma(M)$ is $M = 0$, the magnetization vanishes. When v increases, in general at some value of v other local minima appear which eventually become the absolute minima of $\Gamma(M)$. When this occurs, the value of the magnetization M jumps discontinuously from zero to a finite value corresponding to this new absolute minima. The system undergoes a *first order phase transition*. Figure 24.1 describes the evolution. 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 Field Theory calculations and shall again see explicitly later, is proportional to $\Omega^{1/2}\left[\Gamma''(M)\right]^{-1/2}$ at the minimum, remains finite.

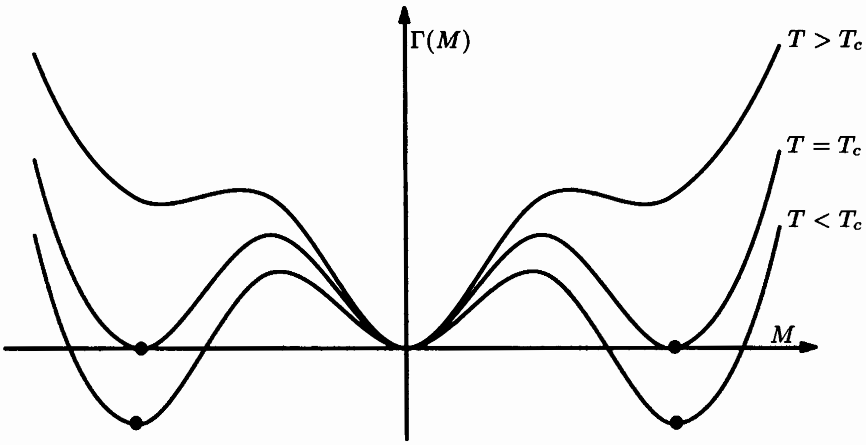


Fig. 24.1 Free energy: first order phase transition.

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, mean field theory 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 temperature T_c corresponding to the value v_c of v,

$$2v_c = B''(0), \tag{24.27}$$

the origin ceases being a minimum of the potential, and below this temperature two minima move continuously away from the origin (see figure 24.2). Since the magnetization

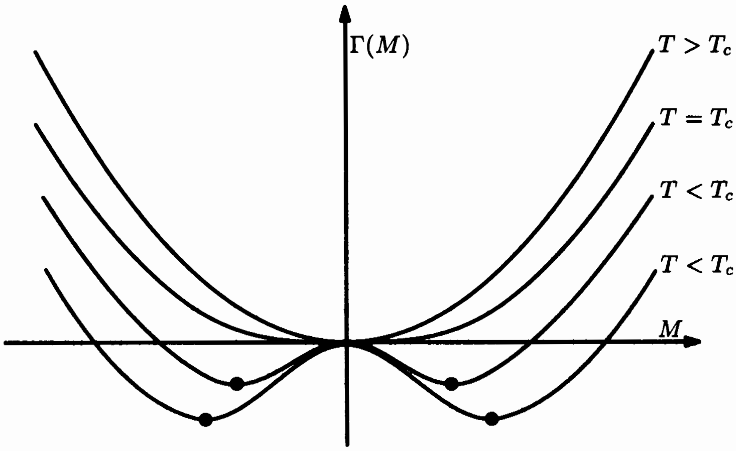


Fig. 24.2 Free energy: second order phase transition.

remains continuous at v_c, the ***phase transition is of second order.*** Because at v_c the second derivative of $\Gamma(M)$ vanishes, the ***correlation length diverges.***

This is the situation we will analyse systematically from now on, first in the framework of the mean field approximation and then by considering corrections coming from higher orders in the mean field expansion.

24.4 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 $B(M)$, is a smooth even function, in a Taylor series in M. We use the parametrization (24.9):

$$\Gamma(M_i) = -\sum_{ij} V_{ij} M_i M_j + \sum_i \left(\frac{b_2}{2!} M_i^2 + \frac{b_4}{4!} M_i^4 + \cdots\right). \tag{24.28}$$

The convexity of $B(M)$ implies that b_2 is positive. The parameter b_4 is also generically positive, because we have assumed that no first order transition occurs at higher temperature. The critical value v_c, in the mean field approximation, is then

$$v_c = b_2/2\,. \tag{24.29}$$

24.4.1 Homogeneous observables

With the parametrization (24.12), the relation (24.26) between field and magnetization becomes

$$H = 2(v_c - v)M + \tfrac{1}{6} b_4 M^3 + O\left(M^5\right). \tag{24.30}$$

In zero field, for $v > v_c$, a spontaneous magnetization M is predicted:

$$M \sim \left[12(v - v_c)/b_4\right]^{1/2} \quad \text{for} \quad |v - v_c| \ll 1\,. \tag{24.31}$$

Near the critical temperature T_c, the magnetization thus has a power behaviour with a mean field or "classical" magnetic exponent β:

$$M \sim (T_c - T)^{\beta}, \qquad \beta = 1/2\,. \tag{24.32}$$

The inverse of the magnetic susceptibility $\chi = \partial M/\partial H$ is the second derivative of $\Gamma(M)$ with respect to M. In zero field,

$$\begin{aligned} \chi_+^{-1} &= 2(v_c - v), \qquad && T > T_c\,, \\ \chi_-^{-1} &= 4(v - v_c), \qquad && T < T_c\,. \end{aligned} \tag{24.33}$$

The magnetic susceptibility, therefore, diverges at T_c with susceptibility exponents γ, γ':

$$\begin{aligned} \chi_+ &\sim C_+(T - T_c)^{-\gamma}, \qquad && \gamma = 1\,, \\ \chi_- &\sim C_-(T_c - T)^{-\gamma'}, \qquad && \gamma' = 1\,, \end{aligned} \tag{24.34}$$

and

$$C_+/C_- = 2\,. \tag{24.35}$$

At T_c ($v = v_c$), for small uniform applied magnetic field H, the equation of state (24.30) becomes

$$H \sim \tfrac{1}{6} b_4 M^3. \tag{24.36}$$

In general, one defines $H \propto M^\delta$, and thus the exponent δ has the mean field value:

$$\delta = 3\,. \tag{24.37}$$

More generally, for v close to v_c, the equation of state (24.30), which is the relation between field, temperature and magnetization, can be cast into a universal scaling form. Rescaling field, temperature and magnetization, we can write

$$H = M^\delta f\big((T - T_c) M^{-1/\beta}\big), \tag{24.38}$$

where the function $f(x)$ is simply

$$f(x) = 1 + x\,, \tag{24.39}$$

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 $\mathcal{C}$:

$$\mathcal{C}(T \to T_{c+}) = 0\,, \qquad \mathcal{C}(T \to T_{c-}) = 12/b_4\,. \tag{24.40}$$

In the mean field approximation, the specific heat has a non-universal jump at T_c.

24.4.2 The two-point function

More generally, the relation between magnetic field and local magnetization can then be written as

$$H_i = \frac{\partial \Gamma}{\partial M_i} = -2\sum_j V_{ij} M_j + B'(M_i). \tag{24.41}$$

Differentiating again with respect to M_i, we obtain the inverse two-point function:

$$W^{(2)}_{ij} = \left(\frac{\partial^2 \Gamma}{\partial M_i \partial M_j}\right)^{-1}\bigg|_{M_i = M} \equiv W^{(2)}(\mathbf{r}_{ij}),$$

whose Fourier transform $\widetilde{W}^{(2)}(\mathbf{k})$ is

$$\widetilde{W}^{(2)}(\mathbf{k}) = \sum_{\mathbf{r}} W^{(2)}(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}) = \left[B''(M) - 2\tilde{V}(\mathbf{k})\right]^{-1}, \tag{24.42}$$

where $\tilde{V}(\mathbf{k})$ is defined by equation (24.3).

One assumes in general that at T_c the two-point function behaves like

$$W^{(2)}(\mathbf{r}) \underset{|\mathbf{r}|\to\infty}{\propto} 1/|\mathbf{r}|^{d-2+\eta} \ \Leftrightarrow\ \widetilde{W}^{(2)}(\mathbf{k}) \underset{k\to 0}{\propto} 1/k^{2-\eta}. \tag{24.43}$$

Inserting the expansion (24.5) into expression (24.42), we find at $v = v_c$, in zero field, for k small,

$$\widetilde{W}^{(2)}(\mathbf{k}) \propto 1/k^2\,, \tag{24.44}$$

result which yields the classical or mean field value of the exponent η:

$$\eta = 0\,. \tag{24.45}$$

More generally for $k^2 = O(v_c - v)$ small, $\widetilde{W}^{(2)}(\mathbf{k})$ can be expanded as:

$$\widetilde{W}^{(2)}(\mathbf{k}) = \widetilde{W}^{(2)}(0)\left[1 + k^2\xi^2 + O\left(k^4/(v - v_c)\right)\right]^{-1}. \tag{24.46}$$

This shows that the two-point function has an Ornstein and Zernike or free field form. We have parametrized $\widetilde{W}^{(2)}(\mathbf{k})$ by introducing ξ^2 which is proportional to the second moment of $W^{(2)}(\mathbf{r})$. 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.

Note that the expansions (24.5) and thus (24.46) have up to order k^2 an $O(d)$ symmetry, and thus a larger symmetry that the complete function $\tilde{V}$.

For non-vanishing magnetization, $M^2 = O(v - v_c)$, from equations (24.42,24.46) and with the parametrization (24.12), we then derive

$$\xi^{-2} = \frac{1}{2v_c a^2}\left(b_2 + \frac{1}{2}b_4 M^2 - 2v\right). \tag{24.47}$$

In zero magnetic field, this yields

$$\begin{aligned} \xi_+^{-2} &= a^{-2}(1 - v/v_c) \qquad &\text{for}\quad T > T_c\,,\\ \xi_-^{-2} &= 2a^{-2}(v/v_c - 1) \qquad &\text{for}\quad T < T_c\,. \end{aligned} \tag{24.48}$$

Introducing general correlation length exponents ν, ν':

$$\xi_+ = f_+(T - T_c)^{-\nu}, \qquad \xi_- = f_-(T_c - T)^{-\nu'}, \tag{24.49}$$

we infer from the relations (24.48) the classical values of exponents

$$\nu = \nu' = 1/2\,,$$

and the universal ratio of amplitudes,

$$f_+/f_- = \sqrt{2}\,. \tag{24.50}$$

24.4.3 Continuous symmetries

If the initial spin variable S_i is a 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 found. Let us illustrate these properties with the example of the $O(N)$ symmetry.

Mean field theory. The mean field thermodynamic potential now takes the form

$$\Gamma(\mathbf{M}) = -\sum_{ij} V_{ij}\mathbf{M}_i \cdot \mathbf{M}_j + \sum_i \mathcal{B}(\mathbf{M}_i^2), \tag{24.51}$$

where the function $\mathcal{B}(\mathbf{M}^2)$ has the same general properties as the function $B(M)$ of the Ising-like case, and can be parametrized as

$$\mathcal{B}(X) = \frac{b_2}{2}X + \frac{b_4}{24}X^2 + \cdots .$$

Differentiating with respect to $M_{\alpha,i}$, where $M_{\alpha,i}$ ($\alpha = 1, \ldots, N$) are the components of the magnetization vector $\mathbf{M}_i$, we find

$$\frac{\partial \Gamma(M)}{\partial M_{\alpha,i}} = -2\sum_j V_{ij} M_{\alpha,j} + 2M_{\alpha,i}\mathcal{B}'(\mathbf{M}^2). \tag{24.52}$$

For uniform field and magnetization this yields

$$H = M\left(-2v + \mathcal{B}'(M^2)\right) \sim (-2v + b_2)M + \tfrac{1}{6}b_4 M^3, \tag{24.53}$$

in which H, M now are the lengths of the vectors $\mathbf{M}, \mathbf{H}$.

From equation (24.51), we compute $\Gamma_{ij}^{(2)\alpha\beta}$, the inverse two-point correlation function:

$$\Gamma^{(2)}_{\alpha\beta,ij} = \left.\frac{\partial^2 \Gamma(M)}{\partial M_{\alpha,i}\partial M_{\beta,j}}\right|_{\mathbf{M}_i=\mathbf{M}} = \left(-2V_{ij} + 2\delta_{ij}\mathcal{B}'(M^2)\right)\delta_{\alpha\beta} + 4\delta_{ij}M_\alpha M_\beta \mathcal{B}''(M^2). \tag{24.54}$$

in which M_α ($\alpha = 1, ..., N$) are the components of the uniform magnetization vector. Let us denote by $\mathbf{u}$ the unit vector along the direction of spontaneous magnetization:

$$\mathbf{M} = M\mathbf{u}_\alpha .$$

The two-point function (24.54) can be decomposed into a transverse and longitudinal part and can be written as

$$\Gamma^{(2)}_{ij,\alpha\beta} = u_\alpha u_\beta \Gamma^{(2)}_{\mathrm{L},ij} + \left(\delta_{\alpha\beta} - u_\alpha u_\beta\right)\Gamma^{(2)}_{\mathrm{T},ij} . \tag{24.55}$$

The two components are given in Fourier representation by

$$\tilde{\Gamma}^{(2)}_{\mathrm{L}}(\mathbf{k}) = 2\mathcal{B}'(M^2) + 4M^2\mathcal{B}''(M^2) - 2\tilde{V}(\mathbf{k}), \tag{24.56a}$$

$$\tilde{\Gamma}^{(2)}_{\mathrm{T}}(\mathbf{k}) = 2\mathcal{B}'(M^2) - 2\tilde{V}(\mathbf{k}). \tag{24.56b}$$

The expressions (24.56a) and (24.42) are similar. Using equation (24.53), we can rewrite $\tilde{\Gamma}^{(2)}_{\mathrm{T}}(\mathbf{k})$ as

$$\tilde{\Gamma}^{(2)}_{\mathrm{T}}(\mathbf{k}) = H/M + 2\left[\tilde{V}(0) - \tilde{V}(\mathbf{k})\right]. \tag{24.57}$$

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, implying the presence of $N-1$ Goldstone (massless) modes.

Goldstone modes: generalization. We remark that for a uniform magnetization the function $\Gamma(M)$ in (24.51), is the most general consistent with $O(N)$ symmetry:

$$\Gamma(M) = -v\mathbf{M}^2 + \mathcal{B}(\mathbf{M}^2).$$

Therefore, the restriction of the previous results to zero momentum are not specific to mean field theory. In the ordered phase, the transverse correlation length always diverges, and the Goldstone phenomenon is completely general, in agreement with the results proven in Sections 13.3, 13.4 (to which we refer for an extensive discussion of this question).

24.5 Beyond Mean Field Approximation

Landau's theory. Although all preceding results have been established in the framework of a special approximation scheme, the mean field approximation, they also follow, as shown by Landau, from more general assumptions:

(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,

$$\Gamma(M) = \sum_n \frac{1}{n!}\int \mathrm{d}^d k_1 \ldots \mathrm{d}^d k_n\, \delta^d\left(\sum k_i\right) \tilde{M}(k_1)\ldots\tilde{M}(k_n)\tilde{\Gamma}^{(n)}(k_1,\ldots,k_n),$$

(we assume translation invariance) are regular for real values of k_i. Finally, second order transition implies $\tilde{\Gamma}^{(4)}(0,0,0,0)$ positive.

Hidden in these 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 even more puzzling that the universal mean field predictions are in quantitative (sometimes even qualitative) disagreement with the empirical information available. By examining the leading corrections to the saddle point approximation (24.21), we will now find some clues. A remark is, here, in order: the validity of the saddle point approximation relies not 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* approximation, since this property is exact in a gaussian model.

An expansion parameter. Since from the point of view of the representation (24.58) of the partition function, the mean field approximation appears 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 discuss the mean field expansion it is useful 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 with identical distribution $\mathrm{d}\rho(S)$:

$$\sigma_i = \frac{1}{\ell}\sum_{k=1}^{\ell} S_i^{(k)} .$$

The case $\ell = 1$ corresponds to the initial distribution. We calculate the σ distribution $R(\sigma)$ by the usual method, introducing a Fourier representation

$$\begin{aligned} R(\sigma) &= \int \prod_k \mathrm{d}\rho(S^{(k)})\delta\left(\ell\sigma - \sum S^{(k)}\right) \\ &= \frac{1}{2i\pi}\int \prod_k \mathrm{d}\rho(S^{(k)}) \int \mathrm{d}\lambda \, \exp\left[\lambda\left(\sum_{k=1}^{\ell} S^{(k)} - \ell\sigma\right)\right] \\ &= \frac{1}{2i\pi}\int \mathrm{d}\lambda \exp\left[\ell(A(\lambda) - \lambda\sigma)\right], \end{aligned}$$

where the λ integration runs along the imaginary axis. 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 reads

$$\mathcal{Z}(H) = \int \prod_i \mathrm{d}\sigma_i \mathrm{d}\lambda_i \, \exp\left[-\ell\beta\,\mathcal{H}(\sigma) + \ell\sum_i(-\lambda_i\sigma_i + A(\lambda_i))\right]. \tag{24.58}$$

This expression clearly shows that a steepest descent calculation of the partition function generates an expansion in powers of $1/\ell$. Mean field theory is the leading order approximation, analogous to the tree approximation of quantum field theory.

It is convenient to define

$$\mathcal{W}(H) \equiv \frac{1}{\ell}\ln \mathcal{Z}\,, \tag{24.59}$$

in such a way that mean field expressions remain unchanged.

We now set

$$\int \mathrm{d}\lambda \; \mathrm{e}^{\ell[A(\lambda)-\lambda\sigma]} = \mathrm{e}^{-\ell\Sigma(\sigma,\ell)} . \tag{24.60}$$

Then, at leading order, for $\ell \to \infty$,

$$\Sigma(\sigma,\ell) = B(\sigma) + O(1/\ell).$$

In terms of Σ the partition function $\mathcal{Z}(H)$ has the representation

$$\mathcal{Z}(H) = \int \prod_i \mathrm{d}\sigma_i \mathrm{d}\lambda_i \, \exp\left[-\ell\beta\,\mathcal{H}(\sigma) - \ell\sum_i \Sigma(\sigma_i,\ell)\right]. \tag{24.61}$$

We recognize a lattice regularized functional integral for a scalar boson field of the type studied in Chapter 7. The mean field approximation is analogous to the tree level approximation of quantum field theory.

Interpretation. A more physical interpretation of the parameter ℓ is the following: we decompose the initial lattice into identical cells containing each ℓ spins. We take ℓ large but finite. Then, from the remark at the end of Section 24.2, the distribution of the average spin σ in a cell is close to a gaussian distribution. In expression (24.58), it is obtained by completely neglecting the interaction. The effective interaction between average, macroscopic spins is complicated but can be approximated by a pair interaction because the average spins are small. The partition function (24.58) thus can be considered as the initial partition function expressed in terms of macroscopic spins of a new lattice, with a quasi-gaussian distribution and an effective pair interaction (in general different from the initial spin interaction).

Perturbative expansion. At order $1/\ell$ two types of corrections appear. First, corrections to $\Sigma(\sigma,\ell)$ are generated by expanding (24.60) to order $1/\ell$. However, these corrections are equivalent to a modification of the coefficients of $B(\sigma)$ 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 replace $\Sigma(\sigma,\ell)$ by $B(\sigma)$:

$$\mathcal{Z}(H)=\int\prod_i \mathrm{d}\sigma_i \exp \ell\left[\sum_{ij}\sigma_i V_{ij}\sigma_j-\sum_i B(\sigma_i)+\sum_i H_i\sigma_i\right]. \tag{24.62}$$

We now recognize in the hamiltonian the thermodynamic potential of the mean field approximation. This suggests that in general fluctuations can be calculated by taking as a hamiltonian the thermodynamic potential of Landau's theory.

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 (7.101)) $\Gamma(M)$ is given by

$$\Gamma(M)=-\sum_{ij}M_iV_{ij}M_j+\sum_i B(M_i)+\frac{1}{2\ell}\operatorname{tr}\ln\left[-2V_{ij}+B''(M_i)\delta_{ij}\right]. \tag{24.63}$$

Expanding the tr ln in powers of M to order M^2 we obtain the inverse two-point function $\Gamma^{(2)}$ to order $1/\ell$. We introduce the propagator Δ_{ij}:

$$\Delta_{ij}^{-1}=b_2\delta_{ij}-2V_{ij},$$

Its Fourier transform $\tilde{\Delta}(\mathbf{k})$ has for k small, the expansion

$$\tilde{\Delta}^{-1}(\mathbf{k})=b_2-2\tilde{V}(\mathbf{k})=b_2-2v+2va^2k^2+O\left(k^4\right). \tag{24.64}$$

Then, in momentum representation, we find

$$\tilde{\Gamma}^{(2)}(\mathbf{p})=b_2-2\tilde{V}(\mathbf{p})+\frac{b_4}{2\ell}\int\frac{\mathrm{d}^dk}{(2\pi)^d}\tilde{\Delta}(\mathbf{k})+O\left(\frac{1}{\ell^2}\right). \tag{24.65}$$

In particular, the zero momentum value is the inverse of the magnetic susceptibility χ:

$$\chi^{-1} = \tilde{\Gamma}^{(2)}(0) = b_2 - 2v + \frac{b_4}{2\ell} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \tilde{\Delta}(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \tag{24.66}$$

The critical value v_c now is given by

$$0 = b_2 - 2v_c + \frac{b_4}{2\ell} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right), \tag{24.67}$$

Subtracting this equation from equation (24.66), we find

$$\chi^{-1} = 2(v_c - v) + \frac{b_4}{\ell} \int \frac{\mathrm{d}^d k}{(2\pi)^d} [\tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k})] \tilde{\Delta}(\mathbf{k}) \tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \tag{24.68}$$

In the $1/\ell$ contribution, we can replace v_c by its mean field value $b_2/2$. Then, for $v \sim v_c$ and k small (equation (24.5)),

$$\begin{aligned} \tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k}) &= (v - v_c)\tilde{U}(\mathbf{k}) \sim (v - v_c)\left(1 + O(k^2)\right) \\ \tilde{\Delta}^{-1}(\mathbf{k}) &= 2\left(v_c - v + v_c a^2 k^2\right) + O\left(k^4\right). \end{aligned}$$

Discussion. From these two expressions, it is clear that we can safely take the limit $v = v_c$ inside the integrand provided the integral $\int \mathrm{d}^d k/(k^2)^2$ converges at zero momentum, that is, for $d > 4$. Thus, χ^{-1} vanishes linearly in $(v - v_c)$ or $T - T_c$ as in mean field theory for $d > 4$. The susceptibility exponent γ 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, we find

$$\begin{aligned} \chi^{-1} &= 2(v_c - v)\left[1 + \frac{b_4}{8v_c^2 a^d \ell} \frac{\Gamma(1 - d/2)}{(4\pi)^{d/2}} \left(\frac{v_c - v}{v_c}\right)^{d/2-2}\right] + O\left(\frac{1}{\ell^2}\right), \\ &= 2(v_c - v)\left[1 + \frac{b_4}{8v_c^2 a^d \ell} \frac{\Gamma(1 - d/2)}{(4\pi)^{d/2}} (\xi/a)^{4-d}\right] + O\left(\frac{1}{\ell^2}\right), \end{aligned} \tag{24.69}$$

where equation (24.48) 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 mean field theory can no longer be trusted:

(i) For dimension $d > 4$ the "universal" predictions of mean field theory remain unchanged.

(ii) For dimension $d \leq 4$ "infrared" (IR) divergences, that is, singularities due to the small momentum behaviour of the propagator, show that the predictions of mean field theory cannot be correct in general. The hypothesis of quasi-gaussian behaviour is invalidated by the singular behaviour of fluctuations at the critical temperature. Inspection in Section 24.6 of higher order corrections will confirm this result.

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 that will be examined in the next chapters.

In particular, the zero momentum value is the inverse of the magnetic susceptibility χ:

$$\chi^{-1} = \tilde{\Gamma}^{(2)}(0) = b_2 - 2v + \frac{b_4}{2\ell}\int \frac{\mathrm{d}^d k}{(2\pi)^d}\tilde{\Delta}(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \tag{24.66}$$

The critical value v_c now is given by

$$0 = b_2 - 2v_c + \frac{b_4}{2\ell}\int \frac{\mathrm{d}^d k}{(2\pi)^d}\tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right), \tag{24.67}$$

Subtracting this equation from equation (24.66), we find

$$\chi^{-1} = 2(v_c - v) + \frac{b_4}{\ell}\int \frac{\mathrm{d}^d k}{(2\pi)^d}[\tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k})]\tilde{\Delta}(\mathbf{k})\tilde{\Delta}_c(\mathbf{k}) + O\left(\frac{1}{\ell^2}\right). \tag{24.68}$$

In the $1/\ell$ contribution, we can replace v_c by its mean field value $b_2/2$. Then, for $v \sim v_c$ and k small (equation (24.5)),

$$\begin{aligned}\tilde{V}(\mathbf{k}) - \tilde{V}_c(\mathbf{k}) &= (v - v_c)\tilde{U}(\mathbf{k}) \sim (v - v_c)\left(1 + O(k^2)\right)\\ \tilde{\Delta}^{-1}(\mathbf{k}) &= 2\left(v_c - v + v_c a^2 k^2\right) + O\left(k^4\right).\end{aligned}$$

Discussion. From these two expressions, it is clear that we can safely take the limit $v = v_c$ inside the integrand provided the integral $\int \mathrm{d}^d k/(k^2)^2$ converges at zero momentum, that is, for $d > 4$. Thus, χ^{-1} vanishes linearly in $(v - v_c)$ or $T - T_c$ as in mean field theory for $d > 4$. The susceptibility exponent γ remains $\gamma = 1$. However, for $d \le 4$ 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, we find

$$\begin{aligned}\chi^{-1} &= 2(v_c - v)\left[1 + \frac{b_4}{8v_c^2 a^d \ell}\frac{\Gamma(1 - d/2)}{(4\pi)^{d/2}}\left(\frac{v_c - v}{v_c}\right)^{d/2-2}\right] + O\left(\frac{1}{\ell^2}\right),\\ &= 2(v_c - v)\left[1 + \frac{b_4}{8v_c^2 a^d \ell}\frac{\Gamma(1 - d/2)}{(4\pi)^{d/2}}(\xi/a)^{4-d}\right] + O\left(\frac{1}{\ell^2}\right),\end{aligned} \tag{24.69}$$

where equation (24.48) 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 mean field theory can no longer be trusted:

(i) For dimension $d > 4$ the "universal" predictions of mean field theory remain unchanged.

(ii) For dimension $d \le 4$ "infrared" (IR) divergences, that is, singularities due to the small momentum behaviour of the propagator, show that the predictions of mean field theory cannot be correct in general. The hypothesis of quasi-gaussian behaviour is invalidated by the singular behaviour of fluctuations at the critical temperature. Inspection in Section 24.6 of higher order corrections will confirm this result.

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 that will be examined in the next chapters.

24.6 Power Counting and the Role of Dimension 4

In the preceding section, we have shown that the universal predictions of mean field theory cannot be trusted below four dimensions. We also note that the first divergent corrections depend on additional details of the initial microscopic model. At first sight, this would suggest that large distance properties are short distance sensitive and, therefore, no form of universality survives, in contradiction with experimental and numerical indications. However, a closer inspection of equation (24.69), for example, reveals that, at least at leading order, only one new parameter appears, the coefficient of the σ^4 term in the small σ expansion of the interaction. We now want to extend the analysis of "infrared divergences" to all orders in the mean field expansion, to understand whether the roles of dimension 4 and σ^4 interaction survive. Since the analysis is perturbative, we have to assume that, in some sense, deviations from mean field theory are not too large. This point will be further discussed in the coming chapters.

In the spirit of what has been done in Section 24.5, we first perform, order by order in the mean field expansion, a magnetic susceptibility or mass (in field theoretical language) renormalization. We trade the parameter v for the parameter μ:

$$\mu^2 = \chi^{-1} = \tilde{\Gamma}^{(2)}(0), \tag{24.70}$$

by inverting order by order the relation between v and μ (see Section 10.1).

The σ-propagator, for what concerns the most singular contribution at low momentum, can then be replaced by

$$\tilde{\Delta}(k) = \frac{1}{\mu^2 + k^2}. \tag{24.71}$$

We now consider the individual contributions coming from all (even due to $\sigma \mapsto -\sigma$ symmetry) powers of σ in $B(\sigma)$, 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 (24.71), integrated up to a "cut-off" of order unity since the momenta $\mathbf{k}$ are limited to a Brillouin zone $|k_\nu| \le \pi$.

In Chapter 9, we have associated with each diagram a dimension given by power counting. If v_α is the number of vertex of type α, which connect n_α lines and contain s_α powers of momenta, the dimension $\delta(\gamma)$ of a diagram γ contributing to the n-point function $\tilde{\Gamma}^{(n)}$ is (equations (9.10,9.18,9.22))

$$\delta(\gamma) = d - \frac{n}{2}(d-2) + \sum_\alpha v_\alpha \left[s_\alpha + n_\alpha (d/2 - 1) - d\right]. \tag{24.72}$$

Let us for simplicity consider a diagram γ with vanishing external momenta and perform the change of variables (for all integration momenta):

$$\mathbf{k} = \mu \mathbf{k}'. \tag{24.73}$$

After this rescaling, the contribution $D(\gamma)$ of the diagram γ takes the form

$$D(\gamma) = \mu^{\delta(\gamma)} D'(\gamma),$$

by definition of $\delta(\gamma)$. $D'(\gamma)$, which is calculated with a propagator $1/(k^2+1)$, is no longer IR divergent when μ goes to zero, but may have large momentum (UV) divergences since the cut-off, which is now of order $1/\mu$, becomes infinite in the same limit.

If the condition

$$\sum_{\alpha} v_\alpha \left[s_\alpha + n_\alpha (d/2 - 1) - d\right] > 0\,, \tag{24.74}$$

is satisfied for all vertices, $\delta(\gamma)$ increases with the order in perturbation theory. However, at the same time $D'(\gamma)$ has UV divergences proportional to $1/\mu$ to a power given by power counting. For UV divergent diagrams the two powers of μ exactly cancel and, therefore, all contributions in perturbation theory have the same behaviour, that is, the behaviour of mean field theory. This in particular applies to the two-point function for which $\delta(\gamma) > 2$. For higher correlation functions, a finite number of diagrams may be UV convergent and, therefore, are more IR singular than predicted by the tree approximation. Then, the leading contribution is given by the one-loop diagram with only vertices having the smallest possible dimension.

The condition (24.74) is satisfied if all interactions are non-renormalizable. For a theory with interactions even in σ this implies $d > 4$. The leading order result of Section 24.5 then extends to all orders: mean field theory is qualitatively correct order by order in perturbation theory; the effect of loop corrections is to modify the non-universal parameters of the mean field approximation.

If instead there exist interactions such that

$$\sum_{\alpha} v_\alpha \left[s_\alpha + n_\alpha (d/2 - 1) - d\right] < 0\,, \tag{24.75}$$

then one can find diagrams in which increasingly large negative powers of μ have been factorized. These diagrams are superficially convergent at large momenta. This implies that either they have no divergent subdiagrams and their singularity is indeed given by the power $\delta(\gamma)$ or they are even more singular when μ goes to zero. In all cases, the singularity of diagrams increases without bound with the order of perturbation theory.

We now characterize the most divergent diagrams order by order in perturbation theory. If we forget about the divergent subdiagrams we have just to find the minimum of $\delta(\gamma)$ at each order. This minimum is obtained for the least UV divergent interaction, that is, for $s_\alpha = 0$ and n_α as small as possible (we assume $d > 2$). The smallest even value is $n_\alpha = 4$. This corresponds to the σ^4 interaction.

We then examine the problem of divergent subdiagrams. By trading v for μ, we have performed a mass renormalization. According to the analysis of Chapters 9–10, the σ^4 diagrams are thus UV finite below four dimensions. The nature of the IR singularity deduced from the simple scaling (24.73) is thus the correct one.

This is no longer true for more UV divergent interactions which even below four dimensions require additional renormalizations. However, it will be shown in Chapter 27 that these interactions can be decomposed into the sum of an effective σ^4 interaction and a part which is less singular.

Therefore, the most IR singular terms, order by order in perturbation theory are generated by an effective σ^4 field theory, the coefficient of the σ^4 interaction being not the coefficient read in the original hamiltonian but a coefficient renormalized by the other interactions. In Chapters 25,26, we shall investigate the large distance behaviour of correlation functions in this effective σ^4 field theory. In Chapter 27, we shall discuss the effect of the neglected subleading divergences to show the internal consistency of the method.

Note, finally, that for $d = 4$, mean field behaviour is modified by logarithmic corrections.

If the condition

$$\sum_{\alpha} v_\alpha \left[s_\alpha + n_\alpha \left(d/2 - 1\right) - d\right] > 0\,, \tag{24.74}$$

is satisfied for all vertices, $\delta(\gamma)$ increases with the order in perturbation theory. However, at the same time $D'(\gamma)$ has UV divergences proportional to $1/\mu$ to a power given by power counting. For UV divergent diagrams the two powers of μ exactly cancel and, therefore, all contributions in perturbation theory have the same behaviour, that is, the behaviour of mean field theory. This in particular applies to the two-point function for which $\delta(\gamma) > 2$. For higher correlation functions, a finite number of diagrams may be UV convergent and, therefore, are more IR singular than predicted by the tree approximation. Then, the leading contribution is given by the one-loop diagram with only vertices having the smallest possible dimension.

The condition (24.74) is satisfied if all interactions are non-renormalizable. For a theory with interactions even in σ this implies $d > 4$. The leading order result of Section 24.5 then extends to all orders: mean field theory is qualitatively correct order by order in perturbation theory; the effect of loop corrections is to modify the non-universal parameters of the mean field approximation.

If instead there exist interactions such that

$$\sum_{\alpha} v_\alpha \left[s_\alpha + n_\alpha \left(d/2 - 1\right) - d\right] < 0\,, \tag{24.75}$$

then one can find diagrams in which increasingly large negative powers of μ have been factorized. These diagrams are superficially convergent at large momenta. This implies that either they have no divergent subdiagrams and their singularity is indeed given by the power $\delta(\gamma)$ or they are even more singular when μ goes to zero. In all cases, the singularity of diagrams increases without bound with the order of perturbation theory.

We now characterize the most divergent diagrams order by order in perturbation theory. If we forget about the divergent subdiagrams we have just to find the minimum of $\delta(\gamma)$ at each order. This minimum is obtained for the least UV divergent interaction, that is, for $s_\alpha = 0$ and n_α as small as possible (we assume $d > 2$). The smallest even value is $n_\alpha = 4$. This corresponds to the σ^4 interaction.

We then examine the problem of divergent subdiagrams. By trading v for μ, we have performed a mass renormalization. According to the analysis of Chapters 9–10, the σ^4 diagrams are thus UV finite below four dimensions. The nature of the IR singularity deduced from the simple scaling (24.73) is thus the correct one.

This is no longer true for more UV divergent interactions which even below four dimensions require additional renormalizations. However, it will be shown in Chapter 27 that these interactions can be decomposed into the sum of an effective σ^4 interaction and a part which is less singular.

Therefore, the most IR singular terms, order by order in perturbation theory are generated by an effective σ^4 field theory, the coefficient of the σ^4 interaction being not the coefficient read in the original hamiltonian but a coefficient renormalized by the other interactions. In Chapters 25,26, we shall investigate the large distance behaviour of correlation functions in this effective σ^4 field theory. In Chapter 27, we shall discuss the effect of the neglected subleading divergences to show the internal consistency of the method.

Note, finally, that for $d = 4$, mean field behaviour is modified by logarithmic corrections.

24.7 Tricritical Points

So far we have assumed that we could vary one control parameter, the temperature, and that, therefore, the coefficient b_4 of the term M^4 in the expansion of $\Gamma(M)$ was 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. This occurs for instance in He^3–He^4 mixtures or some metamagnetic systems. In the Ising-like models we have considered so far, this can be achieved by introducing parameters in the spin distribution.

Let us expand $\Gamma(M)$ up to order M^6:

$$\Gamma(M) = -\sum_{ij} V_{ij} M_i M_j + \sum_i \left(\frac{b_2}{2!} M_i^2 + \frac{b_4}{4!} M_i^4 + \frac{b_6}{6!} M_i^6 + \cdots\right); \qquad (24.76)$$

If the coefficient b_6 of M^6 is positive, when b_4 decreases it is possible to follow a line of critical points until b_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, b_4 becomes negative and the transition becomes first order.

Corrections to the tricritical theory can be studied by the method of Section 24.6. One finds that the upper critical dimension now is 3, that is, above three dimensions mean field theory predicts correctly 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 field theory.

Bibliographical Notes

Some sections of this chapter has been inspired by the review of E. Brézin, J.C. Le Guillou and J. Zinn-Justin, published in

Phase Transitions and Critical Phenomena vol. 6, C. Domb and M.S. Green eds. (Academic Press, London 1976),

in which a brief account of the general method to generate the mean field approximation can also be found. See also

R. Brout, *Phys. Rep.* 10C (1974) 1.

For earlier articles and reviews about mean field theory see, for example,

P. Weiss, *J. Phys. Radium, Paris* 6 (1907) 667; M.E. Fisher, *Rep. Prog. Phys.* 30 (1967) 615; L.P. Kadanoff, W. Götze, D. Hamblen, R. Hecht, E.A.S. Lewis, V.V. Palciauskas, M. Rayl, J. Swift, D. Aspnes and J. Kane, *Rev. Mod. Phys.* 39 (1967) 395; H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York 1971).

The scaling form of the equation of state was proposed by

B. Widom, *J. Chem. Phys.* 43 (1965) 3892, 3898.

Landau's theory is presented in

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).

The role of the $\phi^4(x)$ field theory was emphasized by

K.G. Wilson, *Phys. Rev.* B4 (1971) 3174, 3184.

For a discussion of tricritical systems see
E.H. Graf, D.M. Lee and J.D. Reppy, *Phys. Rev. Lett.* 19 (1967) 417; R.J. Birgeneau, W.B. Yelon, E. Cohen and J. Makovsky, *Phys. Rev.* B5 (1972) 2607; M. Blume, *Phys. Rev.* 141 (1966) 517; H.W. Capel, *Physica* 32 (1966) 966; E.K. Riedel and F.J. Wegner, *Phys. Rev. Lett.* 29 (1972) 349.
General spin systems have been considered by
G. Horwitz and H.B. Callen, *Phys. Rev.* 124 (1961) 1757; F. Englert, *Phys. Rev.* 129 (1963) 567.
Appendix A24 directly follows from
J. Zinn-Justin, ***Perspectives in Particles and Fields***, Cargèse 1983, vol. B126, M. Lévy *et l* eds. (Plenum Press, New York 1985).

APPENDIX A24
MEAN FIELD EXPANSION: GENERAL FORMALISM

We now explain how the mean field approximation can be derived for a general spin model on a lattice. Again, the mean field approximation can be identified with a saddle point approximation, and, therefore, is the first term in a systematic expansion. For completeness, we finally indicate the relation between the mean field approximation and high and low temperature expansions.

In what follows we consider systems with one degree of freedom per lattice site but the generalization of the formalism to several degrees of freedom per site is straightforward. Mean field theory can also be generalized to lattice models with fermions. The equivalent of the variables S_i are then products of the form $\bar{\psi}_i \Gamma_A \psi_i$, in which 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.

A24.1 Mean Field Approximation

We consider a lattice model described in terms of lattice variables S_i in which i is the 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 $\mathcal{H}(S, S^2, \ldots, S^n)$. For example, a general pair potential has to be written as

$$\sum_{ij} V_{ij} S_i S_j = \sum_{i \neq j} V_{ij} S_i S_j + \sum_i V_{ii} S_i^2 .$$

The corresponding partition function is (the temperature factor β is incorporated in $\mathcal{H}$)

$$\mathcal{Z} = \int \left(\prod_i \mathrm{d}\rho(S_i) \right) \exp\left[-\mathcal{H}(S, S^2, \ldots, S^n) \right], \qquad (A24.1)$$

in which $\mathrm{d}\rho(S)$ is the spin distribution.

In order to construct the mean field approximation as the leading order in a steepest descent calculation, it is necessary to first transform expression $(A24.1)$. 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

$$\sigma_i^{(m)} = S_i^m , \qquad m = 1, \ldots, n . \qquad (A24.2)$$

The partition function can then be rewritten as

$$\begin{aligned} \mathcal{Z} = \int & \prod_i \mathrm{d}\rho(S_i) \prod_{i,m} \mathrm{d}\sigma_i^{(m)} \mathrm{d}\lambda_i^{(m)} \\ & \times \exp\left[-\mathcal{H}(\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(n)}) + \sum_{i,m} \lambda_i^{(m)} \left(S_i^m - \sigma_i^{(m)} \right) \right] . \end{aligned} \qquad (A24.3)$$

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 (A24.3), the integrations over all variables S_i are decoupled. We introduce the function A:

$$\exp\left[A(\lambda^{(1)},\ldots,\lambda^{(n)})\right]=\int \mathrm{d}\rho(S)\exp\left(\sum_{m=1}^{n}S^m\lambda^{(m)}\right). \qquad (A24.4)$$

Here, we assume that the integration measure $\mathrm{d}\rho(S)$ is either of compact support or decreasing fast enough so that the integral exists. The partition function can then be rewritten as

$$\mathcal{Z}=\int\prod_{i,m}\mathrm{d}\sigma_i^{(m)}\mathrm{d}\lambda_i^{(m)}\exp\left[-\mathcal{H}(\sigma)-\sum_{i,m}\lambda_i^{(m)}\sigma_i^{(m)}+\sum_i A(\lambda_i^{(1)},\ldots,\lambda_i^{(n)})\right]. \qquad (A24.5)$$

The mean field approximation is the leading order in the evaluation of this expression by steepest descent. The saddle point equations are

$$\frac{\partial\mathcal{H}}{\partial\sigma_i^{(m)}}=-\lambda_i^{(m)}, \qquad (A24.6a)$$

$$\frac{\partial A}{\partial\lambda_i^{(m)}}=\sigma_i^{(m)}. \qquad (A24.6b)$$

Equation (A24.6b) shows that, at leading order, the original weight $\prod_i \mathrm{d}\rho(S_i)\,\mathrm{e}^{-\mathcal{H}}$ has been approximated by a product of weights for each lattice site given by the r.h.s. of equation (A24.4), 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 Chapter 24. 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:

$$\langle S_{i_1}S_{i_2}\ldots S_{i_n}\rangle\mapsto\langle S_{i_1}\rangle\langle S_{i_2}\rangle\ldots\langle S_{i_n}\rangle.$$

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. 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 allow 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 above 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

$$\sum_i \frac{a}{2} S_i^2 + \frac{b}{4!} S_i^4 + \cdots , \tag{A24.7}$$

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 mean field theory depend on the formulation. The introduction of additional variables ensures that this is not the case.

Let us, for example, make the transformation

$$\begin{aligned} \mathcal{H}(S_i, S_i^2, \ldots, S_i^n) &\mapsto \mathcal{H}(S_i, \ldots, S_i^n) + \frac{a}{2} \sum_i S_i^2 , \\ d\rho(S) &\mapsto e^{aS^2/2}\, d\rho(S). \end{aligned} \tag{A24.8}$$

Obviously, the lattice theory is independent of a.

Consider now the modifications this transformation induces into equations ($A24.4$) and ($A24.6$):

$$\exp\left(\tilde{A}(\lambda^{(1)}, \ldots, \lambda^{(n)})\right) = \int d\rho(S)\, e^{aS^2/2} \exp\left(\sum_{m=1}^{n} S^m \lambda^{(m)}\right). \tag{A24.9}$$

Comparing the equation with equation ($A24.4$), we conclude

$$\tilde{A}(\lambda^{(1)}, \ldots, \lambda^{(n)}) = A(\lambda^{(1)}, \lambda^{(2)} + a/2, \lambda^{(3)}, \ldots, \lambda^{(n)}), \tag{A24.10}$$

and the saddle point equations become

$$\begin{aligned} &\frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} + \frac{a}{2} \delta_{l2} = -\lambda_i^{(m)}, \\ &\frac{\partial A}{\partial \lambda_i^{(m)}} (\lambda^{(1)}, \lambda^{(2)} + a/2, \lambda^{(3)}, \ldots, \lambda^{(n)}) = \sigma_i^{(m)} . \end{aligned} \tag{A24.11}$$

We see that the equations are identical up to the substitution

$$\lambda_i^{(2)} + a/2 \;\mapsto\; \lambda_i^{(2)} , \tag{A24.12}$$

which does not change mean field results.

Finally, we also now understand the technical assumption of Chapter 24, $V_{ii} = 0$: it simplified the mean field formalism by avoiding the introduction of variables related to S_i^2.

A24.2 Mean Field Expansion

It is then possible to correct the mean field approximation, by systematically expanding the effective hamiltonian in ($A24.3$) around the saddle points ($A24.6$) and integrating term by term as one does in usual perturbation theory. We call this expansion the mean field expansion. In Chapter 24, 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 ($A24.5$), then the mean field expansion is just a formal expansion in powers of $1/\ell$:

$$\begin{aligned}\mathcal{Z}_\ell &= \int \prod_{i,m} \mathrm{d}\sigma_i^{(m)} \mathrm{d}\lambda_i^{(m)} \exp\left[-\ell \mathfrak{H}(\lambda,\sigma)\right], \\ \mathfrak{H}(\lambda,\sigma) &= \mathcal{H}(\sigma) + \sum_{i,m} \lambda_i^{(m)}\sigma_i^{(m)} - \sum_i A(\lambda_i^{(1)},\ldots,\lambda_i^{(n)}) \end{aligned} \qquad .(A24.13)$$

As in Section 24.3.1, 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 $\mathrm{d}\rho(S)$:

$$\int \mathrm{d}\rho(S_i) \mapsto \int \prod_{\alpha=1}^{\ell} \mathrm{d}\rho(S_i^{(\alpha)}). \qquad (A24.14)$$

Then,

$$\exp\left[\ell A(\lambda^{(1)},\ldots,\lambda^{(n)})\right] = \int \prod_\alpha \mathrm{d}\rho(S^{(\alpha)}) \exp\left(\sum_{m=1}^{n} \lambda^{(m)} \sum_\alpha \left(S^{(\alpha)}\right)^m\right). \qquad (A24.15)$$

We insert this expression into the partition function ($A24.13$). We can then integrate over the Lagrange multipliers $\lambda^{(m)}$ and find

$$\sigma_i^{(m)} = \frac{1}{\ell}\sum_\alpha \left(S^{(\alpha)}\right)^m . \qquad (A24.16)$$

Therefore, as in the simpler example of Section 24.3.1 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 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.

A24.3 High, Low Temperature and Mean Field Approximations

There are interesting connections between mean field and high temperature or low temperature expansion. We show here only that the mean field approximation has the correct high temperature and low temperature limit.

High temperature. We now put an explicit factor β back 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 ($A24.6$) 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 ($A24.16$), we infer that a diagram of order K in β involving vertices which connect $v_1, v_2, \ldots, v_K$ lattice sites, respectively, is affected by ℓ to the power $-\sum v_k$. Finally, 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

$$\beta^K \mapsto K - \sum_{k=1}^{K} v_k + s \,. \tag{A24.17}$$

For a connected tree diagram, one has

$$s = \sum_{k=1}^{K} v_k - (K-1). \tag{A24.18}$$

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

$$s = \sum_{k=1}^{K} v_k - (K-1) - B \,. \tag{A24.19}$$

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.

Low temperature expansion. The mean field expansion also contains the low temperature expansion. We rewrite equations ($A24.6$) with a parameter β in front of the hamiltonian:

$$\beta \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} = -H_i^{(m)}, \quad \frac{\partial A}{\partial H_i^{(m)}} = \sigma_i^{(m)} \,. \tag{A24.20}$$

For large β (low temperature or weak coupling), the $H^{(m)}$ variables become large. To study this limit, we have thus to evaluate A for H large. This will in general select for S some classical value S_c. As a direct consequence, we have

$$\frac{\partial A}{\partial H_i^{(m)}} = (S_{i,c})^m = \sigma_i^{(m)} . \qquad (A24.21)$$

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 $\mathrm{d}\rho(S)$:

$$\mathrm{d}\rho(S) = \mathrm{d}S \frac{1}{2i\pi} \int_{-i\infty}^{+i\infty} \mathrm{e}^{\mu S}\, \tilde{\rho}(\mu) \mathrm{d}\mu . \qquad (A24.22)$$

A low temperature expansion is possible only if $\mathrm{d}\rho(S)$ decrease fast enough for large $|S|$, and thus $\tilde{\rho}(\mu)$ has to be an entire function.

For $H^{(m)}$ large, A can be calculated by the steepest descent method. The saddle point equations are

$$\mu = -\sum_m m H^{(m)} S^{m-1}, \qquad (A24.23)$$

$$S + \frac{\partial}{\partial \mu} \ln \tilde{\rho}(\mu) = 0 , \qquad (A24.24)$$

and

$$A(H^{(m)}) = \sum_m H^{(m)} S^m + \mu S + \ln \tilde{\rho}(\mu). \qquad (A24.25)$$

The mean field saddle point equations then are

$$\sigma_i^{(m)} = \frac{\partial A}{\partial H_i^{(m)}} = S_i^m , \qquad (A24.26)$$

$$H_i^{(m)} = -\beta \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} . \qquad (A24.27)$$

In the same notation, the relevant configuration at low temperature is given by the equations

$$\beta \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}S_i} \equiv \sum_i m S_i^{m-1} \frac{\partial \mathcal{H}}{\partial \sigma_i^{(m)}} = \mu_i , \qquad (A24.28)$$

and

$$S_i + \frac{\mathrm{d}}{\mathrm{d}\mu_i} \ln \tilde{\rho}(\mu) = 0 . \qquad (A24.29)$$

Now by summing equation ($A24.27$) over m after multiplication by $m\,(S_i)^{m-1}$, and using equation ($A24.23$), one reproduces equation ($A24.28$), while equations ($A24.24$) and ($A24.29$) are identical.

The mean field expansion is also a partial summation of the low temperature expansion.

25 GENERAL RENORMALIZATION GROUP. THE CRITICAL THEORY NEAR DIMENSION FOUR

In Chapter 24, we have considered Ising-like systems (and more generally ferromagnetic systems with $O(N)$ symmetry) with short range interactions and determined the behaviour of thermodynamical functions near a second order phase transition, in the mean field approximation, which is as we have shown a *quasi-gaussian* approximation. The mean field approximation predicts a set of *universal* properties, that is, properties which are 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; on the contrary, below four dimensions the corrections diverge at the critical temperature and the universal predictions of the mean field approximation can certainly not be correct. Moreover, we have identified in an expansion around the mean field the most singular terms at criticality and shown that they can be formally summed to give a continuum, massless ϕ^4 field theory, a result we will use in the second part of this chapter.

Such an analysis, which reflects the property that different scales of physics cannot be decoupled, seems to indicate that below four dimensions long distance properties are sensitive to the detailed microscopic structures. Surprisingly enough, however, some universality survives, although this universality is less general than in mean field theory in the sense that long distance properties depend on the space dimension, on symmetries of the system and some other qualitative properties. To understand this universality beyond the mean field (quasi-gaussian) approximation it is necessary to call upon a new idea. The *renormalization group* (RG) will provide us with the essential tool we need.

The RG theory, as applied to Critical Phenomena, has been developed by Kadanoff, Wilson, Wegner and many others. We will first describe the basic renormalization group ideas in a somewhat abstract and intuitive framework. The formulation will lack precision and the arguments will be largely heuristic. The importance of fixed points in hamiltonian space will be stressed. The special role of gaussian models and their universal properties will be related to the existence of fixed point, the gaussian fixed point.

The abstract RG theory, although its formulation is vague, is extremely suggestive, and indeed it has been implemented in many approximate forms and induced a wealth of practical calculations. It is not our purpose to review them here. However, in the more specific context of Quantum Field Theory (QFT), assumptions at the basis of the renormalization group have been clarified and verified in many cases of physical interest, confirming in a very direct manner the deep connection, first recognized by Wilson, between QFT and the theory of Critical Phenomena. The methods of perturbative QFT then have allowed to efficiently calculate universal quantities for a large class of statistical models.

Therefore, in Section 25.3, we will use what we have learned both from the analysis of corrections to mean field theory in Chapter 24 and of the relevant eigenoperators at the gaussian fixed point near dimension 4 and construct an approximate renormalization group for an effective ϕ^4 field theory. We will show that the RG equations which appear as a consequence of the necessity of renormalization of local field theories are directly connected with the abstract RG equations we have introduced in Section 25.1.

Universality in the theory of critical phenomena is thus directly related to the property that local field theories are insensitive to the short distance structure, and physics can, therefore, be described by renormalized correlation functions. Conversely, in the statistical sense, QFTs are always close to criticality and their existence, beyond perturbation theory, relies, from the abstract RG point of view, on the presence of IR fixed points in hamiltonian space.

Finally, it will be possible to solve RG equations in $d = 4 - \varepsilon$ dimensions and, following Wilson and Fisher, to calculate universal quantities in an ε-expansion.

25.1 Renormalization Group: The General Idea

We consider a general hamiltonian $\mathcal{H}(\phi)$, function of a field $\phi(x)$ $(x \in \mathbb{R}^d)$. We assume that $\mathcal{H}$ is expandable in powers of ϕ:

$$\mathcal{H}(\phi) = \sum_{n=0} \frac{1}{n!} \int \mathrm{d}^d x_1 \mathrm{d}^d x_2 \ldots \mathrm{d}^d x_n \, \mathcal{H}_n(x_1, x_2, \ldots, x_n) \phi(x_1) \ldots \phi(x_n), \tag{25.1}$$

and has all the properties of the thermodynamic potential of Landau's theory (Section 24.5). For example, the Fourier transforms of the functions $\mathcal{H}_n$ are regular at low momenta (assumption of short range forces or locality). The hamiltonian will in general depend on an infinite number of parameters or coupling constants.

To the hamiltonian $\mathcal{H}(\phi)$ (really the configuration energy), corresponds a set of connected correlation functions $W^{(n)}(x_1, \ldots, x_n)$:

$$W^{(n)}(x_1, x_2, \ldots, x_n) = \left[\int [\mathrm{d}\phi] \, \phi(x_1) \ldots \phi(x_n) \, \mathrm{e}^{-\beta \mathcal{H}(\phi)} \right]_{\text{connect.}} . \tag{25.2}$$

We want to study the long distance behaviour of critical correlation functions, that is, the behaviour of $W^{(n)}(\lambda x_1, \ldots, \lambda x_n)$ when the dilatation parameter λ becomes large.

25.1.1 The renormalization group 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 hamiltonian $\mathcal{H}_\lambda(\phi)$ which has essentially the same correlation functions at fixed space variables. More precisely we want to construct a hamiltonian $\mathcal{H}_\lambda(\phi)$ which has correlation functions $W_\lambda^{(n)}(x_i)$ satisfying

$$W_\lambda^{(n)}(x_1, \ldots, x_n) - Z^{-n/2}(\lambda) W^{(n)}(\lambda x_1, \ldots, \lambda x_n) = R_\lambda^{(n)}(x_1, \ldots, x_n), \tag{25.3}$$

where the functions $R^{(n)}$ decrease faster than any power of λ for $\lambda \to \infty$. The mapping $\mathcal{H}(\phi) \mapsto \mathcal{H}_\lambda(\phi)$ is called a RG transformation. We define the transformation such that $\mathcal{H}_{\lambda=1}(\phi) \equiv \mathcal{H}(\phi)$.

In the case of models invariant under space translations, equation (25.3) after a Fourier transformation reads

$$\widetilde{W}_\lambda^{(n)}(p_1, \ldots, p_n) = Z^{-n/2}(\lambda) \lambda^{(1-n)d} \widetilde{W}^{(n)}(p_1/\lambda, \ldots, p_n/\lambda) + \widetilde{R}_\lambda^{(n)} . \tag{25.4}$$

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 rescalings of space and field. However, this transformation has a fixed point only in exceptional cases (gaussian models) and thus more general transformations have to be considered (see Appendix A10.1). Below 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 $\mathcal{H}_\lambda$ are now all explicit functions of λ. Let us assume that we have found a RG transformation such that, when λ becomes large, the hamiltonian $\mathcal{H}_\lambda(\phi)$ has a limit $\mathcal{H}^*(\phi)$, the fixed point hamiltonian. If such a fixed point exists in hamiltonian space, then the correlation functions $W^{(n)}_\lambda$ have corresponding limits $W^{(n)}_*$ and equation (25.3) becomes

$$W^{(n)}(\lambda x_1,\ldots,\lambda x_n) \underset{\lambda\to\infty}{\sim} Z^{n/2}(\lambda) W^{(n)}_*(x_1,\ldots,x_n). \tag{25.5}$$

We now introduce a second scale parameter μ and calculate $W^{(n)}(\lambda\mu x_i)$ from equation (25.5) in two different ways, we obtain a relation involving only $W^{(n)}_*$:

$$W^{(n)}_*(\mu x_1,\ldots,\mu x_n) = Z^{n/2}_*(\mu) W^{(n)}_*(x_1,\ldots,x_n) \tag{25.6}$$

with

$$Z_*(\mu) = \lim_{\lambda\to\infty} Z(\lambda\mu)/Z(\lambda). \tag{25.7}$$

Equation (25.6) being valid for arbitrary μ immediately implies that Z_* forms a representation of the dilatation semi-group:

$$Z_*(\lambda_1)Z_*(\lambda_2) = Z_*(\lambda_1\lambda_2). \tag{25.8}$$

Thus, under reasonable assumptions,

$$Z_*(\lambda) = \lambda^{-2d_\phi}. \tag{25.9}$$

The fixed point correlation functions have a power law behaviour characterized by a positive number d_ϕ which is called the dimension of the order parameter $\phi(x)$.

Returning now to equation (25.7), we conclude that $Z(\lambda)$ also has asymptotically a power law behaviour. Equation (25.5) then shows that the correlation functions $W^{(n)}$ have a scaling behaviour at large distances:

$$W^{(n)}(\lambda x_1,\ldots,\lambda x_n) \underset{\lambda\to\infty}{\sim} \lambda^{-nd_\phi} W^{(n)}_*(x_1,\ldots,x_n) \tag{25.10}$$

with a power d_ϕ which is a property of the fixed point. The r.h.s. of the equation, which determines the critical behaviour of correlation functions, therefore, depends only on the fixed point hamiltonian. In other words, the correlation functions corresponding to all hamiltonians which 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 mean field theory, 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, that the corresponding hamiltonians are necessarily critical. Critical hamiltonians define in hamiltonian space the *critical surface* which is invariant under a RG flow. In the generic case where ξ is finite the correlation length ξ/λ corresponding to $\mathcal{H}_\lambda$ 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 $2d_\phi = d$.

25.1.2 Hamiltonian flows. Scaling operators

We now try to write the RG transformation, consequence of equation (25.5), for the hamiltonian itself. For this purpose, we 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 $\lambda(1+\mathrm{d}\lambda/\lambda)$.

We look for RG transformations which will bring the transformed hamiltonian closer to a fixed point. This condition involves only $\mathcal{H}_\lambda$ and not the previous history which has led from $\mathcal{H}$ to $\mathcal{H}_\lambda$: the equations should depend on λ only through $\mathcal{H}_\lambda$. Therefore, we can write the RG transformation in differential form in terms of a mapping $\mathcal{T}$ of the space of hamiltonians into itself and a real function η defined on the space of hamiltonians as

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathcal{H}_\lambda = \mathcal{T}[\mathcal{H}_\lambda], \tag{25.11}$$

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln Z(\lambda) = 2-d-\eta[\mathcal{H}_\lambda]. \tag{25.12}$$

As a function of the "time" $\ln\lambda$ these equations define a Markov's process. Note that we have written the r.h.s. of equation (25.12) in an unnatural way for later convenience.

A fixed point hamiltonian $\mathcal{H}^*$ is necessarily a solution of the fixed point equation:

$$\mathcal{T}[\mathcal{H}^*] = 0. \tag{25.13}$$

The dimension d_ϕ of the field ϕ then is

$$d_\phi = \tfrac{1}{2}(d-2+\eta[\mathcal{H}^*]). \tag{25.14}$$

To understand, at least locally, which hamiltonians flow into the fixed point it is necessary to study its stability. We, therefore, linearize the RG equations near the fixed point.

Linearized flow equations. The RG transformation being determined, we apply it to a hamiltonian $\mathcal{H}$ close to the fixed point $\mathcal{H}^*$. Setting $\mathcal{H}_\lambda = \mathcal{H}^* + \Delta\mathcal{H}_\lambda$, we linearize the RG equations:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\Delta\mathcal{H}_\lambda = L^*(\Delta\mathcal{H}_\lambda), \tag{25.15}$$

where L^* is a linear operator independent of λ, acting on hamiltonian space. Let us assume that L^* has a discrete set of eigenvalues l_i corresponding to a set of eigenoperators $\mathcal{O}_i$. Then, $\Delta\mathcal{H}_\lambda$ can be expanded on the $\mathcal{O}_i$'s:

$$\Delta\mathcal{H}_\lambda = \sum h_i(\lambda)\mathcal{O}_i\,, \tag{25.16}$$

and the transformation (25.15) becomes

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}h_i(\lambda) = l_i h_i(\lambda). \tag{25.17}$$

The integration then yields

$$h_i(\lambda) = \lambda^{l_i}h_i(1). \tag{25.18}$$

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:

(i) Eigenvalues which have a positive real part. The corresponding eigenoperators are called *relevant.* If $\mathcal{H}_\lambda$ has a component on one of these operators, this component will grow with λ, and $\mathcal{H}_\lambda$ will move away from the neighbourhood of $\mathcal{H}^*$. Operators associated with a deviation from criticality are clearly relevant since a dilatation decreases the effective correlation length. In Section 25.6, we calculate the corresponding eigenvalue for the ϕ^4 field theory.

(ii) Eigenvalues for which $\mathrm{Re}(l_i) = 0$. Then, two situations can arise: either $\mathrm{Im}(l_i)$ does not vanish, and the corresponding component has a periodic behaviour (no example will be met in this work) or $l_i = 0$. Eigenoperators corresponding to a vanishing eigenvalue are called *marginal.* In Section 25.5, we show that in the ϕ^4 field theory the operator $\phi^4(x)$ is marginal in four dimensions. To solve the RG equation (25.11) and determine the behaviour of the corresponding component it is necessary to expand beyond the linear approximation. Generically, one finds

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} h_i(\lambda) \sim B h_i^2 \,. \tag{25.19}$$

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

$$h_i(\lambda) \sim -1/(B \ln \lambda)\,. \tag{25.20}$$

A marginal 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 which corresponds to a motion along the line is obviously marginal (see Chapter 33).

(iii) Eigenvalues which have 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 the physics. An example is provided by the operator realizing a constant multiplicative renormalization of the dynamical variables $\phi(x)$. These operators are called *redundant.* In quantum field theory quantum, equation of motions 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 codimension 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 emerge 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 will hardly be 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 correlation length and 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.

25.1.3 Explicit RG equations for correlation functions

Equation (25.3) can be rewritten (neglecting R_n) as

$$Z^{n/2}(\lambda)W_\lambda^{(n)}(x_1/\lambda,\ldots,x_n/\lambda)=W^{(n)}(x_1,\ldots,x_n). \tag{25.21}$$

We now assume that the hamiltonian $\mathcal{H}_\lambda$ has been parametrized in terms of constants $h_i(\lambda)$. We can then change notation and write

$$W_\lambda^{(n)}(x_1,\ldots,x_n)\equiv W^{(n)}(\{h(\lambda)\};x_1,\ldots,x_n),$$

where $\{h(\lambda)\}$ stands for the set of all $h_i(\lambda)$. We then differentiate equation (25.21) with respect to λ. The r.h.s. does not depend on λ and, therefore,

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[Z^{n/2}(\lambda)W^{(n)}(\{h(\lambda)\};x_1/\lambda,\ldots,x_n/\lambda)\right]=0\,. \tag{25.22}$$

We introduce the differential operator

$$\mathrm{D_{RG}}\equiv-\sum_\ell x_\ell\frac{\partial}{\partial x_\ell}-\sum_i\beta_i(h)\frac{\partial}{\partial h_i}+\frac{n}{2}(2-d-\eta(h))$$

with

$$\beta_i(h)=-\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}h_i(\lambda), \tag{25.23}$$

$$2-d-\eta(h)=\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln Z(\lambda), \tag{25.24}$$

where we have used properties of equations (25.11,25.12) which imply that β_i et η do not depend on λ explicitly, but only through the $h_i(\lambda)$. Equation (25.22) then reads

$$\mathrm{D_{RG}}W^{(n)}(\{h(\lambda)\};x_1/\lambda,\ldots,x_n/\lambda)=0\,.$$

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,

$$\left[-\sum_\ell x_\ell\frac{\partial}{\partial x_\ell}-\sum_i\beta_i(h)\frac{\partial}{\partial h_i}+\frac{n}{2}(2-d-\eta(h))\right]W^{(n)}(\{h\};x_1,\ldots,x_n)=0\,. \tag{25.25}$$

With this notation, a fixed point is defined by the common solution to all equations

$$\beta_i(h^*)=0\,,$$

and equation (25.25) then implies the scaling behaviour derived more directly.

operators induce a finite correlation length, the maximal dilatation is of the order of the ratio between correlation length and 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.

25.1.3 Explicit RG equations for correlation functions

Equation (25.3) can be rewritten (neglecting R_n) as

$$Z^{n/2}(\lambda)W_\lambda^{(n)}(x_1/\lambda,\ldots,x_n/\lambda)=W^{(n)}(x_1,\ldots,x_n). \tag{25.21}$$

We now assume that the hamiltonian $\mathcal{H}_\lambda$ has been parametrized in terms of constants $h_i(\lambda)$. We can then change notation and write

$$W_\lambda^{(n)}(x_1,\ldots,x_n)\equiv W^{(n)}\bigl(\{h(\lambda)\};x_1,\ldots,x_n\bigr),$$

where $\{h(\lambda)\}$ stands for the set of all $h_i(\lambda)$. We then differentiate equation (25.21) with respect to λ. The r.h.s. does not depend on λ and, therefore,

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[Z^{n/2}(\lambda)W^{(n)}\bigl(\{h(\lambda)\};x_1/\lambda,\ldots,x_n/\lambda\bigr)\right]=0\,. \tag{25.22}$$

We introduce the differential operator

$$\mathrm{D_{RG}}\equiv-\sum_\ell x_\ell\frac{\partial}{\partial x_\ell}-\sum_i\beta_i(h)\frac{\partial}{\partial h_i}+\frac{n}{2}\bigl(2-d-\eta(h)\bigr)$$

with

$$\beta_i(h)=-\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}h_i(\lambda), \tag{25.23}$$

$$2-d-\eta(h)=\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln Z(\lambda), \tag{25.24}$$

where we have used properties of equations (25.11,25.12) which imply that β_i et η do not depend on λ explicitly, but only through the $h_i(\lambda)$. Equation (25.22) then reads

$$\mathrm{D_{RG}}W^{(n)}\bigl(\{h(\lambda)\};x_1/\lambda,\ldots,x_n/\lambda\bigr)=0\,.$$

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,

$$\left[-\sum_\ell x_\ell\frac{\partial}{\partial x_\ell}-\sum_i\beta_i(h)\frac{\partial}{\partial h_i}+\frac{n}{2}\bigl(2-d-\eta(h)\bigr)\right]W^{(n)}(\{h\};x_1,\ldots,x_n)=0\,. \tag{25.25}$$

With this notation, a fixed point is defined by the common solution to all equations

$$\beta_i(h^*)=0\,,$$

and equation (25.25) then implies the scaling behaviour derived more directly.

25.2 The Gaussian Fixed Point

In a rather general framework, we have shown that universality can emerge as a consequence of the existence of RG fixed points in hamiltonian space. To be able to quantitatively describe the critical behaviour, we, therefore, have 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.

A subspace of hamiltonian space, however, can be explored completely: the subspace of quadratic hamiltonians (free field theories in the field theory language) which correspond 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 shown in Section 24.5, the weakly perturbed gaussian (or quasi-gaussian) model reproduces all results of the mean field approximation, which makes the study of gaussian models specially 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 Chapter 24, in terms of the Fourier components $\tilde{\phi}(q)$ of the field $\phi(x)$ the hamiltonian has a convergent expansion for small momentum:

$$\mathcal{H}_{\mathrm{G}}(\phi) = \frac{1}{2}\int \sum_{r=0} \tilde{\phi}(q) u_r^{(2)} q^{2r} \tilde{\phi}(-q) \mathrm{d}^d q \,. \tag{25.26}$$

This implies that the hamiltonian has a derivative expansion

$$\mathcal{H}_{\mathrm{G}}(\phi) = \frac{1}{2}\int \mathrm{d}^d x \sum_{r=0} \phi(x) u_r^{(2)} (-\nabla^2)^r \phi(x). \tag{25.27}$$

The simple RG transformations $\phi \mapsto \phi\sqrt{Z(\lambda)}$, $x \mapsto \lambda x$, and thus

$$\mathcal{H}_{\mathrm{G},\lambda}(\phi) = \frac{1}{2} Z(\lambda) \int \lambda^d \mathrm{d}^d x \sum_{r=0} \phi(x) u_r^{(2)} \lambda^{-2r} (-\nabla^2)^r \phi(x), \tag{25.28}$$

implies the relation (25.3) with $R^{(2)} = 0$,

$$W_\lambda^{(2)}(x) = Z^{-1}(\lambda) W^{(2)}(\lambda x).$$

Thus, the coefficient of the term with $2r$ derivatives becomes

$$u_r^{(2)} \mapsto u_r^{(2)}(\lambda) = Z(\lambda)\lambda^{d-2r} u_r^{(2)}(1). \tag{25.29}$$

For $\lambda \to \infty$, the terms with the smallest number of derivatives are the most important. For $u_0^{(2)} \neq 0$, the hamiltonian is non-critical. If we take for renormalization factor $Z(\lambda) = \lambda^{-d}$, we obtain the trivial fixed point

$$\mathcal{H}_{\mathrm{G}}^*(\phi) = \frac{1}{2} u_0^{(2)} \int \mathrm{d}^d x \, \phi^2(x).$$

The two-point correlation function has a δ-function limit ($d_\phi = d/2$) with a vanishing correlation length.

For $u_0^{(2)} = 0$ the hamiltonian is critical. If we take

$$Z(\lambda) = \lambda^{-(d-2)}, \Rightarrow d_\phi = \tfrac{1}{2}(d-2), \tag{25.30}$$

and thus

$$u_r^{(2)}(\lambda) = \lambda^{2-2r} u_r^{(2)}(1), \tag{25.31}$$

we find instead the so-called *gaussian fixed point*

$$\mathcal{H}_\mathrm{G}^*(\phi) = \frac{1}{2} u_1^{(2)} \int \mathrm{d}^d x \big(\nabla_x \phi(x)\big)^2, \tag{25.32}$$

which describes the asymptotic behaviour of the two-point function at T_c in mean field theory. Finally, if we take $u_0^{(2)} \neq 0$ but small enough such that when $\lambda^2 u_0^{(2)} = O(1)$ for all $r > 1$ $u_r^{(2)}(\lambda)$ is small, we find a universal two-point function of Ornstein–Zernicke form which describes the critical domain above T_c.

25.2.1 Eigenoperators

We now perform an eigenoperators analysis for the gaussian fixed point (25.32) within the subspace of *even* hamiltonians, that is, having an Ising-like symmetry. We consider perturbations $\Delta\mathcal{H}$ of the form of a general Landau hamiltonian, which can thus be expanded both in powers of ϕ and derivatives:

$$\Delta\mathcal{H}(\phi) = \sum_{n=2}^{\infty} \sum_{r=0}^{\infty} \sum_{\alpha} \mathcal{O}_\alpha^{n,r}(\phi),$$

where the operators $\mathcal{O}_\alpha^{n,r}(\phi)$ are integrals of monomials $V_\alpha^{n,r}(\phi)$ in $\phi(x)$ and its derivatives, of degree n (even) in ϕ and with exactly r derivatives (r also even):

$$\mathcal{O}_\alpha^{n,r}(\phi) = \int \mathrm{d}^d x\, V_\alpha^{n,r}\big(\phi(x), \partial_\mu \phi(x), \ldots\big).$$

The index α emphasizes that to a set of values n, r correspond in general several homogeneous polynomials.

The same RG transformations then yields

$$\mathcal{O}_\alpha^{n,r}(\lambda, \phi) = Z^{n/2}(\lambda) \lambda^{d-r} \mathcal{O}_\alpha^{n,r}(\phi) = \lambda^{d-n(d-2)/2-r} \mathcal{O}_\alpha^{n,r}(\phi). \tag{25.33}$$

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

$$\Delta\mathcal{H}(\phi) = \sum_{n=2}^{\infty} \sum_{r=0}^{\infty} \sum_{\alpha} h_\alpha^{(n,r)} \mathcal{O}_\alpha^{n,r}(\phi).$$

Then,

$$h_\alpha^{(n,r)}(\lambda) = \lambda^{d-n(d-2)/2-r} h_\alpha^{(n,r)}(1),$$

or in differential form

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} h_\alpha^{(n,r)} = L^* h_\alpha^{(n,r)} = l_{n,r} h_\alpha^{(n,r)}, \tag{25.34}$$

which provides an example of equation (25.17), with

$$l_{n,r} = d - \tfrac{1}{2}n(d-2) - r\,. \tag{25.35}$$

We can now classify all even operators:

(i) The operator $n = 2$, $r = 0$ is relevant and corresponds to a deviation from the critical temperature.

(ii) The operator $n = 2$, $r = 2$ is redundant, it corresponds to a simple renormalization of the dynamical variable.

(iii) Above dimension 4, all other operators are irrelevant: on the critical surface, the gaussian fixed point is stable. At dimension 4, $\int \phi^4(x)\mathrm{d}^d x$ $(n = 4, r = 0)$ is marginal and logarithmic corrections are expected.

Below dimension 4 $\int \phi^4(x)\mathrm{d}^d x$ 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 mean field approximation. It is also equivalent to power counting in quantum field theory as discussed in Chapter 9. The number of relevant operators is exactly the number of parameters on which depends the renormalized field theory.

Finally, among the odd operators, which break Ising symmetry, one is always relevant $n = 1, r = 0$ and thus $l_{1,0} = d/2 + 1$.

25.2.2 Beyond the gaussian fixed point

The gaussian fixed point is unstable for $d < 4$. Let us assume that in some sense, we can define models for continuous, non integer, dimension of space d. We then consider a critical hamiltonian in the neighbourhood of dimension 4, that is, for $\varepsilon = 4 - d$ small. If the dimensions of operators vary continuously with the dimension of space, 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 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

$$\lambda \frac{\mathrm{d}g(\lambda)}{\mathrm{d}\lambda} \equiv -\beta\big(g(\lambda)\big) = (4-d)g(\lambda) - \beta_2 g^2(\lambda) + O(g^3(\lambda)). \tag{25.36}$$

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) $\beta_2 < 0$. In this situation for $d < 4$ the first terms have the same sign and $g(\lambda)$ increases until the expansion becomes meaningless. Nothing can be concluded. Note that for $d > 4$ one finds a non-trivial repulsive fixed point

$$g^* = \varepsilon/\beta_2 + O(\varepsilon^2). \tag{25.37}$$

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(λ) increases and we again cannot conclude.

(ii) $\beta_2 = 0$. In this exceptional case, the term of order g^3 is needed.

(iii) $\beta_2 > 0$. In this case for $d > 4$ only the gaussian is stable. For $d < 4$ instead the gaussian fixed point is unstable, as we already know, but for ε small, another fixed point of the form (25.37) 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 a specially interesting situation and, as we shall see in Section 25.4, that it is realized in the $O(N)$ symmetric spin model, as first noticed by Wilson and Fisher. Then, universality is predicted and all universal quantities can be calculated in a power series in the deviation ε from dimension 4.

25.3 Critical Behaviour: The Effective ϕ^4 Field Theory

The main difficulty with the general RG approach is that it requires an explicit construction of renormalization group transformations for hamiltonians, which has a chance to lead to a fixed point. Although the general idea is to integrate over large momentum modes of the dynamical variables, its practical implementation is far from being straightforward. We instead write RG equations directly for correlation functions. The limitation of the method we use 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 25.2.2.

In Section 24.6, we have shown that, for an Ising-like system with short range ferromagnetic interactions, in the critical domain and for $d \le 4$, 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 just given by the first relevant terms of Landau–Ginzburg–Wilson's hamiltonian:

$$\mathcal{H}(\phi) \equiv \beta H = \int \mathrm{d}^d x \left[\frac{1}{2}c(\nabla\phi)^2 + \frac{1}{2}a\phi^2 + b\frac{1}{4!}\phi^4\right], \tag{25.38}$$

with a, b and c being *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 (25.38) should contain all relevant operators at least in some neighbourhood of dimension 4.

The hamiltonian (25.38) generates a perturbative expansion of field theory type which can be described in terms of Feynman diagrams. These have to be calculated with a cut-off of order unity, reflection of the initial microscopic structure. We shall eventually show that the precise cut-off procedure is irrelevant except that it should satisfy some general conditions. For example, the propagator can be modified (as in Pauli–Villars's regularization) but the Fourier transform of the inverse propagator must remain a regular function of momentum (the forces are short range).

In this chapter, we restrict ourselves to Ising-like systems, the field ϕ has only one component. Generalizations to the N-vector model with $O(N)$ symmetry, which is straightforward, and to several-component models with more than one ϕ^4 coupling constant will be briefly analysed in Section 26.6.

We have already seen in Chapter 24 that a convenient way to study the problem of infrared singularities 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.

Let us perform such a rescaling here, and rescale also the field $\phi(x)$ in such a way that the coefficient of $[\nabla\phi(x)]^2$ becomes the standard 1/2:

$$x \mapsto \Lambda x, \tag{25.39}$$

$$\phi(x) \mapsto \zeta\phi(x). \tag{25.40}$$

After this rescaling all quantities have a dimension in units of Λ. Our choice of normalization for the gradient term implies

$$\zeta = c^{-1/2}\Lambda^{1-d/2}, \tag{25.41}$$

which shows that ϕ now has in terms of Λ its canonical dimension $d/2 - 1$. The action $\mathcal{H}(\phi)$, then, becomes

$$\mathcal{H}(\phi) = \int \mathrm{d}^d x \left\{\frac{1}{2}\left[\nabla\phi(x)\right]^2 + \frac{1}{2} r\phi^2(x) + \frac{1}{4!} g\Lambda^{4-d}\phi^4(x)\right\} \tag{25.42}$$

with

$$r = a\Lambda^2/c, \qquad g = b/c^2. \tag{25.43}$$

Let us call r_c the parameter which corresponds, at g fixed, to the critical temperature T_c at which the correlation length ξ diverges. In terms of the scale Λ the critical domain is then defined by

$$\begin{gathered} \text{physical mass } = \xi^{-1} \ll \Lambda \ \Rightarrow\ |r - r_c| \ll \Lambda^2 \\ \text{distances} \gg 1/\Lambda \quad \text{or momenta} \ll \Lambda, \\ \text{magnetization } M \equiv \langle\phi(x)\rangle \ll \zeta^{-1} \sim \Lambda^{(d/2)-1}. \end{gathered} \tag{25.44}$$

Note that these conditions are met if Λ is identified with the cut-off of a usual field theory. However, an inspection of the action (25.42) also shows that, in contrast with conventional QFT, 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. In the usual formulation of QFT instead the *bare* coupling constant is also an adjustable parameter and the resulting QFT thus is less generic.

25.4 Renormalization Group Equations near Four Dimensions

The hamiltonian (25.42) can now be studied by field theoretical methods. Using a power counting argument, we have shown in Chapter 24 that the critical theory does not exist in perturbation theory for any dimension smaller than 4. If we define, by dimensional continuation, a critical theory in $d = 4 - \varepsilon$ dimensions, even for arbitrarily small ε there always exists an order in perturbation ($\sim 2/\varepsilon$) at which IR (infrared) divergences appear. Therefore, the idea, originally due to Wilson and Fisher, is to perform a double series expansion in powers of the coupling constant g and ε. Order by order in this expansion, the critical behaviour differs from the mean field behaviour only by powers of logarithm, and we can construct a perturbative critical theory by adjusting r to its critical value $r_c(T = T_c)$.

To study the large cut-off limit, we then use methods developed for the construction of the renormalized massless ϕ^4 field theory. We introduce rescaled correlation functions, defined by renormalization conditions at a new scale $\mu \ll \Lambda$, and functions of a renormalized coupling constant g_r:

$$\begin{aligned} \Gamma_r^{(2)}(p; g_r, \mu, \Lambda)|_{p^2=0} &= 0\,, \\ \frac{\partial}{\partial p^2}\Gamma_r^{(2)}(p; g_r, \mu, \Lambda)|_{p^2=\mu^2} &= 1\,, \\ \Gamma_r^{(4)}(p_i = \mu\theta_i; g_r, \mu, \Lambda) &= \mu^{\varepsilon} g_r\,, \end{aligned} \tag{25.45}$$

in which θ_i is a numerical vector. These correlation functions are related to the original ones by the equations:

$$\Gamma_r^{(n)}(p_i; g_r, \mu, \Lambda) = Z^{n/2}(g, \Lambda/\mu)\Gamma^{(n)}(p_i; g, \Lambda). \tag{25.46}$$

Renormalization theory (more precisely a slightly extended version adapted to the ε-expansion) tells us that the functions $\Gamma_r^{(n)}(p_i; g_r, \mu, \Lambda)$ of equation (25.46) have at p_i, g_r and μ fixed, a large cut-off limit which are the renormalized correlation functions $\Gamma_r^{(n)}(p_i; g_r, \mu)$. A detailed analysis actually shows

$$\Gamma_r^{(n)}(p_i; g_r, \mu, \Lambda) = \Gamma_r^{(n)}(p_i; g_r, \mu) + O(\Lambda^{-2}(\ln\Lambda)^L), \tag{25.47}$$

in which the power of $\ln\Lambda$ increases with the order in g and ε (see equation (25.53)). Moreover, the renormalized functions $\Gamma_r^{(n)}$ do not depend on the specific cut-off procedure and, given the normalization conditions (25.45), are, therefore, universal. Since the renormalized functions $\Gamma_r^{(n)}$ and the initial ones $\Gamma^{(n)}$ 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. We could, therefore, study only the renormalized field theory, which indeed is the only one really useful for explicit calculations of universal quantities (see Chapter 29). 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. For these reasons, it is useful to study the implications of equation (25.46) also directly in the initial theory.

Bare RG equations. Differentiating equation (25.46) with respect to Λ at g_r and μ fixed, and taking into account (25.47), we can derive a new identity

$$\Lambda\frac{\partial}{\partial\Lambda}\bigg|_{g_r,\mu \text{ fixed}} Z^{n/2}(g, \Lambda/\mu)\Gamma^{(n)}(p_i; g, \Lambda) = O(\Lambda^{-2}(\ln\Lambda)^L). \tag{25.48}$$

In the first part of our study, we neglect corrections subleading by powers of Λ. We shall return to this point in Chapter 27.

Then, using chain rule, we infer from equation (25.48),

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g, \Lambda/\mu)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g, \Lambda/\mu)\right]\Gamma^{(n)}(p_i; g, \Lambda) = 0\,. \tag{25.49}$$

To study the large cut-off limit, we then use methods developed for the construction of the renormalized massless ϕ^4 field theory. We introduce rescaled correlation functions, defined by renormalization conditions at a new scale $\mu \ll \Lambda$, and functions of a renormalized coupling constant g_r:

$$\begin{aligned} \Gamma_\mathrm{r}^{(2)}\,(p; g_\mathrm{r}, \mu, \Lambda)\big|_{p^2=0} &= 0\,, \\ \frac{\partial}{\partial p^2}\Gamma_\mathrm{r}^{(2)}\,(p; g_\mathrm{r}, \mu, \Lambda)\big|_{p^2=\mu^2} &= 1\,, \\ \Gamma_\mathrm{r}^{(4)}(p_i = \mu\theta_i; g_\mathrm{r}, \mu, \Lambda) &= \mu^{\varepsilon} g_\mathrm{r}\,, \end{aligned} \tag{25.45}$$

in which θ_i is a numerical vector. These correlation functions are related to the original ones by the equations:

$$\Gamma_\mathrm{r}^{(n)}(p_i; g_\mathrm{r}, \mu, \Lambda) = Z^{n/2}(g, \Lambda/\mu)\Gamma^{(n)}(p_i; g, \Lambda). \tag{25.46}$$

Renormalization theory (more precisely a slightly extended version adapted to the ε-expansion) tells us that the functions $\Gamma_\mathrm{r}^{(n)}(p_i; g_\mathrm{r}, \mu, \Lambda)$ of equation (25.46) have at p_i, g_r and μ fixed, a large cut-off limit which are the renormalized correlation functions $\Gamma_\mathrm{r}^{(n)}(p_i; g_\mathrm{r}, \mu)$. A detailed analysis actually shows

$$\Gamma_\mathrm{r}^{(n)}(p_i; g_\mathrm{r}, \mu, \Lambda) = \Gamma_\mathrm{r}^{(n)}(p_i; g_\mathrm{r}, \mu) + O(\Lambda^{-2}(\ln\Lambda)^L), \tag{25.47}$$

in which the power of $\ln\Lambda$ increases with the order in g and ε (see equation (25.53)). Moreover, the renormalized functions $\Gamma_\mathrm{r}^{(n)}$ do not depend on the specific cut-off procedure and, given the normalization conditions (25.45), are, therefore, universal. Since the renormalized functions $\Gamma_\mathrm{r}^{(n)}$ and the initial ones $\Gamma^{(n)}$ 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. We could, therefore, study only the renormalized field theory, which indeed is the only one really useful for explicit calculations of universal quantities (see Chapter 29). 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. For these reasons, it is useful to study the implications of equation (25.46) also directly in the initial theory.

Bare RG equations. Differentiating equation (25.46) with respect to Λ at g_r and μ fixed, and taking into account (25.47), we can derive a new identity

$$\Lambda\frac{\partial}{\partial\Lambda}\bigg|_{g_\mathrm{r},\mu\ \text{fixed}} Z^{n/2}(g, \Lambda/\mu)\Gamma^{(n)}(p_i; g, \Lambda) = O(\Lambda^{-2}(\ln\Lambda)^L). \tag{25.48}$$

In the first part of our study, we neglect corrections subleading by powers of Λ. We shall return to this point in Chapter 27.

Then, using chain rule, we infer from equation (25.48),

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g, \Lambda/\mu)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g, \Lambda/\mu)\right]\Gamma^{(n)}(p_i; g, \Lambda) = 0\,. \tag{25.49}$$

The functions β and η, which are dimensionless and may thus depend only on the dimensionless quantities g and Λ/μ, are defined by

$$\beta(g, \Lambda/\mu) = \Lambda \frac{\partial}{\partial \Lambda}\bigg|_{g_{\mathrm{r}},\mu} g\,, \tag{25.50a}$$

$$\eta(g, \Lambda/\mu) = -\Lambda \frac{\partial}{\partial \Lambda}\bigg|_{g_{\mathrm{r}},\mu} \ln Z(g, \Lambda/\mu). \tag{25.50b}$$

However, the functions β and η can also be directly calculated from equation (25.49) in terms of functions $\Gamma^{(n)}$ which do not depend on μ. Therefore, the functions β and η cannot depend on the ratio Λ/μ. Of course, if we examine the definitions (25.50) we see that the r.h.s. has a weak dependence in Λ/μ, but consistency requires that this dependence, which goes to zero like some power of μ/Λ, should be neglected, as in equation (25.48). Therefore, equation (25.49) simplifies as

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right] \Gamma^{(n)}(p_i; g, \Lambda) = 0\,. \tag{25.51}$$

Equation (25.51) is satisfied, when the cut-off is large, by the physical correlation functions of statistical mechanics which are also the bare correlation functions of quantum field theory. It is, as we have shown, a direct consequence of the existence of a renormalized theory. It will be implicit in the solution of equation (25.51) that it also implies the existence of a renormalized theory. Both statements are thus equivalent. We also note that equation (25.51) is a simplified form of equation (25.25) when only one operator is retained.

Beyond leading order. We finally characterize more precisely the terms neglected in equation (25.51). In a series expansion in powers of g and ε:

$$\Gamma^{(n)}(p_i; g, \Lambda) = \sum_{r,s} \Gamma^{(n)}_{rs} g^{r} \varepsilon^{s}, \tag{25.52}$$

the coefficients $\Gamma^{(n)}_{rs}$ have an asymptotic expansion for $|p_i|/\Lambda$ small of the form

$$\Gamma^{(n)}_{rs} = \sum_{l=0}^{L(n,r,s)} \left[(\ln\Lambda)^l A^{(1)}_{lrs} + \frac{1}{\Lambda^2}(\ln\Lambda)^l A^{(2)}_{lrs} + \cdots\right] \tag{25.53}$$

with $L(n,r,s) = 1 + r + s - n/2$ for $n = 4$ and $L(n,r,s) = r + s - n/2$ for $n \neq 4$.

The RG equation (25.51) is exact for the sum of the perturbative contributions which do not vanish for Λ large, as can be verified by expanding equation (25.51) in powers of g and ε.

25.5 Solution of the RG Equations: The ε-Expansion

Equation (25.51) can be solved by the method of characteristics. One introduces a dilatation parameter λ and looks for functions $g(\lambda)$ and $Z(\lambda)$ such that

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[Z^{-n/2}(\lambda)\Gamma^{(n)}\bigl(p_i; g(\lambda), \lambda\Lambda\bigr)\right] = 0 . \tag{25.54}$$

Differentiating explicitly with respect to λ, we find that equation (25.54) is consistent with equation (25.51) provided that

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} g(\lambda) = \beta\bigl(g(\lambda)\bigr), \qquad g(1) = g\,, \tag{25.55a}$$

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} \ln Z(\lambda) = \eta\bigl(g(\lambda)\bigr)\bigr), \qquad Z(1) = 1\,. \tag{25.55b}$$

The function $g(\lambda)$ is the effective coupling at the scale λ, and is governed by the flow equation (25.55a). Equation (25.54) implies

$$\Gamma^{(n)}(p_i; g, \Lambda) = Z^{-n/2}(\lambda)\Gamma^{(n)}\bigl(p_i; g(\lambda), \lambda\Lambda\bigr).$$

It is actually convenient to rescale Λ by a factor $1/\lambda$ and write the equation

$$\Gamma^{(n)}(p_i; g, \Lambda/\lambda) = Z^{-n/2}(\lambda)\Gamma^{(n)}\bigl(p_i; g(\lambda), \Lambda\bigr). \tag{25.56}$$

Equations (25.55) and (25.56) implement approximately (because terms subleading by powers of Λ have been neglected) the RG ideas as presented in Section 25.1: since the coupling constant $g(\lambda)$ characterizes the hamiltonian $\mathcal{H}_\lambda$, equation (25.55a) is the equivalent of equation (25.11) (up to the change $\lambda \mapsto 1/\lambda$); equations (25.12) and (25.55b) differ only by the definition of $Z(\lambda)$.

The solutions of equations (25.55) can be written as

$$\int_g^{g(\lambda)} \frac{\mathrm{d}g'}{\beta(g')} = \ln\lambda\,, \tag{25.57a}$$

$$\int_1^{\lambda} \frac{\mathrm{d}\sigma}{\sigma}\eta\bigl(g(\sigma)\bigr) = \ln Z(\lambda). \tag{25.57b}$$

Equation (25.51) is the renormalization group equation in differential form. Equations (25.56) and (25.57) are the integrated RG equations. In what follows we explicitly assume that the RG functions $\beta(g)$ and $\eta(g)$ are regular functions of g for $g \geq 0$. In equation (25.56), we see that it is equivalent to increase Λ or to decrease λ. To investigate the large Λ limit we, therefore, study the behaviour of the effective coupling constant $g(\lambda)$ when λ goes to zero. Equation (25.57a) 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 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 exactly a fixed point value. Conversely those where the slope is positive are IR attractive.

The RG functions β and η can be calculated in perturbation theory. The relation between bare and renormalized coupling constant

$$\mu^{\varepsilon} g_{\mathrm{r}} = \Lambda^{\varepsilon} \left[g - \frac{3g^2}{16\pi^2} (\ln(\Lambda/\mu) + \text{ const.}) \right] + O(g^3, g^2\varepsilon), \tag{25.58}$$

($\varepsilon = 4 - d$) and the definition (25.50a) imply

$$\beta(g, \varepsilon) = -\varepsilon g + \frac{3g^2}{16\pi^2} + O(g^3, g^2\varepsilon). \tag{25.59}$$

Equations (25.55a) and (25.59) combined, have exactly the form (25.36) (still with $\lambda \mapsto 1/\lambda$), and the discussion now is similar. We see from the explicit expression (25.59) of the β-function that above four dimensions, that is, $\varepsilon < 0$, if initially g is small, $g(\lambda)$ decreases approaching the origin $g = 0$. We recover that the gaussian fixed point is IR stable.

Below four dimensions, if initially g is very small, expression (25.59) shows that $g(\lambda)$ first increases, a behaviour reflecting the instability of the gaussian fixed point.

However, the explicit expression (25.59) shows that, in the sense of an expansion in powers of ε, $\beta(g)$ has another zero g^*:

$$\beta(g^*) = 0 \quad \text{for} \quad g^* = 16\pi^2\varepsilon/3 + O(\varepsilon^2) \tag{25.60}$$

with a positive slope for ε infinitesimal:

$$\omega \equiv \beta'(g^*) = \varepsilon + O(\varepsilon^2) > 0\,. \tag{25.61}$$

This agrees with one of the possibilities envisaged in Section 25.2.2. Then, equation (25.57a) shows that $g(\lambda)$ has g^* as an asymptotic limit. Linearizing equation (25.57a) around the fixed point we find

$$\int_g^{g(\lambda)} \frac{\mathrm{d}g'}{\omega(g' - g^*)} \sim \ln \lambda\,. \tag{25.62}$$

Integrating, we get

$$|g(\lambda) - g^*| \underset{\lambda \to 0}{=} O(\lambda^{\omega}). \tag{25.63}$$

Below dimension 4, at least for ε infinitesimal, this non-gaussian fixed point is IR stable. In dimension 4 this 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 25.2.

Let us now assume that $\Gamma^{(n)}(g^*)$ and $\eta(g^*)$ are finite, conditions which are satisfied within the framework of the ε-expansion. From equation (25.57b), we derive the behaviour of $Z(\lambda)$ for λ small. The integral in the l.h.s. is dominated by small values of σ. It follows that

$$\ln Z(\lambda) \underset{\lambda \to 0}{\sim} \eta \ln \lambda\,, \tag{25.64}$$

where we have set

$$\eta = \eta(g^*).$$

Equation (25.56) then determines the behaviour of $\Gamma^{(n)}(p_i; g, \Lambda)$ for Λ large:

$$\Gamma^{(n)}(p_i; g, \Lambda/\lambda) \sim \lambda^{-n\eta/2}\Gamma^{(n)}(p_i; g^*, \Lambda). \tag{25.65}$$

On the other hand, from simple dimensional considerations, we know that

$$\Gamma^{(n)}(p_i; g, \Lambda/\lambda) = \lambda^{-d+(n/2)(d-2)}\Gamma^{(n)}(\lambda p_i; g, \Lambda). \tag{25.66}$$

Combining this equation with equation (25.65), we obtain

$$\Gamma^{(n)}(\lambda p_i; g, \Lambda) \underset{\lambda\to 0}{\sim} \lambda^{d-(n/2)(d-2+\eta)}\Gamma^{(n)}(p_i; g^*, \Lambda). \tag{25.67}$$

This equation shows that the critical 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 other zero.

Equation (25.67) for $n = 2$ yields the small momentum behaviour of the inverse two-point function, and thus of the correlation function

$$W^{(2)}(p) = \left[\Gamma^{(2)}(p)\right]^{-1} \underset{|p|\to 0}{\sim} 1/p^{2-\eta}. \tag{25.68}$$

One verifies that the definition of equation (25.64) coincides with the usual definition of the critical exponent η. The spectral representation of the two-point function (Section 6.9) 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 short calculation yields

$$\eta = \frac{\varepsilon^2}{54} + O(\varepsilon^3). \tag{25.69}$$

A semi-quantitative prediction of the theory is that η is numerically small in three dimensions.

Finally, we note that equation (25.67) can be interpreted by saying that the field $\phi(x)$, which had at the gaussian fixed point a canonical dimension $(d-2)/2$, has now acquired an anomalous dimension d_ϕ (equation (25.14)):

$$d_\phi = \tfrac{1}{2}(d - 2 + \eta).$$

All 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 long distance behaviour different from the one predicted by mean field theory. 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 Chapter 24 remains valid and thus it does not depend on any other coupling either (this point will be further discussed in Section 27.4). Therefore, the critical behaviour is ***universal***, although less universal than in MFT, in the sense that it depends only on some small number of qualitative features of the system under consideration.

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 (25.51), define implicitly a one parameter family of critical hamiltonians which correspond to a 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 fixed dimension perturbation theory are generated by an expansion around the wrong, since IR repulsive, fixed point. The ε-expansion allows us to interchange limits and to follow perturbatively the attractive IR fixed point.

25.6 Critical Correlation Functions with $\phi^2(x)$ Insertions

In Chapter 26, we will study the whole critical domain (25.44). In particular, following a method already explained in Section 10.10, correlation functions for $T \neq T_c$ will be expanded in terms of *critical* correlation functions with $\int \mathrm{d}^d x\, \phi^2(x)$ insertions. The operator $\phi^2(x)$ has here a direct physical interpretation. It is the most singular part (i.e. the most relevant) of the energy density (25.38).

Therefore, we now discuss the long 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 $\Gamma^{(l,n)}(q_1,\ldots,q_l;p_1,\ldots,p_n;g,\Lambda)$ the 1PI functions of $n\,\phi(x)$ fields and $l\,\frac{1}{2}\phi^2(x)$ operators, (with $(l+n) \geq 2$). Renormalization theory tells us that we can define renormalized correlation functions $\Gamma_{\mathrm{r}}^{(l,n)}(q_i;p_j;g_{\mathrm{r}},\mu)$ which in addition to the conditions (25.45) satisfy

$$\begin{gathered}
\Gamma_{\mathrm{r}}^{(1,2)}(q;p_1,p_2;g_{\mathrm{r}},\mu)\Big|_{p_1^2=p_2^2=\mu^2,\ p_1\cdot p_2=-\frac{1}{3}\mu^2} = 1\,, \\
\Gamma_{\mathrm{r}}^{(2,0)}(q,-q;g_{\mathrm{r}},\mu)\Big|_{q^2=\frac{4}{3}\mu^2} = 0\,,
\end{gathered} \tag{25.70}$$

and are related to the initial ones (see equation (10.70)) by

$$\lim_{\Lambda\to\infty} Z^{n/2}\,(Z_2/Z)^l \left[\Gamma^{(l,n)}(q_i;p_j;g,\Lambda) - \delta_{n0}\delta_{l2}\Lambda^{-\varepsilon}A\right] = \Gamma_{\mathrm{r}}^{(l,n)}(q_i;p_j;g_{\mathrm{r}},\mu). \tag{25.71}$$

$Z_2(g,\Lambda/\mu)$ and $A(g,\Lambda/\mu)$ are two new renormalization constants.

Differentiating with respect to Λ at g_{r} and μ fixed, as in Section 25.4, and using chain rule we obtain a set of RG equations:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g) - l\eta_2(g)\right]\Gamma^{(l,n)} = \delta_{n0}\delta_{l2}\Lambda^{-\varepsilon}B(g). \tag{25.72}$$

In addition to β and η (equations (25.50)), two new RG functions, $\eta_2(g)$ and $B(g)$, appear:

$$\eta_2(g) = -\Lambda\left.\frac{\partial}{\partial\Lambda}\right|_{g_{\mathrm{r}},\mu} \ln\left[Z_2(g,\Lambda/\mu)\,/Z(g,\Lambda/\mu)\right], \tag{25.73}$$

$$B(g) = \left[\Lambda\left.\frac{\partial}{\partial\Lambda}\right|_{g_{\mathrm{r}},\mu} - 2\eta_2(g) - \varepsilon\right]A(g,\Lambda/\mu). \tag{25.74}$$

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 (25.72) 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 12, and we want here only to warn the reader against a too naive application of RG ideas.

In the homogeneous case, equation (25.72) can be solved by the method of characteristics explained in Section 25.5, exactly in the same way as equation (25.49). With the RG function $\eta_2(g)$ is associated a new scale-dependent function $\zeta_2(\lambda)$:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[Z^{-n/2}(\lambda)\zeta_2^{-l}(\lambda)\Gamma^{(l,n)}(q_i;p_j;g(\lambda),\lambda\Lambda)\right]=0\,,$$

which, therefore, satisfies

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln\zeta_2(\lambda)=\eta_2[g(\lambda)].$$

It follows from the arguments of Section 25.5 that the critical behaviour of $\Gamma^{(l,n)}$ is governed by the IR fixed point g^*.

To relate the RG function $\eta_2(g)$ to standard exponents, we also introduce the function $\nu(g)$:

$$1/\nu(g)=\eta_2(g)+2\,. \tag{25.75}$$

For λ small ζ_2 behaves like

$$\ln\zeta_2\sim(1/\nu-2)\ln\lambda\,,$$

where the new critical exponent ν is defined by

$$\nu=\nu(g^*).$$

Combining again the solution of equation (25.72) with simple dimensional analysis, we finally obtain the behaviour of $\Gamma^{(l,n)}$:

$$\Gamma^{(l,n)}(\lambda q_i;\lambda p_j;g,\Lambda)\underset{\lambda\to 0}{\propto}\lambda^{d-n(d-2+\eta)/2-l/\nu}. \tag{25.76}$$

Using equations (25.73,25.75), it is straightforward to calculate $\nu(g)$ at one-loop order:

$$2\nu(g)=1+\frac{g}{32\pi^2}+O(g^2), \tag{25.77}$$

and then, setting $g=g^*$ (equation (25.60)), the exponent ν at first order in ε:

$$2\nu=1+\frac{\varepsilon}{6}+O(\varepsilon^2).$$

The $\langle\phi^2\phi^2\rangle$ correlation function. The ϕ^2 (energy density) two-point function $\Gamma^{(2,0)}$ satisfies an inhomogeneous RG equation. To solve it, one first looks for a particular solution, which can be chosen of the form $\Lambda^{-\varepsilon}C_2(g)$:

$$\beta(g)C_2'(g)-[\varepsilon+2\eta_2(g)]\,C_2(g)=B(g). \tag{25.78}$$

The solution is uniquely determined by imposing its *regularity* at $g=g^*$.

The general solution of equation (25.72) then is the sum of this particular solution and the general solution of the homogeneous equation which has a behaviour given by equation (25.76):

$$\Gamma^{(2,0)}(q;g,\Lambda)-\Lambda^{-\varepsilon}C_2(g)\underset{q\to 0}{\sim}K\Lambda^{2/\nu-4}q^{d-2/\nu}. \tag{25.79}$$

Note that the regular contribution depends on g but not the constant K.

It now remains to study the behaviour of correlation functions in the critical domain (25.44) away from T_c and this will be the subject of Chapter 26.

The dimension of the ϕ^2 operator. The scaling behaviour (25.76) determines the dimension of the operator ϕ^2 at the IR fixed point. More 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

$$\left\langle \prod_i \mathcal{O}_i(y_i/\lambda) \prod_j \phi(x_j/\lambda) \right\rangle \underset{\lambda\to 0}{\propto} \lambda^D , \quad \text{with } D = nd_\phi + \sum_i d_i \,.$$

After Fourier transformation, factorization of the δ-function of momentum conservation and Legendre transformation with respect to the field ϕ one finds (neglecting possible regular terms)

$$\Gamma^{(n)}_{\mathcal{O}_1,\dots,\mathcal{O}_l}(\lambda q_i, \lambda p_j) \propto \lambda^{-D'}$$

with $D' = D - (n+l-1)d - n(2d_\phi - d) = d - nd_\phi + \sum_i d_i - d$. It follows in particular that $d_{\phi^2} = d - 1/\nu$.

Remark. The physics we intend to describe corresponds to integer values of ε, $\varepsilon = 1, 2$. Although we can only prove the validity of all RG results within the framework of the ε-expansion, we will eventually assume that their validity extends beyond an infinitesimal neighbourhood of dimension 4. The comparison of the results obtained from the summed expansion with experimental or numerical data thus provides a crucial test of the theory (see Chapter 29).

The fixed point hamiltonian. Let us note that the fixed point hamiltonian $\mathcal{H}^*$ has never been explicitly constructed for any non-gaussian theory. Indications from field theory are that such a construction is not easy. 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 $W^{(n)}_k$ converge towards fixed point correlation functions $W^{(n)}_*$ in the following way:

$$\left| W^{(n)}_k(\lambda q_i) - W^{(n)}_*(\lambda q_i) \right| = O(\lambda^{2k}).$$

In Sections 25.3–25.5, we have constructed $\mathcal{H}_1$.

Quantum field theory in particle physics. Let us now comment about the significance of these results for field theory in the context of particle physics. From the preceding analysis, we understand that the existence of a renormalizable field theory, beyond perturbation theory, relies on two properties:

(i) Mass renormalization must make sense beyond perturbation theory (the existence of a critical temperature), and must be achieved naturally.

(ii) An IR fixed point in the RG sense must exist, to ensure that the long distance physics is short-distance insensitive and can be described by renormalized quantum field theory.

An inspection of the action (25.38), however, shows that, in contrast with conventional QFT, the ϕ^4 coupling constant in the critical phenomena applications 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. In the usual formulation of QFT instead the *bare* coupling constant is considered as an adjustable parameter: A 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 in the theory (called μ^{-1} in the text), much larger than the microscopic scale, in such a way that it is possible to define, in addition to the universal long distance physics, a short distance or large momentum physics with $\mu \ll |p_i| \ll \Lambda$.

Note, however, that if we demand to particle physics theories to be *natural*, 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.

Finally, the RG analysis indicates that in four dimensions, in the domain of attraction of the $g = 0$ IR fixed point, the renormalized coupling constant of the ϕ^4 field theory goes to zero logarithmically when the cut-off becomes infinite. Since no other fixed point seems to exist, this leads to the so-called *triviality problem.*

Bibliographical Notes

The modern formulation of the RG ideas is due to

K.G. Wilson, *Phys. Rev.* B4 (1971) 3174, *ibidem* 3184;

and presented in an expanded form in

K.G. Wilson and J. Kogut, *Phys. Rep.* 12C (1974) 75.

The idea of RG transformations was earlier proposed in a simplified form in

L.P. Kadanoff, *Physics* 2 (1966) 263.

The systematic classification of operators can be found in

F.J. Wegner, *Phys. Rev.* B5 (1972) 4529, B6 (1972) 1891, and in Wegner's contribution to *Phase Transitions and Critical Phenomena,* vol. 6, C. Domb and M.S. Green eds. (Academic Press, London 1976).

The idea of the ε-expansion is due to

K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* 28 (1972) 240.

After Wilson's original articles, several authors have realized that RG methods, as developed in Quantum Field Theory, could be applied to Critical Phenomena

C. Di Castro, *Lett. Nuovo Cimento.* 5 (1972) 69; G. Mack, *Kaiserslautern 1972*, Lecture Notes in Physics vol. 17, W. Ruhl and A. Vancura eds. (Springer-Verlag, Berlin 1972); E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D8 (1973) 434, *ibidem* 2418; P.K. Mitter, *Phys. Rev.* D7 (1973) 2927; G. Parisi, *Cargèse Lectures 1973*, published in *J. Stat. Phys.* 23 (1980) 49; B. Schroer, *Phys. Rev.* B8 (1973) 4200; C. Di Castro, G. Jona-Lasinio and L. Peliti, *Ann. Phys. (NY)* 87 (1974) 327; F. Jegerlehner and B. Schroer, *Acta Phys. Austr. Suppl.* XI (1973) 389 (Springer-Verlag, Berlin).

The RG equations for the bare theory were first derived in

J. Zinn-Justin, *Cargèse Lectures 1973*, unpublished, incorporated in the contribution of E. Brézin, J.C. Le Guillou and J. Zinn-Justin to *Phase Transitions and Critical Phenomena* vol. 6, quoted in Chapter 24.

26 SCALING BEHAVIOUR IN THE CRITICAL DOMAIN

In Chapter 25, we have established the scaling behaviour of correlation functions at criticality, $T = T_c$. 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.

Using results proven in Chapter 10, 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, this formalism does not allow to cross the critical temperature and thus to describe the whole critical domain. Therefore, we then expand correlation functions in formal power series of the deviation $t = (T - T_c)/T_c$ 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 Chapter 10), whose scaling behaviour has been derived in Section 25.6. Summing the expansion, we obtain RG equations valid for $T \neq T_c$.

To cross the critical temperature, avoiding the critical singularities, it is in addition necessary to explicitly break the symmetry of the hamiltonian. We thus add a small magnetic field to the spin interactions. We then derive RG equations in a field, or at fixed magnetization. In this way, we are able to connect continuously correlation functions above and below T_c, and establish scaling laws in the whole critical domain.

In the first part of the chapter, we restrict ourselves to Ising-like systems, we then generalize the results to N-component order parameters in Section 26.6. In Section 26.7, we show how to expand the universal two-point function when T approaches T_c, using the short distance expansion. Finally, in the appendix, we discuss the energy correlation function when the specific heat exponent α vanishes.

Remark. 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 25.2, the most relevant contribution (the most IR singular) corresponds to the $\phi^2(x)$ operator. We can, therefore, take the difference $t = r - r_c$ between the coefficient of ϕ^2 in (25.42) and its critical value as a linear measure of the deviation from the critical temperature and parametrize the effective hamiltonian as

$$\mathcal{H}(\phi) = \int \left\{ \frac{1}{2} \left[\partial_\mu \phi(x)\right]^2 + \frac{1}{2}(r_c + t)\phi^2(x) + \frac{1}{4!} g\Lambda^\varepsilon \phi^4(x) \right\} \mathrm{d}^d x \,. \tag{26.1}$$

Dimensional analysis then yields the relation

$$\Gamma^{(n)}(p_i; t, g, \Lambda) = \Lambda^{d-n(d-2)/2} \Gamma^{(n)}(p_i \Lambda^{-1}; t\Lambda^{-2}, g, 1). \tag{26.2}$$

With this parametrization, the critical domain corresponds to

$$|t = r - r_c| \ll \Lambda^2 \,.$$

26.1 Strong Scaling above T_c: The Renormalized Theory

We discuss in this section various 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 theory with renormalization conditions imposed at zero momentum. Generalizing to any dimension $d \le 4$ the formalism of Chapter 10, we define renormalized correlation functions by

$$\Gamma_{\mathrm{r}}^{(2)}(p; m_{\mathrm{r}}, g_{\mathrm{r}}) = m_{\mathrm{r}}^2 + p^2 + O(p^4), \qquad (26.3)$$

$$\Gamma_{\mathrm{r}}^{(4)}(p_i = 0; m_{\mathrm{r}}, g_{\mathrm{r}}) = m_{\mathrm{r}}^{\varepsilon} g_{\mathrm{r}}\,. \qquad (26.4)$$

This renormalized theory is obtained from the original "bare" theory by taking the large cut-off limit at g_{r} and m_{r} fixed:

$$\Gamma_{\mathrm{r}}^{(n)}(p_i; m_{\mathrm{r}}, g_{\mathrm{r}}) = \lim_{\Lambda \to \infty} Z^{n/2}(m_{\mathrm{r}}/\Lambda, g_{\mathrm{r}}) \Gamma^{(n)}(p_i; t, g, \Lambda). \qquad (26.5)$$

The condition (26.4) is written in such a way that g_{r} is a dimensionless parameter. A simple dimensional analysis then yields for the renormalized correlation functions, the relation

$$\Gamma_{\mathrm{r}}^{(n)}(p_i; m_{\mathrm{r}}, g_{\mathrm{r}}) = m_{\mathrm{r}}^{d-(n/2)(d-2)} \Gamma_{\mathrm{r}}^{(n)}(p_i/m_{\mathrm{r}}; 1, g_{\mathrm{r}}). \qquad (26.6)$$

Therefore, the correlation length ξ, which characterizes the decay of the connected two-point function, is proportional to m_{r}^{-1}: it fixes the scale of the zero momentum renormalized theory.

Note that equation (26.5) holds for any dimension $d \le 4$ and, therefore, some consequences of this equation are valid beyond the ε-expansion. This is one reason why the consideration of this particular renormalized theory is so useful.

As we have shown in Chapter 10, the renormalized correlation functions satisfy CS equations:

$$\left[m_{\mathrm{r}} \frac{\partial}{\partial m_{\mathrm{r}}} + \beta(g_{\mathrm{r}}) \frac{\partial}{\partial g_{\mathrm{r}}} - \frac{n}{2}\eta(g_{\mathrm{r}}) \right] \Gamma_{\mathrm{r}}^{(n)}(p_i; m_{\mathrm{r}}, g_{\mathrm{r}})$$
$$= m_{\mathrm{r}}^2 \left[2 - \eta(g_{\mathrm{r}})\right] \Gamma_{\mathrm{r}}^{(1,n)}(0; p_i; m_{\mathrm{r}}, g_{\mathrm{r}}), \qquad (26.7)$$

in which the renormalized ϕ^2 insertions are specified by the condition

$$\Gamma_{\mathrm{r}}^{(1,2)}(0; 0, 0; m_{\mathrm{r}}, g_{\mathrm{r}}) = 1\,. \qquad (26.8)$$

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 (26.7) in Section 26.1.3.

26.1.1 Bare and renormalized coupling constants

We first discuss the problem of the bare and renormalized coupling constants. Several remarks are here in order:

(i) Since in dimension $d < 4$ the ϕ^4 field theory is super-renormalizable, all correlation functions have a large cut-off limit after a simple mass renormalization, that is, after one has taken the deviation from the critical temperature as a parameter. The coupling constant and the field amplitude renormalizations are finite and, therefore, are not required in general. If we only renormalize the mass, we obtain a finite theory which is a function of the bare coupling constant u, the coefficient of $\phi^4(x)$ in the hamiltonian. The coupling constant u, however, has a mass dimension ε. Therefore, perturbation theory is really an expansion in powers of $u/m_{\mathrm{r}}^{\varepsilon}$. It only makes sense if this ratio is kept fixed, which implies that the coupling constant u (which characterizes the deviation from the gaussian fixed point) goes to zero with the inverse correlation length as $m_{\mathrm{r}}^{\varepsilon}$. This is indeed what is implicitly assumed in the conventional field theory framework. In the critical phenomena situation instead, the coupling constant, which is related to microscopic parameters of the theory, is fixed. This means, as we already stressed in Chapter 25, that after introduction by rescaling of the cut-off Λ, $g = u/\Lambda^{\varepsilon}$ remains finite (see action (25.42)) when $\Lambda \to \infty$. Therefore, in the critical domain, that is, in the large cut-off limit, u 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 amplitude renormalization and a new expansion parameter g_{r} which, as we verify below, 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 way: $m \ll u^{1/\varepsilon} \ll \Lambda$ instead of $m \ll \Lambda \propto u^{1/\varepsilon}$. We have to assume that the result is the same at leading order.

The renormalized coupling constant. The function $\beta(g_{\mathrm{r}})$ in (26.7) is given by equation (10.32):

$$\beta(g_{\mathrm{r}}) = m_{\mathrm{r}} \left.\frac{\mathrm{d}}{\mathrm{d}m_{\mathrm{r}}}\right|_{\Lambda,g} g_{\mathrm{r}}(g, m_{\mathrm{r}}/\Lambda). \tag{26.9}$$

Let us introduce the variable $\lambda = m_{\mathrm{r}}/\Lambda$. Considering g_{r} as a function of λ at g *fixed*, we can rewrite equation (26.9):

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} g_{\mathrm{r}}(\lambda) = \beta(g_{\mathrm{r}}(\lambda)). \tag{26.10}$$

This equation, which is similar to equation (25.55*a*), is a flow equation for $g_{\mathrm{r}}(\lambda)$. 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 $\beta(g_{\mathrm{r}})$ if such a zero exists. Assuming the existence of such an IR fixed point:

$$\beta(g_{\mathrm{r}}^*) = 0\,, \qquad \text{with} \quad \omega = \beta'(g_{\mathrm{r}}^*) > 0\,, \tag{26.11}$$

we can integrate equation (26.10) in the neighbourhood of g_{r}^*, and estimate $g_{\mathrm{r}} - g_{\mathrm{r}}^*$:

$$|g_{\mathrm{r}} - g_{\mathrm{r}}^*| \sim (m_{\mathrm{r}}/\Lambda)^{\omega}. \tag{26.12}$$

Therefore, as a consequence of the conditions imposed by the physical origin of the problem (the bare coupling constant is associated with microscopic physics), in the critical domain the renormalized or effective coupling constant is close to the IR fixed point value.

26.1.2 Strong scaling

At leading order, we can then replace g_{r} by g_{r}^*. Let us evaluate the behaviour of the renormalization constant Z (equation (10.33)):

$$m_{\mathrm{r}} \left.\frac{\mathrm{d}}{\mathrm{d}m_{\mathrm{r}}}\right|_{g,\Lambda} \ln Z(g_{\mathrm{r}}, m_{\mathrm{r}}/\Lambda) = \eta(g_{\mathrm{r}}). \tag{26.13}$$

Using the parameter $\lambda = m_\mathrm{r}/\Lambda$ and the function $g_\mathrm{r}(\lambda)$ of equation (26.10), we can rewrite equation (26.13) as

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} \ln Z(\lambda) = \eta(g_\mathrm{r}(\lambda)). \tag{26.14}$$

In the case of an IR fixed point, we find for $\lambda = m_\mathrm{r}/\Lambda$ small:

$$Z \sim (m_\mathrm{r}/\Lambda)^\eta, \tag{26.15}$$

where we have called η the value (assumed finite) of the function $\eta(g)$ at the fixed point:

$$\eta = \eta(g_\mathrm{r}^*). \tag{26.16}$$

Let us also assume that the renormalized correlation functions $\Gamma_\mathrm{r}^{(n)}(p_i; m_\mathrm{r}, g_\mathrm{r}^*)$ are finite. Combining then equations (26.5,26.6,26.15), we find in the critical domain the scaling relation:

$$\Gamma^{(n)}(p_i, t, g, \Lambda = 1) \underset{m_\mathrm{r} \ll 1,\ |p_i| \ll 1}{\sim} m_\mathrm{r}^{d-(n/2)(d-2+\eta)} \Gamma_\mathrm{r}^{(n)}(p_i/m_\mathrm{r}; 1, g_\mathrm{r}^*). \tag{26.17}$$

The parameter Λ has only been introduced to provide a scale for all quantities. Once the form of the critical behaviour has been obtained, we can rescale all dimensional parameters (here momenta, deviation from T_c, and correlation length) to eliminate Λ, as we have done above (see equation (26.2)).

Equation (26.17) provides a proof of *strong scaling* and *universality* in the whole critical domain above T_c. The initial correlation functions depend on the detailed form of the interaction. The ϕ^4 theory bare correlation functions depend only on the way the cut-off is introduced and explicitly on p_i, g, and t. The r.h.s. of equation (26.17) involves only the renormalized correlation functions at $g_\mathrm{r} = g_\mathrm{r}^*$, that is, functions of ratios p_i/m_r. Moreover, the result (26.17) holds for any fixed dimension $d \le 4$. Resort 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 11, we can immediately obtain $\beta(g_\mathrm{r})$ at leading order and calculate g_r^* and ω:

$$g_\mathrm{r}^* = \frac{3}{16\pi^2}\varepsilon + O(\varepsilon^2), \quad \omega = \varepsilon + O(\varepsilon^2). \tag{26.18}$$

However, as first suggested by Parisi, it is also possible to analyse numerically the perturbative expansion in powers of g_r at fixed dimension 3 or 2. Such an analysis convincingly demonstrates the existence of an IR fixed point and allows a precise determination of critical exponents as will be shown in Chapter 29 when we discuss numerical results.

Note finally that equations (10.34–10.38) 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 will be derived in Section 26.3 by another, simpler, method. It confirms that the relation between m_r and t is singular. Therefore, the method of this section, based on the introduction of a zero-momentum renormalized theory, does not allow to cross the critical temperature. Indeed, all correlation functions are parametrized in terms of the correlation length which is singular at T_c. To avoid this problem, we introduce in the next section a different formalism which is a natural extension of the formalism of Chapter 25.

Let us conclude this discussion with a few remarks.

(i) We have shown that the renormalized coupling constant g_{r} has a finite limit $g_{\mathrm{r}} = g_{\mathrm{r}}^*$ although the bare coupling $u = g\Lambda^{\varepsilon}$ becomes infinite. In addition, precisely at g_{r}^* the field amplitude renormalization diverges as equation (26.15) shows. The conclusion is that the IR fixed point field theory behaves even below four dimensions as a renormalizable field theory, and a complete renormalization is indeed required.

(ii) A second somewhat related remark is that when u varies from zero to infinity, g_{r} varies from zero to g_{r}^*. This property seems to indicate that it is impossible to construct a theory with $g_{\mathrm{r}} > g_{\mathrm{r}}^*$. Since $g_{\mathrm{r}}^* = O(\varepsilon)$ this argument suggests that it is somewhat unlikely that a non-trivial ϕ^4 field theory exists in four dimensions. On the other hand, to construct the renormalized theory, we have taken the large cut-off limit at u fixed. This procedure is only legitimate if the bare renormalization group has only one IR fixed point. Otherwise, other non-trivial theories might be obtained by sending the cut-off and u to infinity at g fixed. This point will be further discussed in Section 35.1.1.

26.1.3 Large momentum behaviour

We now return to equation (26.7) to investigate the behaviour of correlation functions at large momenta, in the critical domain $m_{\mathrm{r}} \ll |p_i| \ll \Lambda$. At leading order we can replace in equation (26.7) g_{r} by its IR fixed point value g_{r}^*:

$$\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}} - \frac{n}{2}\eta\right)\Gamma_{\mathrm{r}}^{(n)}(p_i; m_{\mathrm{r}}, g_{\mathrm{r}}^*) = m_{\mathrm{r}}^2(2-\eta)\Gamma_{\mathrm{r}}^{(1,n)}(0; p_i; m_{\mathrm{r}}, g_{\mathrm{r}}^*). \tag{26.19}$$

It follows from the scaling relation (26.6) or (26.17) that the large momentum behaviour is directly related to the approach to the critical theory $m_{\mathrm{r}} = 0$. It is thus not too surprising that in order to extract some information from equation (26.19) it is necessary to return to the framework of the ε-expansion. A simple extension of Weinberg's theorem shows that, order by order in g_{r} and ε, the r.h.s. of equation (26.19) is negligible at large (non-exceptional) momenta, or small masses, as one might naively guess from the factor m_{r}^2 which has been factorized. By contrast below four dimensions, at fixed dimension, there always exists an order at which the r.h.s. ceases to be negligible. This is interpreted as a consequence of expanding perturbation theory around the wrong fixed point, the IR unstable gaussian fixed point. The true asymptotic behaviour of correlation functions is different because it is governed by the non-trivial point $g_{\mathrm{r}} = g_{\mathrm{r}}^*$. Asymptotically, the $\Gamma_{\mathrm{r}}^{(n)}$, therefore, satisfy

$$\left(m_{\mathrm{r}}\frac{\partial}{\partial m_{\mathrm{r}}} - \frac{n}{2}\eta\right)\Gamma_{\mathrm{r}}^{(n)}(p_i, m_{\mathrm{r}}) \underset{|p_i|\gg m_{\mathrm{r}}}{\approx} 0\,. \tag{26.20}$$

Combined with the relation (26.6), this leads to

$$\Gamma_{\mathrm{r}}^{(n)}(\lambda p_i; m_{\mathrm{r}}) \underset{|p_i|\gg m_{\mathrm{r}}}{\sim} \lambda^{d-(n/2)(d-2+\eta)}\Gamma_{\mathrm{r}}^{(n)}(p_i; m_{\mathrm{r}}). \tag{26.21}$$

Using then the relation (26.5) between bare and renormalized correlation functions, we recover the scaling behaviour of equation (25.67) derived in Section 25.5.

It is also possible in this approach to analyse the large momentum behaviour of the r.h.s. of equation (26.19), and, therefore, to calculate corrections to the leading behaviour, using the short distance expansion introduced in Chapter 12 (see Section 26.7).

Remark. Note that if g_{r} is fixed with $g_{\mathrm{r}} < g_{\mathrm{r}}^*$ then the large momentum behaviour of correlation functions is given by the UV fixed point $g_{\mathrm{r}} = 0$, that is, by perturbation theory. However, as we have already discussed, this corresponds to a situation where the bare coupling constant u goes to zero as $m_{\mathrm{r}}^{\varepsilon}$, which is rather unnatural.

26.2 Expansion around the Critical Theory

To be able to describe the whole critical domain, we now use a different strategy. First, we consider the functional integral representation of n-point correlation functions corresponding to the hamiltonian (26.1). If we formally expand it in powers of t, we obtain an expansion in terms of *critical* correlation functions with $\frac{1}{2}\int \mathrm{d}^d x\, \phi^2(x)$ (the most IR singular part of the energy operator) insertions (see also Section 10.10). Consequently, to be able to define these correlation functions by their perturbative expansion, we now have to return to the framework developed in Sections 25.4, 25.6 and to the ε-expansion.

However, even so, because the insertion of $\int \mathrm{d}^d x\, \phi^2(x)$ is the insertion of the Fourier transform of $\phi^2(x)$ at zero momentum, the corresponding correlation functions are still IR divergent. We thus first replace in the hamiltonian (26.1) the constant t by a field $t(x)$. We can then expand the 1PI correlation functions $\Gamma^{(n)}$, as functions of space variables, in powers of the field t:

$$\Gamma^{(n)}(x_i;t,g,\Lambda) = \sum_{l=0}^{\infty} \frac{1}{l!} \int \mathrm{d}y_1 \ldots \mathrm{d}y_l\, t(y_1)\ldots t(y_l)\Gamma^{(l,n)}(y_j;x_i;t=0,g,\Lambda). \qquad (26.22)$$

As already noted in Section 10.10, by acting with the functional differential operator $\int \mathrm{d}y\, t(y)\delta/\delta t(y)$ on equation (26.22) we generate in the r.h.s. a factor l in front of $\Gamma^{(l,n)}$. It is then easy to verify that equation (25.72) implies (see equation (10.80)):

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g) - \eta_2(g)\int t(y)\frac{\delta}{\delta t(y)}\right]\Gamma^{(n)}(x_i;t,g,\Lambda) = 0\,. \qquad (26.23)$$

To calculate $\Gamma^{(n)}$, it is possible to perform a partial summation of perturbation theory in order to introduce the mass renormalized free propagator $(p^2+t)^{-1}$, for example, by using the loop expansion. In this new perturbation theory, the constant t limit leads to a massive theory and no IR divergence is generated anymore. Then equation (26.23), in momentum representation, becomes

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g) - \eta_2(g)t\frac{\partial}{\partial t}\right]\Gamma^{(n)}(p_i;t,g,\Lambda) = 0\,. \qquad (26.24)$$

Equation (26.24) is the formal analogue of equation (10.82) and differs from equation (26.7), which also applies to the non-critical (i.e. massive) theory, by the property that it is homogeneous.

26.3 Scaling Laws above T_c

We first discuss again the critical behaviour above T_c, which we have already examined in Section 26.1, within this new formalism. We integrate again equation (26.24) using the method of characteristics. In addition to the functions $g(\lambda)$ and $Z(\lambda)$ of equations (25.57), we now need a function $t(\lambda)$:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[Z^{-n/2}(\lambda)\Gamma^{(n)}(p_i;t(\lambda),g(\lambda),\lambda\Lambda)\right] = 0\,, \qquad (26.25)$$

which we determine by imposing the consistency with equation (26.24). We obtain the set of equations:

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} g(\lambda) = \beta(g(\lambda)), \qquad g(1) = g\,, \tag{26.26}$$

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} \ln t(\lambda) = -\eta_2(g(\lambda)), \qquad t(1) = t\,, \tag{26.27}$$

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} \ln Z(\lambda) = \eta\big(g(\lambda)\big), \qquad Z(1) = 1\,. \tag{26.28}$$

Dimensional analysis yields

$$\Gamma^{(n)}\big(p_i; t(\lambda), g(\lambda), \lambda\Lambda\big) = (\lambda\Lambda)^{d-n(d-2)/2}\Gamma^{(n)}\big(p_i/\lambda\Lambda; t(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1\big). \tag{26.29}$$

The critical region is defined in particular by $|t| \ll \Lambda^2$, and this is the source of the IR singular behaviour which appears in perturbation theory. If we can find a solution λ to the equation

$$t(\lambda) = \lambda^2\Lambda^2, \tag{26.30}$$

then the theory at scale λ will no longer be critical. Combining equations (26.25–26.30) we find

$$\Gamma^{(n)}(p_i; t, g, \Lambda) = Z^{-n/2}(\lambda) m^{(d-n(d-2)/2)}\Gamma^{(n)}(p_i/m; 1, g(\lambda), 1), \tag{26.31}$$

where we have introduced the notation:

$$m = \lambda\Lambda\,. \tag{26.32}$$

The solution of equation (26.27) can be written as

$$t(\lambda) = t \exp\left[-\int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma}\eta_2\big(g(\sigma)\big)\right]. \tag{26.33}$$

Substituting this relation into equation (26.30), we then obtain

$$\ln(t/\Lambda^2) = \int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma}\frac{1}{\nu(g(\sigma))}. \tag{26.34}$$

We look for a solution λ in the limit $t/\Lambda^2 \ll 1$. Since $\nu(g)$ is a positive function, at least for g small enough as can be verified in the explicit expression (25.77), equation (26.34) implies 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 ν and equation (26.34) can be rewritten as

$$\ln(t/\Lambda^2) \sim \frac{1}{\nu}\ln\lambda\,. \tag{26.35}$$

Equation (26.28) then yields

$$Z(\lambda) \sim \lambda^\eta. \tag{26.36}$$

Taking the large Λ, or the small λ limit, and using equations (26.35,26.36) in equation (26.31), we finally obtain

$$\Gamma^{(n)}(p_i;t,g,\Lambda=1) \underset{\substack{t\ll 1\\ |p_i|\ll 1}}{\sim} m^{(d-n(d-2+\eta)/2)} F_+^{(n)}(p_i/m) \tag{26.37}$$

with

$$m(\Lambda=1)=\xi^{-1}\sim t^{\nu}. \tag{26.38}$$

Both equations (26.37) and (26.17) express the same scaling property. However, one new result has been obtained. From equation (26.37), we infer that the quantity m is proportional to the physical mass or inverse correlation length. Equation (26.38) then shows that the divergence of the correlation length $\xi=m^{-1}$ at T_c is characterized by the exponent ν (a result which we could also have derived with the formalism of Section 26.1).

For $t\neq 0$, the correlation functions are finite at zero momentum and behave as

$$\Gamma^{(n)}(0;t,g,\Lambda)\propto t^{\nu(d-n(d-2+\eta)/2)}. \tag{26.39}$$

In particular, for $n=2$, we obtain the inverse magnetic susceptibility:

$$\chi^{-1}=\Gamma^{(2)}(p=0;t,g,\Lambda)\propto t^{\nu(2-\eta)}. \tag{26.40}$$

The exponent which characterizes the divergence of χ is usually called γ. Equation (26.39) establishes the relation between exponents:

$$\gamma=\nu(2-\eta). \tag{26.41}$$

Finally, we note that, for the critical theory to exist, when t goes to zero different powers of t have to cancel in equation (26.37). From this observation, we recover equation (25.67) in the form

$$\Gamma^{(n)}(\lambda p_i;t,g,\Lambda=1) \underset{t^{\nu}\ll\lambda|p_i|\ll 1}{\propto} \lambda^{d-n(d-2+\eta)/2}. \tag{26.42}$$

26.4 Correlation Functions with ϕ^2 Insertions

A differentiation of the functional integral with respect to $t(x)$ before taking the uniform t limit, generates correlation functions with $\left[\frac{1}{2}\phi^2(x)\right]$ insertions (in the statistical formulation insertions of the hamiltonian or configuration energy density). By differentiating equation (26.23) with respect to $t(y_1)\dots t(y_l)$ before taking the same limit, we derive RG equations for the corresponding 1PI correlation functions. One verifies that the resulting equation, except for $l=2$, $n=0$, is

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)\frac{\partial}{\partial g}-\frac{n}{2}\eta(g)-\eta_2(g)\left(l+t\frac{\partial}{\partial t}\right)\right]\Gamma^{(l,n)}(q_i;p_j;t,g,\Lambda)=0\,. \tag{26.43}$$

This equation can be solved exactly in the same way as equation (26.24) and leads, for $t\ll 1$, $|q_i|\ll 1$, $|p_j|\ll 1$, to the critical behaviour:

$$\Gamma^{(l,n)}(q_i;p_j;t,g,\Lambda=1)\sim m^{[d-l/\nu-n(d-2+\eta)/2]}F_+^{(l,n)}(q_i/m;p_j/m) \tag{26.44}$$

with

$$m \sim t^{\nu}. \tag{26.45}$$

The discussion then exactly follows the lines of Section 26.3, and we do not repeat it here. Rather we concentrate on the case $n = 0$, $l = 2$, which corresponds to the energy density two-point correlation function. Starting from equation (25.72), and using the method explained above, we obtain

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \eta_2(g)\left(2 + t\frac{\partial}{\partial t}\right)\right]\Gamma^{(2,0)}(q;t,g,\Lambda) = \Lambda^{-\varepsilon}B(g). \tag{26.46}$$

The function $\Lambda^{-\varepsilon}C_2(g)$ of equation (25.78) is still a solution of the inhomogeneous equation. Equation (26.44) is then replaced by

$$\Gamma^{(2,0)}(q;t,g,\Lambda = 1) \underset{\substack{t\ll 1\\ |q|\ll 1}}{\sim} m^{(d-2/\nu)}F_{+}^{(2,0)}(q/m) + C_2(g). \tag{26.47}$$

At zero momentum, $\Gamma^{(2,0)}(q = 0)$ is also the second derivative with respect to t of the free energy (the connected vacuum amplitude), that is, the specific heat. Its critical behaviour is thus

$$\Gamma^{(2,0)}(0;t,g,\Lambda = 1) \sim A^{+}t^{-(2-\nu d)} + C_2(g). \tag{26.48}$$

The specific heat exponent is called α. We have, therefore, derived the scaling relation

$$\alpha = 2 - \nu d\,. \tag{26.49}$$

Integrating $\Gamma^{(2,0)}$ twice with respect to t, we obtain $\Omega^{-1}\Gamma(t,g,\Lambda)$, the thermodynamic potential density (calling Ω the volume factor) for a vanishing magnetization (and thus also the free energy in zero magnetic field):

$$\begin{aligned}\Omega^{-1}\Gamma(M = 0,t,g,\Lambda) &= \Lambda^{d}C_0(g) + \Lambda^{d-2}C_1(g)\,t + \Lambda^{d-4}C_2(g)\frac{t^2}{2}\\ &\quad + \Lambda^{d}\frac{A^{+}}{(2-\alpha)(1-\alpha)}\left(\frac{t}{\Lambda^2}\right)^{2-\alpha} + \cdots\end{aligned} \tag{26.50}$$

for $|t| \ll \Lambda^2$.

The three first terms correspond to the beginning of the small t 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 can be cancelled in appropriate ratios).

Since we have completely described the critical behaviour above T_c, we examine in the next section the critical behaviour in the ordered phase ($M \neq 0$).

26.5 Scaling Laws in a Magnetic Field and Below T_c

In order to pass continuously from the disordered ($T > T_c$) to the ordered phase ($T < T_c$), it is necessary to add to the hamiltonian an interaction which explicitly breaks symmetry. 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, that is, the equation of state.

As we have already shown in the mean field analysis, in the ordered phase some qualitative differences appear between systems which have a discrete and a continuous symmetry. They will be illustrated with the example of the N-vector model with $O(N)$ symmetry and $(\phi^2)^2$ interaction in Section 26.6.

26.5.1 The equation of state

We denote by M the magnetization, expectation value of $\phi(x)$, in a constant field H. The thermodynamic potential density, as a function of M, is by definition

$$\Omega^{-1}\Gamma(M,t,g,\Lambda) = \sum_{n=0}^{\infty} \frac{M^n}{n!}\Gamma^{(n)}(p_i = 0; t, g, \Lambda). \tag{26.51}$$

The magnetic field H is given by

$$H = \Omega^{-1}\frac{\partial\Gamma}{\partial M} = \sum_{n=1}^{\infty} \frac{M^n}{n!}\Gamma^{(n+1)}(p_i = 0; t, g, \Lambda). \tag{26.52}$$

Noting that $n \equiv M(\partial/\partial M)$, we immediately derive from the RG equation (26.24), the RG equation satisfied by $H(M,t,g,\Lambda)$:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{1}{2}\eta(g)\left(1 + M\frac{\partial}{\partial M}\right) - \eta_2(g)t\frac{\partial}{\partial t}\right] H(M,t,g,\Lambda) = 0\,. \tag{26.53}$$

To integrate equation (26.53) by the method of characteristics, we have to introduce, in addition to the functions $g(\lambda)$, $t(\lambda)$ and $Z(\lambda)$ of equations (26.26–26.28), a new function $M(\lambda)$ that satisfies

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln M(\lambda) = -\tfrac{1}{2}\eta\left[g(\lambda)\right], \qquad M(1) = M\,. \tag{26.54}$$

Comparing equations (26.54) and (26.28), we see that $M(\lambda)$ is given by

$$M(\lambda) = MZ^{-1/2}(\lambda).$$

The solution of equation (26.53) can then be written as

$$H(M,t,g,\Lambda) = Z^{-1/2}(\lambda)H\left[M(\lambda), t(\lambda), g(\lambda), \lambda\Lambda\right]. \tag{26.55}$$

Dimensional analysis shows that

$$H(M,t,g,\Lambda) = \Lambda^{3-\varepsilon/2}H(M/\Lambda^{1-\varepsilon/2}, t/\Lambda^2, g, 1). \tag{26.56}$$

Applying this relation to the r.h.s. of equation (26.55), we obtain

$$H(M,t,g,\Lambda) = (\lambda\Lambda)^{3-\varepsilon/2} Z^{-1/2}(\lambda) H[M(\lambda)/(\lambda\Lambda)^{1-\varepsilon/2}, t(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1]. \quad (26.57)$$

We again use the arbitrariness of λ to move outside the critical domain, in order to remove the critical singularities in the r.h.s. of equation (26.57). Here, a natural choice is

$$M(\lambda) = (\lambda\Lambda)^{1-\varepsilon/2}, \quad (26.58)$$

which implies, using the solution of equation (26.54), that

$$\ln(M/\Lambda^{1-\varepsilon/2}) = \frac{1}{2}\int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma}\left[d-2+\eta(g(\sigma))\right]. \quad (26.59)$$

In the critical domain, the magnetization is small:

$$M \ll \Lambda^{1-\varepsilon/2}.$$

For g small, the expression $d-2+\eta(g)$ is positive because $d \ge 2$ and $\eta(g) > 0$, which again implies that λ is small and thus $g(\lambda)$ close to g^*. In this limit, equation (26.59) leads to

$$M\Lambda^{\varepsilon/2-1} \sim \lambda^{(d-2+\eta)/2}. \quad (26.60)$$

Now, equation (26.27) implies

$$t(\lambda)\,/\lambda^2 \sim t\lambda^{-1/\nu}, \quad (26.61)$$

and we have already seen that

$$Z(\lambda) \sim \lambda^{\eta}. \quad (26.62)$$

Replacing $t(\lambda)$ and $Z(\lambda)$ by their asymptotic forms (26.61,26.62) and using equation (26.60) to eliminate λ, we finally obtain

$$H(M,t,g,1) \sim M^{\delta} f(tM^{-1/\beta}) \quad (26.63)$$

with

$$\beta = \frac{\nu}{2}(d-2+\eta), \quad (26.64)$$

$$\delta = \frac{d+2-\eta}{d-2+\eta}. \quad (26.65)$$

Equation (26.63) establishes the scaling properties of the equation of state. Moreover, equations (26.64,26.65) relate the traditional critical exponents which 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 (26.63) (see Section 27.1).

Properties of the universal function $f(x)$

(i) Griffith's analyticity: equation (26.52) shows that H has a regular expansion in odd powers of M for $t > 0$. This implies that when the variable x becomes large and positive, $f(x)$ has the expansion

$$f(x) = \sum_{p=0}^{\infty} a_p x^{\gamma - 2p\beta}. \tag{26.66}$$

(ii) When M is different from zero, the theory remains massive when t vanishes. In the loop expansion, the corresponding massive propagator has the form $(p^2 + aM^2)^{-1}$ in momentum space. It follows that we can expand $\Gamma(M,t)$ and, therefore, $H(M,t)$ in powers of t without meeting IR divergences. Consequently, $f(x)$ is infinitely differentiable at $x = 0$.

(iii) The appearance of a spontaneous magnetization, below T_c, implies that the function $f(x)$ has a negative zero x_0:

$$f(x_0) = 0\,, \quad x_0 < 0\,. \tag{26.67}$$

Then, equation (26.63) leads to the relation

$$M = |x_0|^{-\beta} (-t)^{\beta} \quad \text{for} \quad H = 0\,, \ \ t < 0\,. \tag{26.68}$$

Equation (26.68) gives the behaviour of the spontaneous magnetization when the temperature approaches the critical temperature from below.

26.5.2 *Correlation functions in a field*

We now examine the behaviour of correlation functions in a field. All expressions will again be written for Ising-like systems. The generalization to the N-vector model with $O(N)$ symmetry is simple and will be briefly discussed in Section 26.6.

The correlation functions at fixed magnetization M are obtained by expanding the generating functional $\Gamma(M(x))$ of 1PI correlation functions, around $M(x) = M$:

$$\Gamma^{(n)}(x_1, \ldots, x_n; t, g, \Lambda) = \frac{\delta^n}{\delta M(x_1) \ldots \delta M(x_n)} \Gamma(M(x), t, g, \Lambda)\bigg|_{M(x)=M} . \tag{26.69}$$

The expansion in powers of M of the r.h.s. of the equation then yields

$$\Gamma^{(n)}(p_1, \ldots, p_n; t, M, g, \Lambda) = \sum_{s=0}^{\infty} \frac{M^s}{s!} \Gamma^{(n+s)}(p_1, \ldots, p_n, 0, \ldots, 0; t, 0, g, \Lambda). \tag{26.70}$$

From the RG equations satisfied by the correlation functions in zero magnetization (equations (26.24)), it is then simple to derive

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2}\eta(g) \left(n + M \frac{\partial}{\partial M}\right) - \eta_2(g) t \frac{\partial}{\partial t}\right] \Gamma^{(n)}(p_i; t, M, g, \Lambda) = 0\,. \tag{26.71}$$

This equation can be solved by exactly the same method as equation (26.53):

$$\Gamma^{(n)}(p_i; t, M, g, \Lambda = 1) \sim m^{[d-(d-2+\eta)/2]} F^{(n)}(p_i/m, t m^{-1/\nu}), \tag{26.72}$$

for $|p_i| \ll 1$, $|t| \ll 1$, $M \ll 1$ and with the definition

$$m = M^{\nu/\beta}. \tag{26.73}$$

The r.h.s. of equation (26.72) now depends on two different mass scales: $m = M^{\nu/\beta}$ and t^{ν}.

26.5.3 Correlation functions in zero field below T_c

We have argued above that correlation functions are regular functions of t for small t, provided M does not vanish. It is, therefore, possible to cross the critical point 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 t, by equation (26.68). After elimination of M in favour of t in equation (26.72), one finds the critical behaviour below T_c:

$$\Gamma^{(n)}(p_i;t,M(t,H=0),g,1) \sim m^{d-n(d-2+\eta)/2}F_{-}^{(n)}(p_i/m) \tag{26.74}$$

with

$$m = |x_0|^{-\nu}(-t)^{\nu}, \qquad H=0\,, \qquad t<0\,. \tag{26.75}$$

We conclude that correlation functions have exactly the same scaling behaviour above and below T_c. In particular, since traditionally one defines below T_c:

$$m^{-1} = \xi \sim (-t)^{-\nu'}, \qquad \left[\Gamma^{(2)}(0)\right]^{-1} = \chi \sim (-t)^{-\gamma'}, \tag{26.76}$$

we have established $\nu' = \nu$ and $\gamma' = \gamma$.

The universal functions $F_{+}^{(n)}$ and $F_{-}^{(n)}$, instead, are different.

The extension of these considerations to the functions with ϕ^2 insertions, $\Gamma^{(l,n)}$ is straightforward. In particular, the same method yields the behaviour of the specific heat below T_c:

$$\Gamma^{(2,0)}(q=0,M(H=0,t)) - \Lambda^{-\varepsilon}C_2(g) \underset{\text{for } t<0}{\sim} A^{-}(-t)^{-\alpha}, \tag{26.77}$$

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 t, H, M, distances or momenta. Physical quantities which are independent of these normalizations are universal pure numbers. Examples are provided by the ratios of the amplitudes of the singularities above and below T_c like A^{+}/A^{-} for the specific heat.

26.6 The N-Vector Model

We now generalize preceding results to models in which the order parameter is a N-vector and which have symmetries such that the Landau–Ginzburg–Wilson hamiltonian has still the form of a ϕ^4 field theory. We first consider a simple but important example: a model in which the hamiltonian has an $O(N)$ symmetry, and then discuss the general situation.

26.5.3 Correlation functions in zero field below T_c

We have argued above that correlation functions are regular functions of t for small t, provided M does not vanish. It is, therefore, possible to cross the critical point 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 t, by equation (26.68). After elimination of M in favour of t in equation (26.72), one finds the critical behaviour below T_c:

$$\Gamma^{(n)}(p_i;t,M(t,H=0),g,1) \sim m^{d-n(d-2+\eta)/2} F_{-}^{(n)}(p_i/m) \tag{26.74}$$

with

$$m = |x_0|^{-\nu}(-t)^{\nu}, \qquad H=0\,, \qquad t<0\,. \tag{26.75}$$

We conclude that correlation functions have exactly the same scaling behaviour above and below T_c. In particular, since traditionally one defines below T_c:

$$m^{-1} = \xi \sim (-t)^{-\nu'}, \qquad \left[\Gamma^{(2)}(0)\right]^{-1} = \chi \sim (-t)^{-\gamma'}, \tag{26.76}$$

we have established $\nu' = \nu$ and $\gamma' = \gamma$.

The universal functions $F_{+}^{(n)}$ and $F_{-}^{(n)}$, instead, are different.

The extension of these considerations to the functions with ϕ^2 insertions, $\Gamma^{(l,n)}$ is straightforward. In particular, the same method yields the behaviour of the specific heat below T_c:

$$\Gamma^{(2,0)}(q=0,M(H=0,t)) - \Lambda^{-\varepsilon}C_2(g) \underset{\text{for } t<0}{\sim} A^{-}(-t)^{-\alpha}, \tag{26.77}$$

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 t, H, M, distances or momenta. Physical quantities which are independent of these normalizations are universal pure numbers. Examples are provided by the ratios of the amplitudes of the singularities above and below T_c like A^{+}/A^{-} for the specific heat.

26.6 The N-Vector Model

We now generalize preceding results to models in which the order parameter is a N-vector and which have symmetries such that the Landau–Ginzburg–Wilson hamiltonian has still the form of a ϕ^4 field theory. We first consider a simple but important example: a model in which the hamiltonian has an $O(N)$ symmetry, and then discuss the general situation.

26.6.1 The $O(N)$ symmetric N-vector model

The $O(N)$ symmetric effective hamiltonian has the form

$$\mathcal{H}(\phi) = \int \left\{ \tfrac{1}{2} \left[\partial_\mu \phi(x)\right]^2 + \tfrac{1}{2}(r_c + t)\phi^2(x) + \tfrac{1}{4!} g \Lambda^{\varepsilon} \left(\phi^2(x)\right)^2 \right\} \mathrm{d}^d x \,, \tag{26.78}$$

in which ϕ is a N-vector.

The RG equations have exactly the same form as the Ising-like case $N = 1$. The coupling constant RG function has the expansion

$$\beta(g) = -\varepsilon g + \frac{N+8}{48\pi^2} g^2 + O(g^3, g^2\varepsilon). \tag{26.79}$$

At leading order in ε, $\beta(g)$ has a zero g^* which is an IR fixed point:

$$g^* = \frac{48\pi^2}{N+8}\varepsilon + O(\varepsilon^2), \tag{26.80}$$

$$\omega \equiv \beta'(g^*) = \varepsilon + O(\varepsilon^2), \tag{26.81}$$

and, therefore, all the scaling laws derived above and in Chapter 25 can also be proven for the N-vector model with $O(N)$ symmetry. We give the expressions of the other RG functions at leading order in Section 26.6.2 by specializing expressions obtained for the general N-vector model.

Correlation functions in a field or below T_c. The addition of a magnetic field term in a $O(N)$ symmetric hamiltonian has several effects. 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 a 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 WT identities, which have been discussed there in a general framework. We recall here in the case of the $O(N)$ symmetry the simplest one, involving the two-point function at zero momentum, also directly established in Section 24.4.3. In terms of Γ_{T} the inverse two-point function transverse to $\mathbf{M}$ at zero momentum or the transverse susceptibility, it reads

$$\Gamma_{\mathrm{T}}(p=0) = \chi_{\mathrm{T}}^{-1} = \frac{H}{M}. \tag{26.82}$$

We recognize, in different notations, equation (13.44). As we have already discussed in Section 13.4, it follows from this equation that if H goes to zero below T_c, H/M and, therefore, Γ_{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 into $O(N-1)$.

Note finally that the inverse longitudinal two-point function $\Gamma_{\mathrm{L}}(p)$ is singular at zero momentum in zero field below T_c as one can infer from its Feynman graph expansion (figure 26.1). 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.

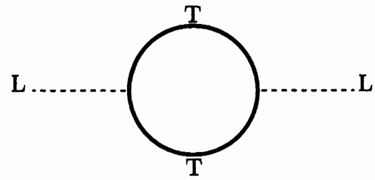


Fig. 26.1 One-loop Goldstone mode contribution to Γ_{L}.

We shall discuss this problem more thoroughly in Chapter 31. In particular, the correlation length (26.75) becomes a crossover scale between critical behaviour and Goldstone mode dominated behaviour at larger distances.

Remark. Previous results also apply to a model in which the hamiltonian has a symmetry smaller than the group $O(N)$ but is still such that the effective hamiltonian has the form (26.78) because the quadratic and quartic group invariants are determined uniquely. 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 a similar phenomenon: the hypercubic symmetry of the lattice has led to a $O(d)$ continuum space symmetry in the critical domain.

26.6.2 The general N-vector model

There exist interesting physical systems in which the 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 9.5. The effect is to renormalize the effective local hamiltonian for the critical components. This remark is related to the decoupling theorem of Particle Physics. We can, therefore, restrict the discussion to theories with only one correlation length. This second category consists in systems in which the hamiltonian is invariant under a symmetry group, subgroup of $O(N)$, which admits a unique quadratic invariant $\Sigma\phi_i\phi_i$ but several quartic invariants.

The general hamiltonian then has the form

$$\mathcal{H}(\phi) = \int \mathrm{d}^d x \left\{ \frac{1}{2} \sum_{i=1}^{N} \left[(\partial_\mu \phi_i)^2 + (r_c + t)\phi_i^2 \right] + \frac{\Lambda^\varepsilon}{4!} \sum_{i,j,k,l=1}^{N} g_{ijkl}\phi_i\phi_j\phi_k\phi_l \right\}. \qquad (26.83)$$

Due to the symmetry, the tensor g_{ijkl} has special properties which imply that the two-point correlation function $\Gamma^{(2)}_{ij}$ is necessarily proportional to the unit matrix:

$$\Gamma^{(2)}_{ij}(p,t,g) = \delta_{ij}\Gamma^{(2)}(p,t,g). \qquad (26.84)$$

This equation implies successive conditions on the tensor g_{ijkl} in perturbation theory.

In Section 11.6, we have discussed the renormalization and RG equations for a general ϕ^4 field theory and we can apply the formalism here.

Renormalization group equations. 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 correlation functions now takes the form

$$\Gamma^{(n)}_{\mathrm{r};i_1,i_2,\ldots,i_n}(p_k, m_\mathrm{r}, g_\mathrm{r}, \mu) = Z^{n/2}\Gamma^{(n)}_{i_1,i_2,\ldots,i_n}(p_k, t, g, \Lambda), \tag{26.85}$$

in which g stands for g_{ijkl} and g_r for $g_{\mathrm{r};ijkl}$.

If we differentiate the equation with respect to Λ at g_r, m_r and μ fixed, we obtain the RG equation:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta_{i'j'k'l'}\frac{\partial}{\partial g_{i'j'k'l'}} - \frac{n}{2}\eta(g) - \eta_2(g)t\frac{\partial}{\partial t}\right]\Gamma^{(n)}_{i_1,i_2,\ldots,i_n} = 0 \tag{26.86}$$

with the definition

$$\beta_{i'j'k'l'}\frac{\partial g_{\mathrm{r};ijkl}}{\partial g_{i'j'k'l'}} = -\varepsilon g_{\mathrm{r};ijkl}\,. \tag{26.87}$$

These equations can be integrated by the same method as before. One introduces a scale-dependent coupling constant $g_{ijkl}(\lambda)$ obeying the flow equation:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}g_{ijkl}(\lambda) = \beta_{ijkl}\bigl(g(\lambda)\bigr). \tag{26.88}$$

The long distance properties of such theories are then governed by fixed points solution of the equation

$$\beta_{ijkl}(g^*) = 0\,. \tag{26.89}$$

The local stability properties of fixed points are governed by the eigenvalues of the matrix $M_{ijkl,i'j'k'l'}$:

$$M_{ijkl,i'j'k'l'} = \frac{\partial\beta_{ijkl}(g^*)}{\partial g_{i'j'k'l'}}. \tag{26.90}$$

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 (26.88) which determine the basin of attraction in coupling space of each IR stable fixed point. We will not discuss this problem further here and refer to the literature where a number of specific models have been considered.

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 $\zeta^{(2)}_{ij,kl}$. This leads to the RG equation:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta_{i'j'k'l'}\frac{\partial}{\partial g_{i'j'k'l'}} - \frac{n}{2}\eta(g) - \eta_2(g)t\frac{\partial}{\partial t}\right]\Gamma^{(l,n)}_{j_1k_1\ldots j_lk_l,i_1\ldots i_n} - \sum_{m=1}^{l}\eta^{(2)}_{j_mk_m,b_mc_m}\Gamma^{(l,n)}_{j_1k_1\ldots b_mc_m\ldots j_lk_l,i_1\ldots i_n} = 0 \tag{26.91}$$

with the definition

$$\eta^{(2)}_{ij,kl} = -\beta_{i'j'k'l'}\left(\frac{\partial\zeta^{(2)}}{\partial g_{i'j'k'l'}}\left[\zeta^{(2)}\right]^{-1}\right)_{ij,kl}. \tag{26.92}$$

Since the insertions of $\Sigma\phi_i\phi_i$, which are generated by a differentiation with respect to t, are multiplicatively renormalized, the matrix $\eta^{(2)}_{ij,kl}(g)$ has δ_{kl} as eigenvector, and

$$\eta^{(2)}_{ij,kk}(g) = \eta_2(g)\delta_{ij}\,. \tag{26.93}$$

RG functions. At one-loop order, the RG functions are given by

$$\begin{aligned}
\beta_{ijkl} &= -\varepsilon g_{ijkl} + \frac{1}{16\pi^2}(g_{ijmn}g_{mnkl} + 2\text{ terms}) + O(g^3), && (26.94)\\
\eta(g)\delta_{ij} &= \frac{1}{24}\frac{1}{(8\pi^2)^2}g_{iklm}g_{jklm} + O(g^3), && (26.95)\\
\eta^{(2)}_{ij,kl} &= -\frac{1}{16\pi^2}g_{ijkl} + O(g^2). && (26.96)
\end{aligned}$$

The condition (26.93) then implies

$$g_{ijkk} = G\delta_{ij}\,, \tag{26.97}$$

and summing over ij, we can rewrite equation (26.95) as

$$\eta(g) = \frac{1}{24N}\frac{1}{(8\pi^2)^2}g_{iklm}g_{iklm} + O(g^3). \tag{26.98}$$

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 (26.94–26.96) to the case of the $(\phi^2)^2$ field theory with $O(N)$ symmetry. We then have to substitute

$$g_{ijkl} = \frac{g}{3}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{26.99}$$

After a short calculation, the expression (26.79) of the β-function is recovered and, in addition,

$$\eta(\tilde{g}) = \frac{(N+2)}{72}\tilde{g}^2 + O(\tilde{g}^3), \tag{26.100}$$

where the notation (11.75) has been used:

$$\tilde{g} = N_d g\,, \qquad N_d = 2(4\pi)^{-d/2}/\Gamma(d/2).$$

As noted in Section 11.6, introducing the identity matrix $\mathbf{I}$ and the projector $\mathbf{P}$:

$$\begin{aligned}
I_{ij,kl} &= \tfrac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), && (26.101)\\
P_{ij,kl} &= \frac{1}{N}\delta_{ij}\delta_{kl}\,, && (26.102)
\end{aligned}$$

we can write the matrix $\eta^{(2)}_{ij,kl}$:

$$\eta^{(2)} = -(N\mathbf{P} + 2\mathbf{I})\frac{\tilde{g}}{6} + O(\tilde{g}^2). \tag{26.103}$$

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

$$\eta_2'(g) = \frac{\varphi(g)}{\nu(g)} - 2 . \tag{26.104}$$

The fixed point value $\varphi = \varphi(g^*)$ is called the crossover exponent. Finally, the stability conditions are given by the eigenvalues of the matrix M. Setting

$$g_{ijkl} = g^*_{ijkl} + s_{ijkl} , \tag{26.105}$$

we find at leading order

$$(Ms)_{ijkl} = -\varepsilon s_{ijkl} + \frac{\varepsilon}{N+8}(\delta_{ij} s_{mmkl} + 5 \text{ terms} + 12 s_{ijkl}). \tag{26.106}$$

Taking s_{ijkl} proportional to g^*_{ijkl} we recover the exponent ω. More generally, the eigenvectors can be classified according to their trace properties. Let us write

$$s_{ijkl} = u g^*_{ijkl} + (v_{ij}\delta_{kl} + 5 \text{ terms}) + w_{ijkl} , \tag{26.107}$$

in which the tensors v_{ij} and w_{ijkl} are traceless:

$$v_{ii} = 0 , \qquad w_{ijkk} = 0 . \tag{26.108}$$

The three eigenvalues corresponding to u, w, v are, respectively,

$$\omega = \varepsilon + O(\varepsilon^2), \ \omega_{\text{anis.}} = \varepsilon \frac{4-N}{N+8} + O(\varepsilon^2), \ \ \omega' = \frac{8\varepsilon}{N+8} + O(\varepsilon^2). \tag{26.109}$$

Note that the perturbation proportional to v_{ij} does not satisfy the trace condition (26.97). It, therefore, 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 (26.97), the relevant eigenvalue is ω_{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. This is again an example of dynamically induced symmetry: correlation functions in the critical domain have a larger symmetry than microscopic correlation functions. The calculation of $\omega_{\text{anis.}}$ at order ε yields

$$N_c = 4 - 2\varepsilon + O(\varepsilon^2). \tag{26.110}$$

Gradient flow. It can be verified in expression (26.94) that the β-function defines a gradient flow. At two-loop order if we introduce the potential U:

$$\begin{aligned} U(g) = &-\frac{1}{2}\varepsilon g_{ijkl} g_{ijkl} + \frac{1}{(4\pi)^2} g_{ijkl} g_{klmn} g_{mnij} \\ &+ \frac{1}{(4\pi)^4}\left(\frac{3}{2} g_{ijkl} g_{ijmn} g_{pqkm} g_{pqln} + \frac{1}{12} g_{ijkl} g_{ijkm} g_{npql} g_{npqm}\right), \end{aligned} \tag{26.111}$$

then

$$\beta_{ijkl} = \frac{\partial U(g)}{\partial g_{ijkl}}.$$

Such a structure 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.

Actually, it has been checked up to three-loop order that this property remains true in the following sense: calling g_α the set of all coupling constants, one finds

$$\beta_\alpha = T_{\alpha\beta}\frac{\partial U}{\partial g_\beta},$$

in which $T_{\alpha\beta}$ is a symmetric and, for g_α small, positive matrix. This is the best one can expect in general, considering the transformation properties of the β-function under reparametrization of the coupling constant space. This form is sufficient to ensure all the properties mentioned above.

At leading order in ε a fixed point g^* is solution of the equation

$$\varepsilon g^*_{ijkl} = \frac{1}{16\pi^2}(g^*_{ijmn}g^*_{mnkl} + \text{ 2 terms}) \Rightarrow \varepsilon g^*_{ijkl}g^*_{ijkl} = \frac{3}{(4\pi)^2}g^*_{ijkl}g^*_{klmn}g^*_{mnij}\,,$$

and, therefore, the value of the potential can then be written as

$$U(g^*) = -\tfrac{1}{6}\varepsilon g^*_{ijkl}g^*_{ijkl} + O(\varepsilon^4).$$

Comparing with the expression (26.98) for the RG function $\eta(g)$, we find a relation between the potential at the fixed point and the corresponding exponent

$$\eta = \frac{1}{6N}\frac{1}{(4\pi)^4}g^*_{ijkl}g^*_{ijkl} + O(\varepsilon^3) = -\frac{1}{N\varepsilon}U(g^*) + O(\varepsilon^3).$$

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. It is not clear whether this property remains true beyond the ε-expansion.

26.7 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 ($T = T_c$)

$$\Gamma^{(n)}(p) \underset{\xi^{-1}\ll p\ll\Lambda}{\propto} p^{2-\eta}. \tag{26.112}$$

The r.h.s. 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 $\partial\Gamma^{(n)}(p_1,\ldots,p_n)/\partial t$ is asymptotically negligible with respect to $\Gamma^{(n)}(p_1,\ldots,p_n)$. However, since

$$\frac{\partial}{\partial t}\Gamma^{(n)}(p_1,\ldots,p_n) = \Gamma^{(1,n)}(0;p_1,\ldots,p_n),$$

then

$$\beta_{ijkl} = \frac{\partial U(g)}{\partial g_{ijkl}}.$$

Such a structure 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.

Actually, it has been checked up to three-loop order that this property remains true in the following sense: calling g_α the set of all coupling constants, one finds

$$\beta_\alpha = T_{\alpha\beta} \frac{\partial U}{\partial g_\beta},$$

in which $T_{\alpha\beta}$ is a symmetric and, for g_α small, positive matrix. This is the best one can expect in general, considering the transformation properties of the β-function under reparametrization of the coupling constant space. This form is sufficient to ensure all the properties mentioned above.

At leading order in ε a fixed point g^* is solution of the equation

$$\varepsilon g^*_{ijkl} = \frac{1}{16\pi^2}(g^*_{ijmn} g^*_{mnkl} + \text{ 2 terms}) \Rightarrow \varepsilon g^*_{ijkl} g^*_{ijkl} = \frac{3}{(4\pi)^2} g^*_{ijkl} g^*_{klmn} g^*_{mnij},$$

and, therefore, the value of the potential can then be written as

$$U(g^*) = -\tfrac{1}{6}\varepsilon g^*_{ijkl} g^*_{ijkl} + O(\varepsilon^4).$$

Comparing with the expression (26.98) for the RG function $\eta(g)$, we find a relation between the potential at the fixed point and the corresponding exponent

$$\eta = \frac{1}{6N} \frac{1}{(4\pi)^4} g^*_{ijkl} g^*_{ijkl} + O(\varepsilon^3) = -\frac{1}{N\varepsilon} U(g^*) + O(\varepsilon^3).$$

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. It is not clear whether this property remains true beyond the ε-expansion.

26.7 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 ($T = T_c$)

$$\Gamma^{(n)}(p) \underset{\xi^{-1} \ll p \ll \Lambda}{\propto} p^{2-\eta}. \tag{26.112}$$

The r.h.s. 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 $\partial\Gamma^{(n)}(p_1, \ldots, p_n)/\partial t$ is asymptotically negligible with respect to $\Gamma^{(n)}(p_1, \ldots, p_n)$. However, since

$$\frac{\partial}{\partial t}\Gamma^{(n)}(p_1, \ldots, p_n) = \Gamma^{(1,n)}(0; p_1, \ldots, p_n),$$

the derivative $\partial\Gamma^{(n)}/\partial t$ cannot be evaluated with the same method because the momenta are exceptional. As we have explained in Chapter 12, it is necessary to use the short distance expansion (SDE) of operator products. Let us now concentrate on the two-point function. We have to evaluate $\Gamma^{(1,2)}(0;p,-p)$. However, we cannot apply directly the SDE to $\Gamma^{(1,2)}$ because this would involve $\Gamma^{(2,0)}$ which needs additional renormalizations.

We, therefore, differentiate once more with respect to t:

$$\frac{\partial^2}{(\partial t)^2}\Gamma^{(2)}(p)=\Gamma^{(2,2)}(0,0;p,-p). \tag{26.113}$$

To $\Gamma^{(2,2)}$ we can now apply the SDE:

$$\Gamma^{(2,2)}(0,0;p,-p)\sim B(p)\Gamma^{(3,0)}(0,0,0) \quad \text{for } \xi^{-1}\ll p\ll\Lambda\,. \tag{26.114}$$

As shown in Section 12.4, $B(p)$ satisfies a RG equation which can be obtained by applying the differential operator

$$D\equiv\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)-\left(\frac{1}{\nu(g)}-2\right)t\frac{\partial}{\partial t}, \tag{26.115}$$

on both sides of equation (26.114) and using the RG equations (26.43).

One finds

$$\left[D+\nu^{-1}(g)-2-\eta(g)\right]B(p)\sim 0\,. \tag{26.116}$$

Taking the $t=0$ limit and setting $g=g^*$, the equation becomes

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+\frac{1}{\nu}-2-\eta\right)B(p)\sim 0\,. \tag{26.117}$$

Equation (26.114) shows also that $B(p)$ has canonical dimension ε. It follows that

$$B(p)\sim\Lambda^{\varepsilon}(p/\Lambda)^{2-\eta-(1-\alpha)/\nu}. \tag{26.118}$$

Differentiating equation (26.48) with respect to t, we find

$$\Gamma^{(3,0)}(0,0,0)\sim\Lambda^{-\varepsilon}(t/\Lambda^2)^{-1-\alpha}. \tag{26.119}$$

Integrating equation (26.113) twice with respect to t and using the set of equations (26.114,26.118,26.119), we obtain finally:

$$\Gamma^{(2)}(p)\underset{\xi^{-1}\ll p\ll\Lambda}{=}p^{2-\eta}\left(a+bt\,p^{-1/\nu}+ct^{1-\alpha}p^{-(1-\alpha)/\nu}+\cdots\right). \tag{26.120}$$

A systematic use of the SDE beyond leading order allows to calculate systematic corrections to expression (26.120).

Note that the effect of the differentiation with respect to t has been simply to generate the regular terms in the temperature which have an order in t comparable to the singular terms for $\varepsilon\to 0$.

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 expansion to each term. The results now are different for Ising-like systems and the N-vector model. For Ising-like systems, one finds

$$\Gamma^{(2)}(p,t,M)=p^{2-\eta}\left[a+bt/p^{1/\nu}+F_1(t/M^{1/\beta})(p/M^{\nu/\beta})^{-(1-\alpha)/\nu}+\cdots\right], \qquad (26.121)$$

in which the function F_1 can be related to the free energy and thus the equation of state by

$$F_1(x)=\int_1^{\infty}\mathrm{d}s\, s^{\delta-1/\beta}\left[f'(0)-f'(x/s^{1/\beta})\right]+\frac{f'(0)}{\delta-(1/\beta)+1}. \qquad (26.122)$$

In the $O(N)$ symmetric case, the SDE involves a second operator of dimension 2:

$$\mathcal{O}_{ij}\left[\phi(x)\right]=\phi_i(x)\phi_j(x)-\frac{\delta_{ij}}{N}\phi^2(x). \qquad (26.123)$$

This operator is multiplicatively renormalizable. Correlation functions with the insertion of $\mathcal{O}_{ij}$ satisfy the RG equations:

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)\frac{\partial}{\partial g}-\frac{n}{2}\eta(g)-\left(\frac{\varphi(g)}{\nu(g)}-2\right)\right]\Gamma^{(n)}_{\mathcal{O}_{ij}}(p_i;g,\Lambda)=0\,, \qquad (26.124)$$

in which the RG function $\varphi(g)$ has been defined in equation (26.104). Consequently, a new term is present at the same order in the asymptotic expansion of the two-point function which becomes

$$\begin{aligned}\Gamma^{(2)}(p,t,M)=p^{2-\eta}&\left\{\left[a+bt/p^{1/\nu}+F_1(t/M^{1/\beta})(p/M^{\nu/\beta})^{-(1-\alpha)/\nu}\right]\delta_{ij}\right.\\&\left.+F_2(t/M^{1/\beta})(p/M^{\nu/\beta})^{-d+\varphi/\nu}\left(\frac{\delta_{ij}}{N}-\frac{M_iM_j}{M^2}\right)\right\}+\cdots\,, \end{aligned}\qquad (26.125)$$

in which φ is the crossover exponent and F_2 a new universal function which may be calculated in an ε-expansion. At order ε, it is given by

$$F_2(x)=1+\frac{\varepsilon}{2(N+8)}\left[(x+3)\ln(x+3)-(x+1)\ln(x+1)\right]+O(\varepsilon^2). \qquad (26.126)$$

Bibliographical Notes

Most references given at the end of Chapter 25 are also relevant here and will not be repeated. The chapter is directly inspired by Sections VI, VII, XI and XII of the review article

E. Brézin, J.C. Le Guillou and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, vol. 6, already quoted in Chapter 24,

which also contains a more detailed discussion of the multi-component ϕ^4 theory. The scaling form of the equation of state has been proposed in

B. Widom, *J. Chem. Phys.* 43 (1965) 3892, 3898; C. Domb and D.L. Hunter, *Proc. Phys. Soc. (London)* 86 (1965) 1147,

and verified in RG calculations by

E. Brézin, D.J. Wallace and K.G. Wilson, *Phys. Rev. Lett.* 29 (1972) 591, *Phys. Rev.* B7 (1973) 232; G.M. Avdeeva and A.A. Migdal, *JETP Lett.* 16 (1972) 178; see also

the contribution of D.J. Wallace in *Phase Transitions and Critical Phenomena* vol. 6 already quoted.
Also relevant is
R.B. Griffiths, *Phys. Rev.* 158 (1967) 176.
The N-vector model with cubic anisotropy has been studied in
D.J. Wallace, *J. Phys.C: Solid State Phys.* 6 (1973) 1390; A. Aharony, *Phys. Rev.* B8 (1973) 3342, 3349, 3358, 3363, 4270, *Phys. Rev. Lett.* 31 (1973) 1494.
See also
I.J. Ketley and D.J. Wallace, *J. Phys. A: Math. Nucl. Gen.* 6 (1973) 1667,
and the contribution of A. Aharony to *Phase Transitions and Critical Phenomena* vol. 6 already quoted.
The general N-vector model is discussed in
E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* B10 (1974) 892.
The decoupling of non-critical or massive modes is discussed in
K. Symanzik, *Acta Phys. Austr. Suppl.* XI (1973) 199; T. Applequist and J. Carazzone, *Phys. Rev.* D11 (1975) 2856.
For the application of the SDE to critical phenomena see
E. Brézin, D. Amit and J. Zinn-Justin, *Phys. Rev. Lett.* 32 (1974) 151; E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. Lett.* 32 (1974) 473.
Some results had been anticipated on physical grounds by
M.E. Fisher and J.S. Langer, *Phys. Rev. Lett.* 20 (1968) 665; M.E. Fisher and A. Aharony, *Phys. Rev.* B10 (1974) 2818.
The crossover exponent has been discussed in
E. Riedel and F. Wegner, *Z. Phys.* 225 (1969) 195; M.E. Fisher and P. Pfeuty, *Phys. Rev.* B6 (1972) 1889; K.G. Wilson, *Phys. Rev. Lett.* 28 (1972) 548; Y. Yamazaki and M. Suzuki, *Prog. Theor. Phys.* 50 (1973) 1097; Y. Yamazaki, *Phys. Lett.* 49A (1974) 215.
The effect of long range forces have been considered in
M.E. Fisher, S.K. Ma and B.G. Nickel, *Phys. Rev. Lett.* 29 (1972) 917.

APPENDIX A26
THE SPECIFIC HEAT FOR $\alpha = 0$

In the case of the 2D Ising model, the exponent α vanishes. In such a situation, it is impossible to find a function $C_2(g)\Lambda^{-\varepsilon}$ solution of equation (25.78) regular at $g = g^*$. Indeed, the equation for $C_2(g)$ then reads

$$\beta(g)C_2'(g) + (d - 2\,/\nu(g)\,)C_2(g) = B(g) \qquad (A26.1)$$

with

$$\alpha = 0 \ \Rightarrow \quad \nu(g^*) = 2/d\,. \qquad (A26.2)$$

Thus, at $g = g^*$ the l.h.s. of equation ($A26.1$) vanishes if $C_2(g)$ is regular, and the r.h.s. does not. The general solution of equation ($A26.1$) has instead the form

$$C_2(g) = D(g)\ln|g - g^*| + E(g), \qquad (A26.3)$$

in which $D(g)$ and $E(g)$ are regular at g^* and

$$D(g^*) = B(g^*)/\omega\,. \qquad (A26.4)$$

As a consequence, the function $\Gamma^{(2,0)}(q;t,g,\Lambda) - \Lambda^{-\varepsilon}C_2(g)$, which satisfies a homogeneous RG equation, is singular at g^*. When $g(\lambda)$ approaches g^*, the theory at scale λ is dominated by this singularity. Since

$$|g(\lambda) - g^*| \sim \lambda^{\omega} \sim t^{\omega\nu}, \qquad (A26.5)$$

equations (26.47,26.48) are replaced by

$$\Gamma^{(2,0)}(q;t,g,\Lambda) - \Lambda^{-\varepsilon}C_2(g) \sim \Lambda^{-\varepsilon}\left[F(q/m) - A_2(g)\ln(t/\Lambda^2)\right], \qquad (A26.6)$$
$$\Gamma^{(2,0)}(0;t,g,\Lambda) - \Lambda^{-\varepsilon}C_2(g) \sim \Lambda^{-\varepsilon}\left[F(0) - A_2(g)\ln(t/\Lambda^2)\right] \qquad (A26.7)$$

with

$$A_2(g) = B(g^*)\nu\exp\left[2\int_0^1 \mathrm{d}\sigma\frac{\eta_2(g(\sigma)) - \eta_2}{\sigma}\right].$$

Alternative derivation. Alternatively, we can solve the RG equation (26.46) by the method of characteristics directly. We set

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\zeta^{-2}(\lambda)\left[\Gamma^{(2,0)}(q;t(\lambda),g(\lambda),\lambda\Lambda) - (\lambda\Lambda)^{-\varepsilon}C(\lambda)\right] = 0\,. \qquad (A26.8)$$

Applying the chain rule to evaluate explicitly the derivative and expressing the compatibility with equation (26.46), we find in addition to equations (26.27–26.28), two equations:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln\zeta(\lambda) = \eta_2(g(\lambda)), \qquad \zeta(1) = 1\,, \qquad (A26.9)$$
$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}C(\lambda) = \left[2\eta_2(g(\lambda)) + \varepsilon\right]C(\lambda) + B(g(\lambda)), \quad C(1) = 0\,. \qquad (A26.10)$$

The solution of the first equation can be written as

$$\zeta(\lambda) = \exp\left[\int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma}\eta_2\big(g(\sigma)\big)\right]. \tag{A26.11}$$

It follows, in particular,

$$t(\lambda) = t/\zeta(\lambda). \tag{A26.12}$$

The solution of equation ($A26.10$) can then be written as

$$C(\lambda) = \int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma} B\big(g(\sigma)\big)(\lambda/\sigma)^\varepsilon \left(\frac{\zeta(\lambda)}{\zeta(\sigma)}\right)^2. \tag{A26.13}$$

Using dimensional analysis and imposing the condition (26.30), $t(\lambda) = \lambda^2\Lambda^2$, we obtain first

$$\Gamma^{(2,0)}(q;t,g,\Lambda) = \zeta^{-2}(\lambda)(\lambda\Lambda)^{-\varepsilon}\left[\Gamma^{(2,0)}(q/m;1,g(\lambda),1) - C(\lambda)\right] \tag{A26.14}$$

with (equation (26.32)) $m = \lambda\Lambda$.

Equation (26.35) gives the asymptotic relation between λ and t. For λ small $\zeta(\lambda)$ has a power behaviour which in terms of the exponent ν, related to η_2 by equation (25.75), has the form

$$\zeta(\lambda) \sim \lambda^{1/\nu-2} \sim t^{1-2\nu}. \tag{A26.15}$$

In the absence of the inhomogeneous term we would recover the scaling form (26.44) applied to $l = 2$, $n = 0$. Then, let us examine the behaviour of the inhomogeneous term for λ small:

$$\zeta^{-2}(\lambda)\lambda^{-\varepsilon}C(\lambda) = \int_1^\lambda \mathrm{d}\sigma\,\sigma^{\alpha/\nu-1}\left[B\big(g(\sigma)\big)\sigma^{2\eta_2(g^*)}\zeta^{-2}(\sigma)\right]. \tag{A26.16}$$

For $\alpha > 0$, the integral in the r.h.s. has a limit when λ goes to zero and becomes a function of g in agreement with equation (26.48). When α vanishes, it becomes necessary to extract one singular term from the integrand before taking the limit. One obtains

$$\begin{aligned}\zeta^{-2}(\lambda)\lambda^{-\varepsilon}C(\lambda) &= B(g^*)\exp\left[2\int_0^1 \mathrm{d}\sigma\frac{\eta_2\big(g(\sigma)\big)-\eta_2}{\sigma}\right]\ln\lambda + \text{ regular terms}\\ &= A_2(g)\ln(t/\Lambda^2) + \text{ regular terms}.\end{aligned} \tag{A26.17}$$

Replacing $C(\lambda)$ by this expression in equation ($A26.14$), we recover equation ($A26.6$).

27 CORRECTIONS TO SCALING BEHAVIOUR

In previous chapters, while deriving the scaling behaviour of correlation functions, we have always kept only the leading term in the critical region. For instance, when we have solved the RG equations, we have always 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 had already 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 obtained from the solutions of the RG equations derived previously and will be 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 will be studied in the second part of this chapter. Finally, one physical application is provided by systems with strong dipolar forces which have three as upper-critical dimension.

27.1 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 in a slightly different manner by introducing a set of coupling constant dependent renormalizations:

$$\begin{aligned} \ln \tilde{Z}(g) &= -\int_{g^*}^{g} \frac{\mathrm{d}g'}{\beta(g')} \left[\eta(g') - \eta\right], \\ \tilde{M}(g) &= M Z^{-1/2}(g), \\ \tilde{t}(g) &= t \, \exp\left[\int_{g^*}^{g} \frac{\mathrm{d}g'}{\beta(g')} \left(\frac{1}{\nu(g')} - \frac{1}{\nu}\right)\right], \end{aligned} \tag{27.1}$$

and a new coupling constant $\tilde{g}$ which characterizes the deviation of g from g^*:

$$\tilde{g} = (g - g^*) \exp\left[\int_{g^*}^{g} \mathrm{d}g' \left(\frac{\omega}{\beta(g')} - \frac{1}{(g' - g^*)}\right)\right]. \tag{27.2}$$

Then, we substitute

$$\begin{aligned} \Gamma^{(n)}(p_i; t, M, g, \Lambda) = {}& \tilde{Z}^{-n/2}(g) \Gamma^{(n)}(p_i; \tilde{t}(g), \tilde{M}(g), g^*, \Lambda) \\ & \times C^{(n)}(p_i; \tilde{t}(g), yM(g), \tilde{g}, \Lambda). \end{aligned} \tag{27.3}$$

The functions $C^{(n)}$ satisfy the following boundary conditions:

$$C^{(n)}(p_i; t, M; 0, \Lambda) = 1\,. \tag{27.4}$$

The finite renormalizations (27.1) eliminate trivial deviations from the fixed point theory which correspond simply to a change of normalization of the different scaling variables.

Equation (26.71) then implies

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\omega\tilde{g}\frac{\partial}{\partial\tilde{g}}-\frac{\eta}{2}M\frac{\partial}{\partial M}-\left(\frac{1}{\nu}-2\right)t\frac{\partial}{\partial t}\right]C^{(n)}(p_i,t,M,\tilde{g},\Lambda)=0\,. \tag{27.5}$$

Solving this equation by expanding $C^{(n)}$ in powers of $\tilde{g}$, we obtain

$$C^{(n)}(p_i,t,M,\tilde{g},1)=1+\sum_{s=1}^{\infty}\tilde{g}^s t^{s\omega\nu}D_s^{(n)}(p_i t^{-\nu},tM^{-1/\beta}), \tag{27.6}$$

in which Λ has been set equal to 1.

The exponent ω, defined as the value of the derivative of $\beta(g)$ at g^*, which characterizes the approach to the fixed point, characterizes also, therefore, the leading corrections to the critical behaviour.

For models with one coupling constant, $\beta(g)$ has the form

$$\beta(g)=-\varepsilon g+ag^2+O\left(g^3,g^2\varepsilon\right). \tag{27.7}$$

In such a case, one always finds

$$\omega=\varepsilon+O\left(\varepsilon^2\right),\qquad \omega\nu=\varepsilon/2+O\left(\varepsilon^2\right). \tag{27.8}$$

Note that to render all terms in the expansion (27.6) 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 since the $\phi^4(x)$ operator is marginal, and the approach to the fixed point is only logarithmic. This will be the subject of the next section.

Scaling for $d>4$. Up to now 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 easy 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 $-\omega$:

$$\omega=d-4\quad\text{for } d>4\,. \tag{27.9}$$

Let us consider, for example, the mean field equation of state, valid for all dimensions $d>4$:

$$H=tM+\tfrac{1}{6}gM^3.$$

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 (26.64,26.65). These values are consistent with the product tM since t 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.

27.2 Logarithmic Corrections at the Upper-Critical Dimension

The upper-critical dimension is the dimension at which deviations from mean field theory appear. For our model this dimension is 4. In this dimension, there generally exists a marginal operator, here $\int \phi^4(x)\mathrm{d}x$, and, therefore, as we have indicated in the general discussion of Section 26.1, logarithmic corrections to mean field behaviour are expected. Although the dimension 4 is not of physical relevance for statistical problems, its study is of special pedagogical value, because exact predictions can be derived from RG arguments. Moreover, a ϕ^4 interaction is present in the Higgs sector of the Standard Model of weak electromagnetic interactions. Because the 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 or ferroelectrics with dipolar uniaxial long range forces. The latter example will be discussed in Section 27.5.

We study below only the equation of state and the specific heat, the generalization to correlation functions being straightforward. We write the relation (26.57) for $\varepsilon = 0$:

$$H(M,t,g,\Lambda) = Z^{-1/2}(\lambda)(\lambda\Lambda)^3 H\bigl(M(\lambda)/\lambda\Lambda, t(\lambda)/\lambda^2\Lambda^2, g(\lambda), 1\bigr). \tag{27.10}$$

We recall the definitions of the various functions:

$$\ln\lambda = \int_g^{g(\lambda)} \frac{\mathrm{d}g'}{\beta(g')}, \tag{27.11}$$

$$\ln Z(\lambda) = \int_g^{g(\lambda)} \mathrm{d}g' \frac{\eta(g')}{\beta(g')}, \tag{27.12}$$

$$M(\lambda)\,/\lambda\Lambda = \left[\lambda Z^{1/2}(\lambda)\right]^{-1} M\,/\Lambda\,, \tag{27.13}$$

$$\ln\frac{t(\lambda)}{\lambda^2\Lambda^2} = \ln\frac{t}{\Lambda^2} - \int_g^{g(\lambda)} \frac{\mathrm{d}g'}{\beta(g')\nu(g')}. \tag{27.14}$$

It is also necessary to here recall the expansions of the RG functions for g small (equations (26.79,26.100,26.103)). In the $(\phi^2)^2$ field theory, ϕ being a N-component vector field, that is, in the $O(N)$ symmetric N-vector model:

$$\begin{aligned}\beta(g) &= \frac{(N+8)}{6}\frac{g^2}{8\pi^2} + O(g^3),\\ \eta(g) &= \frac{(N+2)}{72}\left(\frac{g}{8\pi^2}\right)^2 + O(g^3),\\ \nu^{-1}(g) &= 2 - \frac{(N+2)}{6}\frac{g}{8\pi^2} + O(g^2).\end{aligned} \tag{27.15}$$

As we have done previously, we choose λ such that

$$M(\lambda) = \lambda\Lambda\,, \tag{27.16}$$

or, in zero magnetization, solution to the equation

$$t(\lambda) = \lambda^2\Lambda^2. \tag{27.17}$$

In both cases, the solution λ of these equations is small in the critical domain. The corresponding effective coupling $g(\lambda)$ is then also small. It is easy to integrate equation (27.11) in this regime. One finds

$$g(\lambda) = 8\pi^2 \frac{6}{N+8} \frac{1}{|\ln \lambda|} + O\left[\frac{\ln|\ln \lambda|}{(\ln|\lambda|)^2}\right]. \tag{27.18}$$

Due to conditions (27.16) or (27.17), the theory at scale λ is no longer critical; furthermore, $g(\lambda)$ is small: therefore, $H(\lambda)$ in the r.h.s. of equation (27.10) 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 $\ln(t/\Lambda^2)$ or equivalent. Direct perturbation theory is obviously meaningless. In contrast in the effective theory at scale λ the coupling constant $g(\lambda)$ is small and the perturbative coefficients of order 1 due to conditions (27.16) or (27.17).

For example, equation (27.16) can be used:

$$M(\lambda) = \lambda\Lambda \ \Rightarrow \lambda \sim M/\Lambda\,,$$

and calculate $t(\lambda)$:

$$\frac{t(\lambda)}{\lambda^2\Lambda^2} \sim K(g)\frac{t}{\lambda^2\Lambda^2}(|\ln\lambda|)^{-(N+2)/(N+8)} \sim K(g)\frac{t}{M^2}\left|\ln\frac{M}{\Lambda}\right|^{-(N+2)/(N+8)}, \tag{27.19}$$

$K(g)$ being a finite temperature renormalization. Using now the perturbative expansion

$$H = tM + \frac{g}{6}M^3 + \cdots\,,$$

and the relations (27.10) and (27.18), one can derive

$$H(M,t,g,\Lambda=1) = \frac{K(g)\,tM}{|\ln M|^{(N+2)/(N+8)}} + \frac{8\pi^2}{N+8}\frac{M^3}{|\ln M|} + O\left(M^3\frac{\ln|\ln M|}{|\ln M|^2}\right). \tag{27.20}$$

At T_c one thus finds

$$H \propto M^3\,/|\ln M|\,. \tag{27.21}$$

The spontaneous magnetization is given by

$$M \propto (-t)^{1/2}\,|\ln(-t)|^{3/(N+8)}\,. \tag{27.22}$$

Since the expression (27.20) is not uniform for M small, it is necessary to use equation (27.17) to calculate the susceptibility in zero field. One then finds

$$\chi^{-1} \propto |t|\,|\ln|t||^{-[(N+2)/(N+8)]}\,. \tag{27.23}$$

Finally, the specific heat satisfies the RG equation:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \eta_2(g)\left(2+t\frac{\partial}{\partial t}\right)\right]\Gamma^{(2,0)} = B(g), \tag{27.24}$$

and $B(g)$ has the expansion

$$B(g)=\frac{N}{16\pi^2}+O(g). \tag{27.25}$$

It is easy to verify that a function $C_2(g)$ solution of equation (27.24) and thus of

$$\beta(g)C_2'(g)-2\eta_2(g)C_2(g)=B(g),$$

is necessarily singular at the origin (see also Appendix A26). For example, one can take a solution of the form

$$C_2(g)=\frac{3N}{(N-4)}\frac{1}{g}+O(1)\quad \text{for}\ \ N\neq 4\,. \tag{27.26}$$

For $N=4$ an additional logarithmic singularity is present. The combination $\Gamma^{(2,0)}-C_2(g)$ solution of the homogeneous RG equation is for g small dominated by the pole of $C_2(g)$:

$$\Gamma^{(2,0)}(0;t,g,\Lambda)-C_2(g)\sim-\frac{3N}{(N-4)}\frac{1}{g(\lambda)}\zeta^{-2}(\lambda) \tag{27.27}$$

with

$$\zeta(\lambda)=\exp\left[\int_g^{g(\lambda)}\frac{\mathrm{d}g'}{\beta(g')}\eta_2(g')\right]\propto[g(\lambda)]^{-(N+2)/(N+8)}\,. \tag{27.28}$$

Collecting all factors one obtains the behaviour of the specific heat:

$$\Gamma^{(2,0)}(0;t,g,\Lambda)-C_2(g)\propto|\ln t|^{(4-N)/(N+8)}\left[1+O\left(\frac{\ln|\ln t|}{\ln t}\right)\right]. \tag{27.29}$$

Remark. It is apparent from these expressions that a parametrization in terms of the variables t or M leads to rather complicated expressions. A better 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 (27.11). Let us parametrize $\beta(g)$ as

$$\beta(g)=\beta_2g^2+\beta_3g^3+O(g^4), \tag{27.30}$$

and set

$$s=g(\lambda)\,. \tag{27.31}$$

Then, λ is given by

$$\lambda=s^{-\beta_3/\beta_2^2}\,\mathrm{e}^{-1/\beta_2 s}\,\tilde{\lambda}(s). \tag{27.32}$$

In equation (27.32), 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 (27.14) yields

$$\frac{t(\lambda)}{\lambda^2\Lambda^2}=\frac{t}{\Lambda^2}s^{2[2\nu_1+\beta_3/\beta_2]/\beta_2}\,\mathrm{e}^{2/\beta_2 s}\left[\tilde{t}(s)\right]^{-1}, \tag{27.33}$$

in which $\tilde{t}(s)$ is a regular function of s and $\nu(g)$ has been written as

$$\nu(g)=(2+\eta_2(g))^{-1}=\tfrac{1}{2}+\nu_1g+O(g^2). \tag{27.34}$$

Then, we determine λ from the condition (27.17), equation (27.33) parametrizes t as a function of s. At leading order, all critical behaviours will be described by a singular factor of the form occurring in equations (27.32) or (27.33) multiplied by regular series in $s=g(\lambda)$.

27.3 Irrelevant Operators and the Question of Universality

We now examine, in generic dimensions, the contributions coming from irrelevant operators. 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 $S(\phi)$ (for details, see Chapters 9,12):

$$S(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}(\partial_\mu \phi)^2_\Lambda + \frac{1}{2}m^2\phi^2 + \sum_\alpha u_\alpha \mathcal{O}_\alpha(\phi)\right], \tag{27.35}$$

in which $\mathcal{O}_\alpha(\phi)$ 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

$$[u_\alpha] = d - k_\alpha - \tfrac{1}{2}n_\alpha(d-2). \tag{27.36}$$

We treat all interactions in action (27.35) 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 (27.35) 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 $\mathcal{O}_\alpha$, it is necessary and generically sufficient that the condition

$$\Delta = d - [\mathcal{O}_\alpha] - \sum_l [u_{\beta_l}] \geq 0 \tag{27.37}$$

is satisfied. Then, the coefficient $\delta u_\alpha(\Lambda)$ of the counter-term proportional to $\mathcal{O}_\alpha$ diverges like a positive power of the cut-off Λ (or a power of logarithm if $\Delta = 0$),

$$\delta u_\alpha(\lambda) \sim \Lambda^\Delta. \tag{27.38}$$

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, $\phi^2(\partial_\mu\phi)^2$... which are operators of dimension 6 in four dimensions.

To introduce the cut-off Λ in our statistical problem we have rescaled the lengths and the field ϕ (equations (25.39–25.42)). Therefore, each coupling constant u_α is the product of a pure number g_α by a power of the cut-off which gives it its dimension:

$$u_\alpha = g_\alpha \Lambda^{-\delta_\alpha}. \tag{27.39}$$

Equation (27.36) gives

$$\delta_\alpha = -d + k_\alpha + \tfrac{1}{2}n_\alpha(d-2). \tag{27.40}$$

If δ_α is positive the corresponding operator $\mathcal{O}_\alpha$ leads to a non-renormalizable theory and we have already stated that it is irrelevant. In the tree approximation, the statement follows from equation (27.39): 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 can potentially compensate the powers coming from the coupling constants. To discover what happens it is necessary to analyse the counter-terms generated by these operators, using equations (27.37) and (27.38).

The total power Δ' of the cut-off which multiplies the operator $\mathcal{O}_\alpha$ in a counter-term is the sum of the power Δ generated by the divergence of perturbation theory (equation (27.37)) and the powers already present in the coefficients u_β (equation (27.39)):

$$\Delta' = \Delta - \sum_l \delta_{\beta_l} = \Delta + \sum_l [u_{\beta_l}], \tag{27.41}$$

and, therefore, using the definition (27.37):

$$\Delta' = d - [\mathcal{O}_\alpha] = [u_\alpha]. \tag{27.42}$$

The conclusion, therefore, is simple: due to the divergences of perturbation theory, irrelevant operators give indeed 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

$$\Lambda^{2\varepsilon-2}\int \mathrm{d}^d x\,\phi^6(x),\quad \Lambda^{\varepsilon-2}\int \mathrm{d}^d x(\phi\partial_\mu\phi)^2,\quad \Lambda^{-2}\int \mathrm{d}^d x\,\phi\partial^4\phi\,.$$

In terms of the renormalized operators they have the expansion

$$\int \mathrm{d}^d x\,\mathcal{O}_i^6(x) = \int \mathrm{d}^d x\,\left\{Z_{ij}\left[\mathcal{O}_j^6(x)\right]_\mathrm{r} + A_{ij}\left[\mathcal{O}_j^4(x)\right]_\mathrm{r} + B\left[\phi^2(x)\right]_\mathrm{r}\right\}.$$

We have denoted by $\mathcal{O}_j^4(x)$ the operators which have dimension 4 in four dimensions: ϕ^4 and $(\partial_\mu\phi)^2$. 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 $\Lambda^{-2+O(\varepsilon)}$.

This discussion also clarifies the interpretation of the constants r and g which parametrize the ϕ^4 hamiltonian. These are not the parameters which for instance appear as coefficients of the ϕ^2 or ϕ^4 terms in the mean field analysis, but instead effective parameters taking into account the effect of neglected irrelevant operators. The analysis of previous chapters is, however, 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, we note that a change in the cut-off procedure corresponds generally to a change in the coefficients of the irrelevant part of the propagator ($\phi\Delta^2\phi$...). 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 25,26, etc.

Remark. Another simple consequence of this analysis is the following: let us add to the hamiltonian an irrelevant operator which breaks a symmetry of the hamiltonian. Then, the symmetry of the critical theory will be broken if and only if the irrelevant operator can generate by renormalization relevant or marginal operators breaking the symmetry. An example is the following: on the lattice operators of the form $\sum_\mu \int \phi(x)(\partial_\mu)^4\phi(x)\mathrm{d}x$ 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 $\int (\partial_\mu\phi)^2\mathrm{d}^d x$, due to the hypercubic symmetry, are $O(d)$ symmetric, the $O(d)$ symmetry of the critical theory is not broken. Conversely, the addition of a naively irrelevant term like $\int \phi^5(x)\mathrm{d}^d x$ to a hamiltonian which is symmetric in $\phi \mapsto -\phi$ generates relevant terms linear in ϕ which are equivalent to the addition of a magnetic field.

27.4 Corrections Coming from Irrelevant Operators. Improved Action

For simplicity, we consider the effect, in the critical theory, of an operator $\mathcal{O}_\alpha$ 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 12 concerning the renormalization of composite operators.

27.4.1 Corrections to scaling

In Section 27.3, we have shown that an operator $\mathcal{O}_\alpha$ 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 $(\partial_\mu\phi(x))^2$ 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. Note that we can omit all operators which are total derivatives since only the space integrals appear in the action. We define a subtracted operator $\tilde{\mathcal{O}}_\alpha(\phi)$ by

$$\tilde{\mathcal{O}}_\alpha(\phi) = \mathcal{O}_\alpha(\phi) - \sum_{\beta \text{ such that } [\mathcal{O}_\beta]<[\mathcal{O}_\alpha]} C_{\alpha\beta}(\Lambda, g)\mathcal{O}_\beta(\phi). \tag{27.43}$$

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:

$$\tilde{\phi}^6(x) = \phi^6(x) - C_1\phi^2(x) - C_2\left[\partial_\mu\phi(x)\right]^2 - C_3\phi^4(x).$$

The coefficients C_1, C_2 and C_3 can be determined by a set of renormalization conditions at zero momentum:

$$\begin{aligned} \Gamma^{(2)}_{\tilde{\phi}^6}(p) &= O(p^4 \times \text{ powers of } \ln p) \text{ for } p \to 0\,, \\ \Gamma^{(4)}_{\tilde{\phi}^6}(p_i = 0) &= 0\,. \end{aligned} \tag{27.44}$$

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 25,26, etc.

Remark. Another simple consequence of this analysis is the following: let us add to the hamiltonian an irrelevant operator which breaks a symmetry of the hamiltonian. Then, the symmetry of the critical theory will be broken if and only if the irrelevant operator can generate by renormalization relevant or marginal operators breaking the symmetry. An example is the following: on the lattice operators of the form $\sum_\mu \int \phi(x)(\partial_\mu)^4\phi(x)\mathrm{d}x$ 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 $\int(\partial_\mu\phi)^2\mathrm{d}^d x$, due to the hypercubic symmetry, are $O(d)$ symmetric, the $O(d)$ symmetry of the critical theory is not broken. Conversely, the addition of a naively irrelevant term like $\int \phi^5(x)\mathrm{d}^d x$ to a hamiltonian which is symmetric in $\phi \mapsto -\phi$ generates relevant terms linear in ϕ which are equivalent to the addition of a magnetic field.

27.4 Corrections Coming from Irrelevant Operators. Improved Action

For simplicity, we consider the effect, in the critical theory, of an operator $\mathcal{O}_\alpha$ 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 12 concerning the renormalization of composite operators.

27.4.1 Corrections to scaling

In Section 27.3, we have shown that an operator $\mathcal{O}_\alpha$ 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 $(\partial_\mu\phi(x))^2$ 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. Note that we can omit all operators which are total derivatives since only the space integrals appear in the action. We define a subtracted operator $\tilde{\mathcal{O}}_\alpha(\phi)$ by

$$\tilde{\mathcal{O}}_\alpha(\phi) = \mathcal{O}_\alpha(\phi) - \sum_{\beta \text{ such that } [\mathcal{O}_\beta]<[\mathcal{O}_\alpha]} C_{\alpha\beta}(\Lambda, g)\mathcal{O}_\beta(\phi). \tag{27.43}$$

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:

$$\tilde{\phi}^6(x) = \phi^6(x) - C_1\phi^2(x) - C_2\left[\partial_\mu\phi(x)\right]^2 - C_3\phi^4(x).$$

The coefficients C_1, C_2 and C_3 can be determined by a set of renormalization conditions at zero momentum:

$$\begin{aligned}\Gamma^{(2)}_{\tilde{\phi}^6}(p) &= O(p^4 \times \text{ powers of } \ln p) \text{ for } p\to 0\,, \\ \Gamma^{(4)}_{\tilde{\phi}^6}(p_i=0) &= 0\,.\end{aligned} \tag{27.44}$$

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 $\mathcal{O}^{\mathrm{r}}_\alpha$ by

$$\mu^{-\delta_\alpha} \int \mathrm{d}^d x\, \mathcal{O}^{\mathrm{r}}_\alpha(x) = \sum_\beta Z_{\alpha\beta}(g, \Lambda/\mu)\Lambda^{-\delta_\beta} \int \mathrm{d}^d x\, \tilde{\mathcal{O}}_\beta(x), \qquad (27.45)$$

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 $\Gamma^{(n)}_{\tilde{\mathcal{O}}_\alpha}$ with one $\Lambda^{-\delta_\alpha} \int \mathrm{d}^d x\, \tilde{\mathcal{O}}_\alpha$ insertion and the renormalized functions with $\mu^{-\delta_\alpha} \int \mathrm{d}^d x\, \mathcal{O}_\alpha(x)$ insertion then is

$$\sum_\beta Z_{\alpha\beta} Z^{n/2} \Gamma^{(n)}_{\tilde{\mathcal{O}}_\beta}(p_i; g, \Lambda) = \left[\Gamma^{(n)}_{\mathcal{O}_\alpha}(p_i, g_{\mathrm{r}}, \mu)\right]_{\mathrm{ren}}. \qquad (27.46)$$

This leads to the RG equations:

$$\sum_\beta \left\{ \left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right] \delta_{\alpha\beta} - \eta_{\alpha\beta}(g) \right\} \Gamma^{(n)}_{\tilde{\mathcal{O}}_\beta} = 0 \qquad (27.47)$$

with

$$\eta_{\alpha\beta}(g) = -\sum_\gamma \left(\Lambda \frac{\partial}{\partial \Lambda} Z_{\alpha\gamma}\right) (Z^{-1})_{\gamma\beta}. \qquad (27.48)$$

Effects of the insertions of the operators $\tilde{\mathcal{O}}_\alpha$ are then governed by the values of η_α the eigenvalues of the matrix $\eta_{\alpha\beta}(g^*)$. From relation (27.45), we see that the renormalization matrix $Z_{\alpha\beta}$ has the form

$$Z_{\alpha\beta} = \delta_{\alpha\beta}(\Lambda/\mu)^{\delta_\alpha}\bigl(1 + O(g)\bigr). \qquad (27.49)$$

Therefore, the eigenvalues η_α are at leading order given by

$$\eta_\alpha = \delta_\alpha + O(\varepsilon).$$

The effects 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, $\delta = 2$. 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.

Remark. Note finally that in some cases the irrelevant effects may be specially 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 irrelevant corrections.

27.4.2 Fixed point in hamiltonian space and improved actions

We have seen 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 r.h.s. of RG equations (25.51), we have just neglected terms of the same order. Symanzik has shown in perturbation theory 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 r.h.s. of the RG equations (25.51), the operators of dimension 8 to cancel the order $1/\Lambda^4$ and so on. An iteration of this procedure leads to a theory which 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 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 constructing directly the renormalization group for a cut-off field theory, it is possible to prove the existence of the renormalized field theory (see Appendix A10.1).

Finally, let us point out that Symanzik has advocated the use of such improved actions (adding, for example, all terms of dimension 6) for numerical calculations on the lattice. It should be mentioned, however, that the applications of this ingenious idea are sometimes disappointing and 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.

27.5 Application: Uniaxial Systems with Strong Dipolar Forces

In Section 27.2, we have stressed that the renormalization group leads to exact predictions for critical systems at the upper-critical dimension. Unfortunately in the case of the N-vector model, the upper-critical dimension is 4 and, therefore, the predictions are not useful for experimental physics. The main application is numerical physics, for example, the Higgs sector ($N = 4$) of the Standard Model has been investigated numerically.

Therefore, here we present another system on which accurate measurements have been made, and which has dimension 3 as the upper-critical dimension: a uniaxial ferromagnet or ferroelectric system with strong dipolar forces.

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:

$$-\beta\mathcal{H}(\mathbf{S}) = \sum_{\mathbf{x},\mathbf{x}'} V_{\mu\nu}(\mathbf{x}-\mathbf{x}')S^\mu_{\mathbf{x}}S^\nu_{\mathbf{x}'} + \gamma(\mathbf{S}_{\mathbf{x}}\cdot\nabla_{\mathbf{x}})(\mathbf{S}_{\mathbf{x}'}\cdot\nabla_{\mathbf{x}'})\frac{1}{|\mathbf{x}-\mathbf{x}'|^{d-2}}, \qquad (27.50)$$

$V_{\mu\nu}(\mathbf{x})$ being the short range potential. We assume that the long range dipolar forces are strong enough to play a role in the part of the critical domain accessible experimentally. 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

$$-\beta\mathcal{H}(S)=\sum_{\mathbf{x},\mathbf{x}'} S_{\mathbf{x}}S_{\mathbf{x}'}\left[V(\mathbf{x}-\mathbf{x}')+\gamma\partial_z\partial_{z'}\frac{1}{|\mathbf{x}-\mathbf{x}'|^{d-2}}\right], \tag{27.51}$$

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

$$-\beta\mathcal{H}(S)=\int \mathrm{d}^d q\, S(\mathbf{q})S(-\mathbf{q})\left[\tilde{V}(\mathbf{q})+\tilde{\gamma}\frac{q_z^2}{q^2}\right], \tag{27.52}$$

where $\tilde{V}(\mathbf{q})$ is a regular function of $\mathbf{q}$ which, due to hypercubic symmetry, has the expansion

$$\tilde{V}(\mathbf{q})=a+\tfrac{1}{2}bq^2+O\left(q^4\right). \tag{27.53}$$

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:

$$q^2\sim q_z^2/\,q^2. \tag{27.54}$$

This implies that q_z, the z component of the vector $\mathbf{q}$, is much smaller than the other components $\mathbf{q}_\perp$:

$$|q_z|\sim(\mathbf{q}_\perp)^2. \tag{27.55}$$

We can, therefore, simplify further the interaction potential. In the case of an even spin distribution, we finally obtain an effective hamiltonian $\mathcal{H}(\phi)$ of the form

$$\begin{aligned}\mathcal{H}(\phi)=\frac{1}{2}\int \mathrm{d}^d q\,\phi(-\mathbf{q})\left(\mathbf{q}_\perp^2+A_0^2q_z^2/\mathbf{q}_\perp^2+r_c+t_0\right)\phi(\mathbf{q})\\ +\frac{u_0}{4!}\int \mathrm{d}^d q_1\ldots \mathrm{d}^d q_4\,\delta(\textstyle\sum\mathbf{q}_i)\phi(\mathbf{q}_1)\ldots\phi(\mathbf{q}_4).\end{aligned} \tag{27.56}$$

The upper-critical dimension. Usual power counting is now modified because space is no longer isotropic. In units of the transverse components of $\mathbf{q}$, the dimension of q_z is 2 (equation (27.55)):

$$[q_z]=2\quad\Rightarrow\quad[z]=-2\,.$$

The volume element in configuration space $\mathrm{d}\mathbf{x}_\perp\mathrm{d}z$ 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

$$d+1=4\quad\Rightarrow\quad d=3\,.$$

The renormalization group method thus leads to exact analytic predictions in the physical dimension $d=3$.

Renormalization group equations. From expression (27.56), we read off the propagator

$$\Delta(\mathbf{q})=\frac{\mathbf{q}_\perp^2}{(\mathbf{q}_\perp^2)^2+A^2q_z^2+t\mathbf{q}_\perp^2}\,. \tag{27.57}$$

Diagrams calculated with this propagator are regular for $\mathbf{q}$ small, therefore,

$$\int \mathrm{d}^d q\, \phi(-\mathbf{q}) \frac{q_z^2}{\mathbf{q}_\perp^2} \phi(\mathbf{q}),$$

will never appear as a counter-term.

The renormalized hamiltonian thus has the form

$$\begin{aligned}\mathcal{H}_\mathrm{r} = {} & \frac{1}{2}\int \mathrm{d}^d q\, \phi(-\mathbf{q}) \left(Z\mathbf{q}_\perp^2 + A^2 q_z^2/\mathbf{q}_\perp^2 + \delta m^2 + Z_2 t\right)\phi(\mathbf{q}) \\ & + \frac{\mu^\varepsilon}{4!} A\, g\, Z_g(g) \int \mathrm{d}^d q_1 \ldots \mathrm{d}^d q_4\, \delta(\textstyle\sum \mathbf{q}_i)\phi(\mathbf{q}_1)\ldots\phi(\mathbf{q}_4),\end{aligned} \qquad (27.58)$$

in which μ is the renormalization scale and ε now is defined as

$$d = 3 - \varepsilon\,. \qquad (27.59)$$

We introduce also a bare dimensionless coupling constant

$$u_0 = \Lambda^\varepsilon A_0 g_0\,. \qquad (27.60)$$

The relation between bare and renormalized correlation functions reads

$$\Gamma_\mathrm{r}^{(n)}(p_i; t, g, A, \mu) = Z^{n/2}\Gamma^{(n)}(p_i; t_0, g_0, A_0, \Lambda). \qquad (27.61)$$

In addition, comparing expressions (27.56) and (27.58), we find

$$A = Z^{1/2} A_0\,. \qquad (27.62)$$

RG equations follow

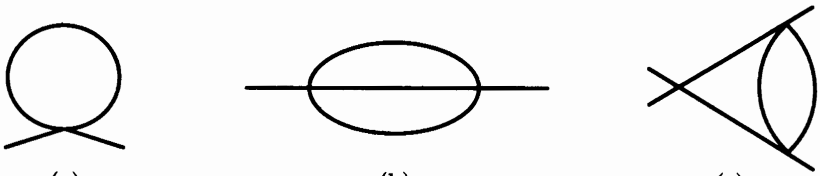
$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g_0)\frac{\partial}{\partial g_0} + \frac{1}{2}\eta(g_0)\left(A_0\frac{\partial}{\partial A_0} - n\right) - \eta_2(g_0)t_0\frac{\partial}{\partial t_0}\right]\Gamma^{(n)}(p_i; g_0, A_0, t_0, \Lambda) = 0\,. \qquad (27.63)$$

Two-loop calculation of RG functions. As was explained in Section 25.4, for practical calculations, we use the renormalized theory and minimal subtraction (see Chapter 11). In what follows we omit, for convenience, the index r on renormalized functions. The renormalized RG equations are formally identical to equations (27.63):

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} + \frac{1}{2}\eta(g)\left(A\frac{\partial}{\partial A} - n\right) - \eta_2(g)t\frac{\partial}{\partial t}\right]\Gamma^{(n)}(p_i; g, A, t, \Lambda) = 0\,, \qquad (27.64)$$

but the RG functions are different at two-loop order.

The calculations are here somewhat similar to the dynamic calculations as described in Appendix 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



(a) (b) (c)

Fig. 27.1 Diagrams needed at two-loop order.

momentum variables in the z direction (called ω hereafter). Figure 27.1 gives the three diagrams needed at two-loop order. In the massless theory, in terms of the propagator

$$\Delta(\omega, k) = \frac{1}{k^2 + \omega^2/k^2},$$

the expressions are

$$\begin{aligned}
(\mathrm{a}) &\equiv \frac{1}{(2\pi)^d} \int \mathrm{d}\omega\, \mathrm{d}^{d-1}k\, \Delta(\omega, k)\Delta(\omega, p+k) \\
&= \frac{1}{(2\pi)^{d-1}} \frac{1}{2} \int \frac{\mathrm{d}^{d-1}k}{k^2 + (p+k)^2} = \frac{1}{(16\pi)^{(d-1)/2}} \Gamma(\varepsilon/2) p^{-\varepsilon}. \\
(\mathrm{b}) &\equiv \frac{1}{(2\pi)^{2d}} \int \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, \mathrm{d}^{d-1}k_1\, \mathrm{d}^{d-1}k_2\, \Delta(\omega_1, k_1)\Delta(\omega_2, k_2)\Delta(\omega_1 + \omega_2, k_1 + k_2 + p) \\
&= \frac{1}{(2\pi)^{d-2}} \frac{1}{4} \int \frac{\mathrm{d}^{d-1}k_1 \mathrm{d}^{d-1}k_2}{k_1^2 + k_2^2 + (k_1 + k_2 + p)^2} = \frac{1}{(16\pi)^{d-1}} \frac{3}{8} \left(\frac{16}{27}\right)^{(d-1)/2} \Gamma(2-d) p^{2d-4}.
\end{aligned}$$

After integration over the corresponding ω variables (c) takes the form

$$\begin{aligned}
(\mathrm{c}) &= \frac{1}{(2\pi)^{2(d-1)}} \frac{1}{8} \left[C_1(p_1, p_2) + C_1(p_2, p_1) + C_2(p_1, p_2)\right], \\
C_1(p,q) &= \int \frac{\mathrm{d}^{d-1}k_1 \mathrm{d}^{d-1}k_2}{[k_1^2 + (k_1 + p + q)^2]\,[k_1^2 + k_2^2 + (k_1 + k_2 + p)^2]}, \\
C_2(p,q) &= \int \frac{\mathrm{d}^{d-1}k_1 \mathrm{d}^{d-1}k_2}{[(k_1 + p + q)^2 + k_2^2 + (k_1 + k_2 + p)^2]} \frac{1}{[k_1^2 + k_2^2 + (k_1 + k_2 + p)^2]}.
\end{aligned}$$

The final result is

$$\begin{aligned}
C_1(\varepsilon) &\simeq \frac{1}{(16\pi)^{d-1}} \left(\frac{2}{\sqrt{3}}\right)^{\varepsilon} \frac{\Gamma(1+\varepsilon)}{\varepsilon^2} (|p_1 + p_2|)^{-2\varepsilon} + O(1), \\
C_2(\varepsilon) &= \frac{2}{3} \frac{1}{(16\pi)^2 \varepsilon} + O(1).
\end{aligned}$$

The renormalization constants Z_g, Z and Z_2 are then determined up to order g^2:

$$Z_g = 1 + \frac{3N_d}{\varepsilon} g + \left(\frac{9}{\varepsilon^2} - \frac{3}{\varepsilon} \ln\frac{4}{3} - \frac{2}{\varepsilon}\right) N_d^2 g^2 + O(g^3), \tag{27.65}$$

$$Z = 1 - \frac{2}{27} N_d^2 \frac{g^2}{\varepsilon} + O(g^3), \tag{27.66}$$

$$Z_2^{-1} = 1 - N_d \frac{g}{\varepsilon} + \left(-\frac{1}{\varepsilon^2} + \frac{1}{3\varepsilon} + \frac{1}{2\varepsilon} \ln\frac{4}{3}\right) N_d^2 g^2 + O(g^3) \tag{27.67}$$

with

$$N_d = (16\pi)^{\varepsilon/2-1}\Gamma(1+\varepsilon/2). \tag{27.68}$$

The RG functions follow

$$\beta(g) = -\varepsilon \left[\frac{\mathrm{d}}{\mathrm{d}g}\ln(gZ_g Z^{-3/2})\right]^{-1}, \tag{27.69}$$

$$\eta(g) = \beta(g)\frac{\mathrm{d}}{\mathrm{d}g}\ln Z(g), \tag{27.70}$$

$$\eta_2(g) = \frac{1}{\nu(g)} - 2 = \frac{\mathrm{d}}{\mathrm{d}g}\ln(Z_2 Z^{-1}), \tag{27.71}$$

and, therefore,

$$N_d\beta(\tilde{g}) = -\varepsilon\tilde{g} + 3\tilde{g}^2 - \left(-6\ln\frac{4}{3} + \frac{34}{9}\right)\tilde{g}^3 + O(\tilde{g}^4), \tag{27.72}$$

$$\eta(\tilde{g}) = \frac{4}{27}\tilde{g}^2 + O(\tilde{g}^3), \tag{27.73}$$

$$\eta_2(\tilde{g}) = -\tilde{g} + \left(\frac{14}{27} + \ln\frac{4}{3}\right)\tilde{g}^2 + O(\tilde{g}^3) \tag{27.74}$$

with

$$\tilde{g} = N_d g\,. \tag{27.75}$$

Scaling behaviour below three dimensions. Dimensional analysis in the critical theory yields

$$\Gamma^{(n)}(\lambda\mathbf{p}_\perp, \rho p_z; t, g, A, \mu) = \lambda^{n+(n-2)(1-d)/2}\rho^{(2-n)/2} \times \Gamma^{(n)}\left(\mathbf{p}_\perp, p_z; t/\lambda^2, g, A\rho/\lambda^2, \mu/\lambda\right). \tag{27.76}$$

In $d = 3-\varepsilon$ dimensions, the model has an IR fixed point $g^*(\varepsilon)$. At the fixed point, we find

$$\Gamma^{(n)}(\mathbf{p}_\perp, p_z, t, A=\mu=1) = t^{\gamma-(n-2)d_\phi}\Gamma^{(n)}(\mathbf{p}_\perp t^\nu, p_z/t^{\nu(2-\eta/2)}) \tag{27.77}$$

with

$$\gamma = \nu(2-\eta), \qquad d_\phi = \tfrac{1}{2}(d-1+\eta).$$

At two-loop order, the exponents are

$$\tilde{g}^*(\varepsilon) = \frac{\varepsilon}{3} + \left(\frac{2}{9}\ln\frac{4}{3} + \frac{34}{243}\right)\varepsilon^2 + O(\varepsilon^3), \tag{27.78}$$

$$\eta = \frac{4}{243}\varepsilon^2 + O(\varepsilon^3), \tag{27.79}$$

$$\nu^{-1} = 2 - \frac{\varepsilon}{3} - \left(\frac{1}{9}\ln\frac{4}{3} + \frac{20}{243}\right)\varepsilon^2 + O(\varepsilon^3), \tag{27.80}$$

$$\omega = \varepsilon - \left(\frac{2}{3}\ln\frac{4}{3} + \frac{34}{81}\right)\varepsilon^2 + O(\varepsilon^3). \tag{27.81}$$

Logarithmic corrections to mean field behaviour in three dimensions. In three dimensions, the RG equations can be solved as indicated in Section 27.2.

For the effective coupling constant at scale λ, one finds

$$\frac{1}{g(\lambda)} \underset{\lambda\to 0}{\sim} \frac{3}{16\pi}\ln\frac{1}{\lambda}\left[1-2\frac{(17+27\ln 4/3)}{81}\frac{\ln|\ln\lambda|}{|\ln\lambda|}+O\left(\frac{1}{\ln\lambda}\right)\right]. \tag{27.82}$$

A short calculation then yields, for example, the susceptibility in zero field:

$$\chi^{-1} \sim C_{\pm}\left|\ln t\right|^{-1/3}\left[1+\frac{1}{243}(108\ln(4/3)+41)\frac{\ln|\ln t|}{|\ln t|}+O\left(\frac{1}{|\ln t|}\right)\right], \tag{27.83}$$

or the specific heat

$$C = A_{\pm}\left|\ln|t|\right|^{1/3}\left[1-\frac{1}{243}(108\ln(4/3)+41)\frac{\ln|\ln t|}{|\ln t|}+O\left(\frac{1}{|\ln t|}\right)\right] \tag{27.84}$$

with the universal ratio

$$\frac{A_+}{A_-}=\frac{1}{4}. \tag{27.85}$$

The specific heat has been measured in a high precision experiment on the dipolar Ising ferromagnet $LiTbF_4$ by Ahlers *et al.* Fitting the specific heat by

$$\begin{aligned} C_+ &= \frac{A_+}{b^z}\left\{[1+b\ln(a/|t|)]^z-1\right\}+B,\\ C_- &= \frac{A_-}{b^{z'}}\left\{[1+b\ln(a/|t|)]^{z'}-\tfrac{1}{4}\right\}+B, \end{aligned}$$

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.

Bibliographical Notes

Correction terms due to irrelevant operators are discussed in

F. Wegner, *Phys. Rev.* B5 (1972) 4529; B6 (1972) 1891, and in Wegner's contribution to *Phase Transitions and Critical Phenomena*, vol. 6, quoted in Chapter 24; F.J. Wegner and E.K. Riedel, *Phys. Rev.* B7 (1973) 248; E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D8 (1973) 434, 2418 and in *Phase Transitions and Critical Phenomena* vol. 6 already quoted, Section VIII.

Recent calculations of dimensions of operators can be found in

S.K. Kehrein and F. Wegner, *Nucl. Phys.* B424 (1994) 521.

The logarithmic corrections in four dimensions for the ϕ^4 theory and the uniaxial system with dipolar forces have been obtained, with the use of the parquet approximation by

A.I. Larkin and D.E. Khmel'nitskii, *Zh. Eksp. Teor. Fiz.* 56 (1969) 647 (*JETP* 29, 1123).

The idea of improved action has been discussed in

K. Symanzik, in *Mathematical Problems in Theoretical Physics (Lecture Notes in Physics 153)*, R. Schrader, R. Seiler and D.A. Uhlenbrock eds. (Springer-Verlag, Berlin 1982); *Nucl. Phys.* B226 (1983) 187, 205.

The application of the renormalization group to the uniaxial system with dipolar forces is found in Wegner and Riedel as quoted above and

A. Aharony, *Phys. Rev.* B8 (1973) 3363; E. Brézin and J. Zinn-Justin, *Phys. Rev.* B13 (1976) 251.

The experimental results are quoted from

G. Ahlers, A. Kornblit and H.J. Guggenheim, *Phys. Rev. Lett.* 34 (1975) 1227.

28 NON-MAGNETIC SYSTEMS AND THE $(\phi^2)^2$ FIELD THEORY

We have already pointed out that many systems to which the RG predictions based on the ϕ^4 field theory apply are non-magnetic. We consider here three systems or classes of systems which are of special interest because they allow precise comparisons between RG results, experiments and simulations: long polymer chains (self-avoiding random walks), liquid–vapour and superfluid Helium (see Chapter 29).

We want to show in this chapter that indeed their critical properties are related to the N-component ϕ^4 field theory: $N = 0$ describes the statistical properties of polymer chains asymptotically when the length becomes large, $N = 1$ and $N = 2$ the universal properties of the liquid–vapour and Helium superfluid transitions, respectively.

28.1 Statistics of Self-Repelling Chains, Approximations

We first show that 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 "unphysical" $N = 0$ limit. More precisely, we study the statistical properties of long chains with repulsive contact self-interaction.

We call u the distance along the chain, $\mathbf{r}(u)$ the position in space of the point on the chain of parameter u. We then characterize the chain by a probability distribution for a chain $\mathbf{r}(u)$ of total length s:

$$[\mathrm{d}\rho(\mathbf{r}(u))] = [\mathrm{d}\mathbf{r}(u)] \exp\left\{-\left[\frac{1}{4}\int_0^s \dot{\mathbf{r}}^2(u)\mathrm{d}u + \frac{g}{6}\int_0^s \mathrm{d}u_1\mathrm{d}u_2\,\delta^d\big(\mathbf{r}(u_1)-\mathbf{r}(u_2)\big)\right]\right\}. \tag{28.1}$$

The special limit $g = 0$ corresponds to the brownian chain or gaussian random walk, as has been discussed in Chapter 4. The self-avoiding random walk (SAW) 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 (28.1). Note finally that any short range potential would yield, at leading order for long chains, the same results as the δ-function interaction (as can be verified 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).

A generating function. Various characteristic properties of the chain can be derived from the two-point function

$$G^{(2)}(\mathbf{k},s) = \left\langle \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{r}(s)-\mathbf{r}(0))}\right\rangle, \tag{28.2}$$

where brackets mean average with respect to the distribution (28.1). Indeed, the expansion of $G^{(2)}(\mathbf{k},s)$ in powers of $\mathbf{k}$ yield the successive moments of the distribution of the origin to end positions:

$$G^{(2)}(\mathbf{k},s) = 1 - \frac{1}{2!}\frac{\mathbf{k}^2}{d}\left\langle(\mathbf{r}(s)-\mathbf{r}(0))^2\right\rangle + \frac{(\mathbf{k}^2)^2}{4!}\frac{1}{d(d+2)}\left\langle(\mathbf{r}(s)-\mathbf{r}(0))^4\right\rangle + \cdots, \tag{28.3}$$

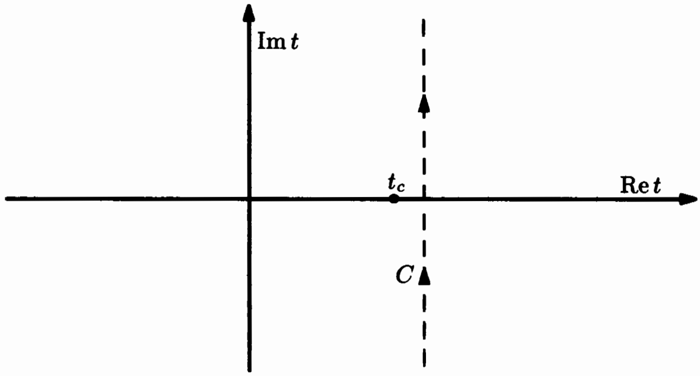


Fig. 28.1

where the rotation symmetry of the distribution (28.1) has been used.

It is actually easier to calculate the Laplace transform of $G^{(2)}(\mathbf{k},s)$ with respect to s:

$$Z^{(2)}(\mathbf{k},t)=\int_0^\infty \mathrm{e}^{-st}\,G^{(2)}(\mathbf{k},s)\mathrm{d}s\,. \tag{28.4}$$

The function $G^{(2)}(\mathbf{k},s)$ is then recovered by inverting the Laplace transformation

$$G^{(2)}(\mathbf{k},s)=\frac{1}{2i\pi}\oint_C \mathrm{e}^{st}\,Z^{(2)}(\mathbf{k},t)\mathrm{d}t\,, \tag{28.5}$$

in which the contour C is a parallel to the imaginary t axis at the right of all singularities of $Z^{(2)}(\mathbf{k},t)$. Since only the expansion of $G^{(2)}(\mathbf{k},s)$ in powers of $\mathbf{k}$ (equation (28.3)) is useful, $\mathbf{k}$ has to be considered as infinitesimal. Thus, only the location of singularities in the limit $\mathbf{k}\to\mathbf{0}$ is relevant. Let us call t_c the first singularity one meets while trying to displace the contour to the left (figure 28.1). On C the real part $\mathrm{Re}(t)$ of t thus satisfies

$$\mathrm{Re}\,(t)>\mathrm{Re}(t_c).$$

One verifies that the behaviour of $Z^{(2)}(\mathbf{k},t)$ near the singularity governs the large positive s behaviour of the integral (28.5) and thus of the moments of $\mathbf{r}(s)-\mathbf{r}(0)$.

28.1.1 Special examples. Flory's approximation

The gaussian random walk. In expression (28.2), we can substitute

$$\mathbf{r}(s)-\mathbf{r}(0)=\int_0^s \dot{\mathbf{r}}(u)\mathrm{d}u\,.$$

In the absence of self-repulsion, we then get

$$G^{(2)}(\mathbf{k},s)=\int[\mathrm{d}\mathbf{r}(u)]\exp\left\{-\int_0^s \mathrm{d}u\left[\frac{1}{4}\dot{\mathbf{r}}^2(u)-i\mathbf{k}\cdot\dot{\mathbf{r}}(u)\right]\right\}.$$

Shifting $\mathbf{r}$, we find

$$G^{(2)}(\mathbf{k},s)=\mathrm{e}^{-s\mathbf{k}^2}, \tag{28.6}$$

which implies the well-known scaling laws of the gaussian random walk:

$$\left\langle (\mathbf{r}(s)-\mathbf{r}(0))^{2n} \right\rangle \sim s^n. \tag{28.7}$$

The corresponding Laplace transform $Z^{(2)}(\mathbf{k},t)$ then takes a free field form

$$Z^{(2)}(\mathbf{k},t) = (t+\mathbf{k}^2)^{-1}. \tag{28.8}$$

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:

$$r(s) = \pm s\,, \tag{28.9}$$

and, therefore,

$$G^{(2)}(k,s) = \cos ks\,, \qquad Z^{(2)}(k,t) = \frac{t}{t^2+k^2}. \tag{28.10}$$

The one-dimensional example shows that the self-avoiding condition can change the statistical properties of the random walk. It also provides an upper bound on the moments. A reasonable guess for generic dimensions, then, is

$$\left\langle (\mathbf{r}(s)-\mathbf{r}(0))^{2n} \right\rangle \sim s^{2\nu n} \qquad \text{for } s\to\infty\,,$$

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. If we consider a very long chain, we can consider the effect of self-avoiding as the influence of one chain onto another. A brownian chain has Hausdorf dimension 2. Therefore, above dimension 4 (2 +2) two chains do no more 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. The approximation is based on an energy balance argument. For s large, the gaussian term should scale as the interaction:

$$\left\langle \int_0^s \dot{\mathbf{r}}^2(u)\mathrm{d}u \right\rangle \sim s^{2\nu-1}, \tag{28.11}$$

$$\int \mathrm{d}u_1 \mathrm{d}u_2\, \delta^d(\mathbf{r}(u_1)-\mathbf{r}(u_2)) \sim s^{2-d\nu}. \tag{28.12}$$

Balancing both terms, we obtain an equation for ν:

$$2\nu-1 = 2-d\nu \;\Rightarrow\; \nu = \frac{3}{d+2}, \tag{28.13}$$

valid for $1\le d\le 4$. This expression yields for $d=1$ the exact value $\nu=1$. It predicts that for $d\ge 4$ 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.

We shall discuss below other dimensions. 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 if this result is not exact. We show below that it is indeed only an approximation. For this purpose, it is convenient to completely reformulate the theory.

28.1.2 Equivalence with the ϕ^4 field theory

We first transform the interaction term in equation (28.1) by writing it as a gaussian integral over an auxiliary field $\sigma(\mathbf{r})$. If $\mathbf{r}(u)$ is a given chain, then,

$$\int [\mathrm{d}\sigma(\mathbf{r})] \exp\left[\frac{3}{2g}\int \mathrm{d}^d r\, \sigma^2(r) - \int \mathrm{d}u\, \sigma\big(\mathbf{r}(u)\big)\right]$$
$$= \exp\left[-\frac{g}{6}\int \mathrm{d}u_1 \mathrm{d}u_2\, \delta^d(\mathbf{r}(u_1)-\mathbf{r}(u_2))\right], \tag{28.14}$$

where we have used

$$\int \mathrm{d}u\, \sigma(\mathbf{r}(u)) = \int \mathrm{d}^d r\, \sigma(r) \int \mathrm{d}u\, \delta^d(\mathbf{r}(u)-\mathbf{r}).$$

Integration over imaginary σ-fields is assumed. After this substitution, we identify the $\mathbf{r}(u)$ measure of integration 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 $Z^{(2)}(\mathbf{k},t)$ can thus be written as

$$Z^{(2)}(\mathbf{k},t) = \int [\mathrm{d}\sigma(\mathbf{r})] \exp\left[\frac{3}{2g}\int \mathrm{d}^d r\, \sigma^2(r)\right] \int_0^\infty \mathrm{e}^{-ts}\, \mathrm{d}s$$
$$\times \int \mathrm{d}^d r\, \mathrm{d}^d r'\, \mathrm{e}^{i\mathbf{k}(\mathbf{r}-\mathbf{r}')} \left\langle \mathbf{r}' \left| \mathrm{e}^{-sH} \right| \mathbf{r} \right\rangle, \tag{28.15}$$

in which H is the quantum hamiltonian:

$$H = -\nabla^2 + \sigma(\mathbf{r}). \tag{28.16}$$

The Laplace transform can be explicitly calculated:

$$Z^{(2)}(\mathbf{k},t) = \int [\mathrm{d}\sigma(\mathbf{r})] \exp\left[\frac{3}{2g}\int \mathrm{d}^d r\, \sigma^2(r)\right] \int \mathrm{d}^d r\, \mathrm{d}^d r'\, \mathrm{e}^{i\mathbf{k}(\mathbf{r}-\mathbf{r}')}$$
$$\times \left\langle \mathbf{r}' \left| (-\nabla^2 + t + \sigma)^{-1} \right| \mathbf{r} \right\rangle. \tag{28.17}$$

We now apply a variant of the so-called *replica trick*:

$$\lim_{N\to 0} \int [\mathrm{d}\boldsymbol{\phi}(\mathbf{r})]\, \phi_1(\mathbf{r})\phi_1(\mathbf{r}') \exp\left\{-\frac{1}{2}\int \mathrm{d}^d r \left[\big(\partial_\mu \boldsymbol{\phi}(r)\big)^2 + \big(t+\sigma(r)\big)\boldsymbol{\phi}^2(r)\right]\right\}$$
$$= \left\langle \mathbf{r}' \left| (-\nabla^2 + t + \sigma)^{-1} \right| \mathbf{r} \right\rangle, \tag{28.18}$$

in which N is the number of components of the field $\boldsymbol{\phi}(x)$. Indeed, the gaussian integration over the field $\boldsymbol{\phi}(r)$ yields the $\phi\phi$ propagator divided by a factor $[\det(-\nabla^2+t+\sigma)]^{N/2}$ which goes to 1 in the "unphysical" $N=0$ limit.

Substituting this identity into expression (28.17), we obtain

$$Z^{(2)}(\mathbf{k},t) = \int [\mathrm{d}\boldsymbol{\phi}]\, \phi_1(\mathbf{k})\phi_1(-\mathbf{k}) \int [\mathrm{d}\sigma(r)]$$
$$\times \exp\left[\frac{1}{2}\int \mathrm{d}^d r\, \left(3\sigma^2/g - \sigma\boldsymbol{\phi}^2 - (\partial_\mu\boldsymbol{\phi})^2 - t\boldsymbol{\phi}^2\right)\right]. \tag{28.19}$$

Then, we can integrate over σ:

$$Z^{(2)}(\mathbf{k},t) = \int [\mathrm{d}\boldsymbol{\phi}]\, \tilde\phi_1(\mathbf{k})\tilde\phi_1(-\mathbf{k}) \exp\left[-\mathcal{S}(\boldsymbol{\phi})\right],$$
$$\mathcal{S}(\boldsymbol{\phi}) = \int \mathrm{d}^d r \left[\frac{1}{2}(\partial_\mu\boldsymbol{\phi})^2 + \frac{1}{2}t\boldsymbol{\phi}^2 + \frac{g}{4!}(\boldsymbol{\phi}^2)^2\right]. \tag{28.20}$$

This is a most remarkable result: the statistical properties of polymers are related to the properties of the $(\boldsymbol{\phi}^2)^2$ field theory in the $N=0$ limit as first noticed by de Gennes.

28.1.3 RG approach to 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 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 $Z^{(2)}(k,t)$ has a scaling behaviour:

$$Z^{(2)}(\mathbf{k},t) \sim (t-t_c)^{-\gamma} f\left[\mathbf{k}(t-t_c)^{-\nu}\right], \tag{28.21}$$

in which ν and γ are the $N=0$ limits of the critical exponents of the $(\phi^2)^2$ field theory.

Inserting the form (28.21) into the integral (28.5), we then obtain the scaling behaviour of $G^{(2)}(s,k)$ for s large:

$$G^{(2)}(s,\mathbf{k}) \sim \mathrm{e}^{st_c}\, s^{\gamma-1} g(|\mathbf{k}|\, s^{\nu}). \tag{28.22}$$

Expanding in powers of $\mathbf{k}^2$, we find a first term

$$\langle 1\rangle \sim \mathrm{e}^{st_c}\, s^{\gamma-1} g(0),$$

instead of 1 which shows that the distribution is not correctly 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 $\langle \exp\left[i\mathbf{k}(r(s)-r(0))\right]\rangle$ has a scaling behaviour entirely characterized by the exponent ν. For example,

$$\left\langle (\mathbf{r}(s)-\mathbf{r}(0))^2\right\rangle \sim s^{2\nu}. \tag{28.23}$$

The exponent can be obtained from the ε-expansion:

$$\nu = \tfrac{1}{2}\left(1+\tfrac{1}{8}\varepsilon+\tfrac{15}{256}\varepsilon^2\right)+O(\varepsilon^3). \tag{28.24}$$

Instead, Flory's formula (28.13) predicts at the same order:

$$\nu(\text{Flory}) = \tfrac{1}{2}\left(1+\tfrac{1}{6}\varepsilon\right)+O(\varepsilon^2). \tag{28.25}$$

Comparing the two expressions, we see 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 for $d=1$ and 2.

For $d=3$, RG calculations yield

$$\nu = 0.5880 \pm 0.0015\,,$$

result which can be compared with $\nu(\text{Flory})=0.6$. The two values are close but definitively different. Actually, one can estimate the difference approximately. If we simply take into account the orders ε and ε^2 and the property that Flory's approximation is exact for $\varepsilon=0,2,3$, we can write

$$\nu_{\text{RG}}^{-1}-\nu^{-1}(\text{Flory}) = \tfrac{1}{12}\varepsilon(1-\tfrac{1}{2}\varepsilon)(1-\tfrac{1}{3}\varepsilon)(1-\tfrac{19}{96}\varepsilon)+O(\varepsilon^3)).$$

For $\varepsilon=1$, we find 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 29.4,29.6). Experimental results, $\nu=0.586\pm0.004$, HT series and Monte Carlo simulations, $\nu=0.5877\pm0.0006$, agree with the RG prediction and exclude the value predicted by Flory's approximation.

Moreover, the renormalization group makes many additional predictions and allows the calculation of many universal quantities which cannot be obtained from Flory's argument.

28.2 Liquid–Vapour Phase Transition and Field Theory

It is not completely obvious *a priori* that the liquid–vapour phase transition in classical systems is in the same universality class as the Ising model and can be described by the ϕ^4 field theory. In particular, no apparent discrete symmetry is broken. There exist several methods to establish a connection between 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, obtain directly a ϕ^4 field theory.

The lattice gas model. One assumes that particles are living on a lattice and that at each site i 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 represented by a nearest neighbour interaction which favours the occupation of neighbour sites. One finally adds a chemical potential term, that is, a source term coupled to the number of particles:

$$-\beta\mathcal{H}(n_i) = J\sum_{\text{n.n.}} n_i n_j - \mu\sum_i n_i . \tag{28.26}$$

This model is directly related to the Ising model by the transformation:

$$S_i = 2n_i - 1 , \tag{28.27}$$

where S_i is an Ising spin. One finds in d dimensions:

$$-\beta\mathcal{H}(S_i) = \frac{1}{4}J\sum_{\text{n.n.}} S_i S_j + \frac{1}{2}(dJ - \mu)\sum_i S_i + \text{ const.}. \tag{28.28}$$

We recognize the energy of the Ising model with n.n. interaction and magnetic field. Since we have shown that the Ising model leads to the ϕ^4 field theory the same thus applies to the lattice gas model, and the equivalence between the two models allows to translate the RG results into the language of classical fluids.

28.2.1 The classical gas in the continuum: functional integrals

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 functional integral representation of the partition function (in the grand canonical formalism). Then, a few additional steps, that are intuitive but harder to justify rigorously, lead to the ϕ^4 field theory.

We consider a real classical gas in the continuum and assumes that the potential between particles is the sum of a hard-core potential, with a core of size a, and a short range attractive two-body potential $V(\mathbf{r})$:

$$V(\mathbf{r}) = V(r = |\mathbf{r}|) \le 0 , \quad \forall\, \mathbf{r} ,$$

but with a range much larger than a. The canonical partition function for n particles of mass m in a d-dimensional volume Ω then has the form

$$\mathcal{Z}(n,\beta,\Omega) = \left(\frac{2\pi m}{\beta h^2}\right)^{nd/2} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^{n} \mathrm{d}^d r_i \exp\left[-\beta \sum_{i<j} V(\mathbf{r}_i - \mathbf{r}_j)\right] . \tag{28.29}$$

Since the potential depends only on the distance r and is short range, its Fourier transform

$$\tilde{V}(\mathbf{k}) = \int \mathrm{d}^d r\, \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}}\, V(\mathbf{r}),$$

is an analytic function of k^2 in a strip, and can be expanded for k^2 small as

$$\tilde{V}(\mathbf{k}) = -v(1-\sigma^2\mathbf{k}^2) + O(k^4), \quad v > 0\,, \tag{28.30}$$

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 now introduce the local gas density

$$\rho(\mathbf{r}) = \sum_{i=1}^{n} \delta(\mathbf{r}-\mathbf{r}_i), \quad \text{and thus} \quad \int \mathrm{d}^d r\, \rho(\mathbf{r}) = n\,, \tag{28.31}$$

where $\mathbf{r}_i$ is the position of a particle in the gas. The gas potential energy can then be written as

$$\sum_{i<j} V(\mathbf{r}_i - \mathbf{r}_j) = \frac{1}{2}\int \mathrm{d}^d r\, \mathrm{d}^d r' V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r})\rho(\mathbf{r}') - \frac{1}{2} n V(0).$$

We now impose the relation (28.31) by a functional integral in the spirit of the method of Section 24.3.1 and insert the identity

$$\int [\mathrm{d}\rho(\mathbf{r})] \int [\mathrm{d}\phi(\mathbf{r})] \exp\left[\int \mathrm{d}^d r\, \phi(\mathbf{r})\rho(\mathbf{r}) - \sum_{i=1}^{n} \phi(\mathbf{r}_i)\right] = \text{const.}$$

into the expression (28.29) (note that the integral over ϕ represents a δ function and thus the contour is parallel to the imaginary axis). After these transformations the partition function becomes

$$\begin{aligned} \mathcal{Z}(n,\beta,\Omega) = \mathrm{e}^{n\beta V(0)/2} \int [\mathrm{d}\phi(r)\mathrm{d}\rho(r)] \exp\Bigg[& -\frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \\ & + \int \mathrm{d}^d r\, \rho(\mathbf{r})\phi(\mathbf{r})\Bigg] \mathcal{Z}_{\text{h.c.}}(n,\beta,\Omega,\phi), \end{aligned} \tag{28.32}$$

where $\mathcal{Z}_{\text{h.c.}}(n,\beta,\Omega,\phi)$ is the canonical partition function of particles interacting through a hard-core two-body potential in a random external (imaginary) potential $\phi(\mathbf{r})$:

$$\mathcal{Z}_{\text{h.c.}}(n,\beta,\Omega,\phi) = \left(\frac{2\pi m}{\beta h^2}\right)^{nd/2} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^{n} \mathrm{d}^d r_i \exp\left[-\sum_i \phi(\mathbf{r}_i)\right]. \tag{28.33}$$

We can transform equation (28.32) into an identity between grand canonical partition functions:

$$\mathcal{Z}(\mu,\beta,\Omega) = \sum_n \frac{z^n}{n!} \mathcal{Z}(n,\beta,\Omega), \tag{28.34}$$

where z is the gas fugacity. We set

$$z = \mathrm{e}^{\beta(\mu - V(0)/2)}, \quad \rho_0 = (2\pi)^{d/2}/\lambda^d\,, \tag{28.35}$$

where μ is some renormalized chemical potential and λ the thermal wavelength:

$$\lambda = \hbar\sqrt{\beta/m}\,. \tag{28.36}$$

The normalization density ρ_0, which can be changed by shifting the chemical potential, thus increases with the temperature.

Writing n as the space integral of the local density ρ, and using the integral representation (28.32), we obtain the functional integral

$$\mathcal{Z}(\mu,\beta,\Omega) = \int [\mathrm{d}\phi(r)\mathrm{d}\rho(r)] \exp\left[-\mathcal{S}(\rho,\phi)\right], \tag{28.37}$$

with the action

$$\begin{aligned}\mathcal{S}(\rho,\phi) = {}& \frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \\ & - \int \mathrm{d}^d r\, \rho(\mathbf{r})\left(\phi(\mathbf{r}) + \beta\mu\right) - \ln \mathcal{Z}_{\text{h.c.}}(\Omega,\phi)\end{aligned} \tag{28.38}$$

and

$$\mathcal{Z}_{\text{h.c.}}(\Omega,\phi) = \sum_n \frac{\rho_0^n}{n!} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^n \mathrm{d}^d r_i \exp\left[-\sum_i \phi(\mathbf{r}_i)\right]. \tag{28.39}$$

Note that if we replace the chemical potential μ by an external potential $\mu(\mathbf{r})$, we transform the partition function (28.39) 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 will face in Chapter 33. Here, instead 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. Therefore, we instead introduce the integral over ϕ, even though we cannot calculate it exactly.

Finally, we note that the expression (28.37) can easily be generalized to many-body forces, leading to an action which is a polynomial of higher degree in the local density ρ.

28.2.2 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 functional integral:

$$\begin{aligned}\mathcal{Z}(\mu,\beta,\Omega) &= \int [\mathrm{d}\phi(r)\mathrm{d}\rho(r)] \exp\left[-\mathcal{S}(\rho,\phi)\right] \\ \mathcal{S}(\rho,\phi) &= \frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \\ & \quad - \int \mathrm{d}^d r \left(\rho(\mathbf{r})\left(\phi(\mathbf{r}) + \beta\mu\right) + \rho_0\, \mathrm{e}^{-\phi(\mathbf{r})}\right).\end{aligned} \tag{28.40}$$

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

$$\rho(\mathbf{r}) = \rho_0\, \mathrm{e}^{-\phi(\mathbf{r})} \;\Rightarrow\; \phi(\mathbf{r}) = -\ln\left(\rho(\mathbf{r})/\rho_0\right).$$

where μ is some renormalized chemical potential and λ the thermal wavelength:

$$\lambda = \hbar\sqrt{\beta/m}\,. \tag{28.36}$$

The normalization density ρ_0, which can be changed by shifting the chemical potential, thus increases with the temperature.

Writing n as the space integral of the local density ρ, and using the integral representation (28.32), we obtain the functional integral

$$\mathcal{Z}(\mu,\beta,\Omega) = \int [\mathrm{d}\phi(r)\mathrm{d}\rho(r)] \exp\left[-\mathcal{S}(\rho,\phi)\right], \tag{28.37}$$

with the action

$$\begin{aligned}\mathcal{S}(\rho,\phi) = {} & \frac{1}{2}\beta\int \mathrm{d}^d r\,\mathrm{d}^d r'\,\rho(\mathbf{r})V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \\ & - \int \mathrm{d}^d r\,\rho(\mathbf{r})\big(\phi(\mathbf{r})+\beta\mu\big) - \ln \mathcal{Z}_{\text{h.c.}}(\Omega,\phi)\end{aligned} \tag{28.38}$$

and

$$\mathcal{Z}_{\text{h.c.}}(\Omega,\phi) = \sum_n \frac{\rho_0^n}{n!}\int_{\substack{\mathbf{r}_i\in\Omega \\ |\mathbf{r}_i-\mathbf{r}_j|>2a}} \prod_{i=1}^{n} \mathrm{d}^d r_i \exp\left[-\sum_i \phi(\mathbf{r}_i)\right]. \tag{28.39}$$

Note that if we replace the chemical potential μ by an external potential $\mu(\mathbf{r})$, we transform the partition function (28.39) 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 will face in Chapter 33. Here, instead 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. Therefore, we instead introduce the integral over ϕ, even though we cannot calculate it exactly.

Finally, we note that the expression (28.37) can easily be generalized to many-body forces, leading to an action which is a polynomial of higher degree in the local density ρ.

28.2.2 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 functional integral:

$$\begin{aligned}\mathcal{Z}(\mu,\beta,\Omega) &= \int [\mathrm{d}\phi(r)\mathrm{d}\rho(r)] \exp\left[-\mathcal{S}(\rho,\phi)\right] \\ \mathcal{S}(\rho,\phi) &= \frac{1}{2}\beta\int \mathrm{d}^d r\,\mathrm{d}^d r'\,\rho(\mathbf{r})V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') \\ & \quad - \int \mathrm{d}^d r\big(\rho(\mathbf{r})\big(\phi(\mathbf{r})+\beta\mu\big) + \rho_0\,\mathrm{e}^{-\phi(\mathbf{r})}\big).\end{aligned} \tag{28.40}$$

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

$$\rho(\mathbf{r}) = \rho_0\,\mathrm{e}^{-\phi(\mathbf{r})} \;\Rightarrow\; \phi(\mathbf{r}) = -\ln\big(\rho(\mathbf{r})/\rho_0\big)\,.$$

The action becomes

$$\mathcal{S}(\rho) = \frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') + \int \mathrm{d}^d r\, \rho(\mathbf{r}) \left[\ln\left(\rho(\mathbf{r})/\rho_0\right) - 1 - \beta\mu\right]. \quad (28.41)$$

We now look for a homogeneous saddle point ρ. With the notation (28.30), we obtain

$$-\beta v \rho + \ln(\rho/\rho_0) = \beta\mu\,, \quad (28.42)$$

which allows us to eliminate the chemical potential. The action per unit volume is

$$\mathcal{S}/\Omega = -\rho + \tfrac{1}{2}\beta v \rho^2 = -\beta p\,,$$

where p is 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 $\ln \mathcal{Z}_{\text{h.c.}}(\phi)$ governed by the length scale a, in the successive terms of the sum (28.39). Using the leading order equation (28.42), 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 combination which appears is $a^d\rho$ 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

$$\mathcal{S}(\rho) = \frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') - \int \mathrm{d}^d r \left[\rho(\mathbf{r})\left(\phi(\mathbf{r}) + \beta\mu\right) + w\left(\phi(\mathbf{r})\right)\right], \quad (28.43)$$

where $w(\phi)$ is the hard-core free energy in the thermodynamic limit, as a function of a chemical potential $-\phi/\beta$:

$$\exp\left[\Omega w(\phi)\right] = 1 + \sum_{n=1} \frac{\rho_0^n\, \mathrm{e}^{-n\phi}}{n!} \int_{\substack{\mathbf{r}_i \in \Omega \\ |\mathbf{r}_i - \mathbf{r}_j| > 2a}} \prod_{i=1}^{n} \mathrm{d}^d r_i\,.$$

Again, we integrate over ϕ by the steepest descent method and replace ϕ by the solution of the saddle point equation:

$$\rho(\mathbf{r}) + w'\left(\phi(\mathbf{r})\right) = 0\,. \quad (28.44)$$

We introduce the Legendre transform of $w(\phi)$:

$$U(\rho) = -w(\phi) - \rho\phi\,, \quad (28.45)$$

in such a way that the approximated action now reads

$$\mathcal{S}(\rho) = \frac{1}{2}\beta \int \mathrm{d}^d r\, \mathrm{d}^d r' \rho(\mathbf{r}) V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') + \int \mathrm{d}^d r \left[U\left(\rho(\mathbf{r})\right) - \beta\mu\rho(\mathbf{r})\right]. \quad (28.46)$$

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

$$\mathrm{e}^{\Omega F(z)} \underset{\Omega\to\infty}{=} 1 + \sum_{n=1} \frac{z^n}{n!} \int_{\substack{r_i \in \Omega \\ |r_i - r_j| > 2a}} \prod_{i=1}^{n} \mathrm{d}r_i \,,$$

one finds that F is solution of the transcendental equation:

$$F(z) = z\,\mathrm{e}^{-aF(z)} \,.$$

Of particular interest is the behaviour of the function for $z \to +\infty$. One finds

$$aF = \ln z - \ln\ln z + \cdots \,.$$

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,

$$F(z) \underset{z\to\infty}{\propto} a^{-d} \ln z \,.$$

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 (28.46) in the tree approximation, following the lines of Section 24.3.2, where the mean field approximation was considered. We look for a homogeneous saddle point ρ:

$$\mathcal{S}'(\rho)/\Omega = -\beta v \rho + U'(\rho) - \beta\mu = 0 \,. \tag{28.47}$$

Depending on the number of solutions of the equation, we will find 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 $\rho \to 0$, and is expected to diverge at the close packing density. It has, therefore, an absolute minimum at a 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 thermodynamical 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 (28.47), 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

$$\mathcal{S}(\rho)/\Omega = \tfrac{1}{2}s_2(\beta)(\rho-\rho_c)^2 + \tfrac{1}{4!}s_4(\rho-\rho_c)^4 + O\big((\rho-\rho_c)^5\big)$$

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 (28.30)):

$$\frac{1}{2}\beta\int \mathrm{d}^d r\,\mathrm{d}^d r'\rho(\mathbf{r})V(\mathbf{r}-\mathbf{r}')\rho(\mathbf{r}') = \frac{1}{2}\beta v\int \mathrm{d}^d r\left[-\rho^2 + \sigma^2(\nabla\rho)^2\right].$$

The universal properties of the liquid–vapour transition near the special second-order transition point can thus be described by the $N=1$ ϕ^4 field theory and belong 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 term odd 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 describe correctly 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.

28.3 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 ideas introduced in Section 5.5. We have shown that the partition function for a Bose fluid in the grand canonical formalism can be expressed as a functional integral with the euclidean action of a non-relativistic quantum field theory.

28.3.1 The model

We have justified in Section 5.5.4 that non-relativistic Bose systems with two-body interactions can be described by an euclidean action of the form

$$\mathcal{S}(\bar\varphi,\varphi) = \int \mathrm{d}t\,\mathrm{d}^d x\,\bar\varphi(x,t)\left(\frac{\partial}{\partial t} - \frac{\hbar^2}{2m}\nabla^2 - \mu\right)\varphi(x,t) + \frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^d x\,\mathrm{d}^d y\,\bar\varphi(x,t)\varphi(x,t)V_2(x,y)\bar\varphi(y,t)\varphi(y,t), \qquad (28.48)$$

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:

$$\mathcal{S}(\bar{\varphi},\varphi)=\int \mathrm{d}t\,\mathrm{d}^d x\left[\bar{\varphi}(x,t)\left(\frac{\partial}{\partial t}-\frac{\hbar^2}{2m}\nabla^2-\mu\right)\varphi(x,t)+\frac{G}{2}\left(\bar{\varphi}(x,t)\varphi(x,t)\right)^2\right]. \quad (28.49)$$

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:

$$\varphi\mapsto\varphi\,\mathrm{e}^{i\theta},\quad \bar{\varphi}\mapsto\bar{\varphi}\,\mathrm{e}^{-i\theta},$$

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 $\mathcal{E}(\varphi)$ for constant fields is

$$\mathcal{E}(\bar{\varphi},\varphi)=-\mu\varphi\bar{\varphi}+\tfrac{1}{2}G(\varphi\bar{\varphi})^2.$$

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 $\Delta(\omega,\mathbf{k})$ in Fourier representation is

$$\Delta(\omega,\mathbf{k})=\frac{1}{i\omega+k^2/2m-\mu}\,.$$

As a consequence of the periodic boundary conditions, the field can be expanded on a basis of periodic functions, of period β:

$$\varphi(x,t)=\sum_{\nu\in\mathbb{Z}}\mathrm{e}^{2i\pi\nu t/\beta}\,\varphi_\nu(x),\quad \bar{\varphi}(x,t)=\sum_{\nu\in\mathbb{Z}}\mathrm{e}^{2i\pi\nu t/\beta}\,\bar{\varphi}_\nu(x),$$

and thus ω is quantized: $\omega_\nu=2\pi\nu/\beta$. Near the transition point $\mu=0$, the large distance behaviour of the equal-time two-point function:

$$\Delta(x,0)=\frac{1}{(2\pi)^d}\frac{1}{\beta}\sum_\nu\int\frac{\mathrm{d}^d k\;\mathrm{e}^{ikx}}{i\omega_\nu+k^2/2m} \quad (28.50)$$

is dominated by the zero-mode $\nu=0$, the contributions of other modes decreasing exponentially,

$$\Delta(x,0)\underset{|x|\to\infty}{\sim}\frac{1}{(2\pi)^d}\frac{1}{\beta}\int\frac{\mathrm{d}^d k\;\mathrm{e}^{ikx}}{k^2/2m}\propto|x|^{2-d}.$$

One finds the usual behaviour of the propagator in a massless field theory in d isotropic dimensions.

28.3.2 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 will again be obtained from 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 37 will confirm this analysis. In fact, the quantity $\sqrt{m/\beta}$ acts as an effective momentum cut-off, or the thermal wavelength:

$$\lambda = \hbar\sqrt{\beta/m}\,, \tag{28.51}$$

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 action for the zero-mode thus reads

$$\mathcal{S}(\bar\varphi_0,\varphi_0) = \beta\int \mathrm{d}^d x\left[\bar\varphi_0(x)\left(-\frac{\hbar^2}{2m}\nabla^2-\mu\right)\varphi_0(x)+\frac{G}{2}\left(\bar\varphi_0(x)\varphi_0(x)\right)^2\right], \tag{28.52}$$

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 now recognize the action of the $O(2)$ $(\phi^2)^2$ field theory,

$$\mathcal{H}(\phi) = \int\left[\tfrac{1}{2}\left(\nabla\phi(x)\right)^2+\tfrac{1}{2}r\phi^2(x)+\tfrac{1}{4!}g\Lambda^{4-d}\left(\phi^2(x)\right)^2\right]\mathrm{d}^d x\,,$$

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, we find the relations between parameters (a is the scattering length)

$$\begin{aligned} g &= 12\lambda^{-d}\beta G \propto (a/\lambda)^{d-2}\,,\\ r &= -2\mu\beta\Lambda^2. \end{aligned}$$

Note that from the derivation of the field theory we infer that the field ϕ by itself is not an 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 ϕ^4 theory. They depend on the two exponents ν or α which are related. A measure of both quantities yields two different determinations of ν.

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.

Superfluid transition and Bose–Einstein condensation. In the case of a Bose gas, the approximation of two-body interaction is very natural because the system is dilute. In the case of a fluid this is no longer the case. However, we know that additional higher powers of $\bar\varphi\varphi$ in the action are irrelevant and thus do not modify the critical behaviour. There is one exception: if the renormalized amplitude of the $(\bar\varphi\varphi)^2$ interaction is small three-body interactions which generate a $(\bar\varphi\varphi)^3$ term become important and lead to a tricritical behaviour.

By contrast in dilute Bose gases, of the kind studied in recent years in Bose–Einstein condensation experiments, three-body are totally negligible, even though two-body interactions are small. In such systems at the transition temperature correlation functions depend only on the RG invariant length scale ξ:

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)\frac{\partial}{\partial g}\right)\xi(\Lambda,g)=0\,,$$

which is thus proportional to

$$\xi\propto\Lambda^{-1}g^{1/(d-4)}\exp\left[\int_0^g \mathrm{d}g'\left(\frac{1}{\beta(g')}-\frac{1}{(d-4)g'}\right)\right].$$

We see 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 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 is still described by the IR fixed point, and similar to the superfluid transition. But for distances much larger than $1/\Lambda$ but much smaller than ξ physics is dominated by the gaussian fixed point (a UV fixed point), and, therefore, one observes a behaviour close to the Bose–Einstein condensation of the free Bose gas.

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 mean field theory are to be expected.

Classical fluids. 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. Instead 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 $\rho(x)=\bar\varphi_0(x)\varphi_0(x)$ which 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 28.2, and which eventually leads to a $N = 1$ ϕ^4 field theory in the critical domain.

Bibliographical Notes

The relation between statistical properties of polymers and the $N = 0$ $(\phi^2)^2$ field theory has been found by

P.G. de Gennes, *Phys. Lett.* 38A (1972) 339.

We have followed the derivation of

V.J. Emery, *Phys. Rev.* B11 (1975) 239; B. Duplantier, *C.R. Acad. Sci. Paris* 290B (1980) 199.

The application of RG methods to polymers is presented in

J. des Cloizeaux, *J. Physique (Paris)* 36 (1975) 281; L. Schäfer and T.A. Witten, *J. Chem. Phys.* 66 (1977) 2121; *J. Physique (Paris)* 41 (1980) 459.

For a comprehensive experimental and theoretical discussion of polymer physics see

J. des Cloizeaux and G. Jannink, *Les Polymères en Solution* (Les Editions de Physique, Paris 1987), English version *Polymers in Solution: Modelling and Structure*, (Oxford Univ. Press 1990).

The equivalence between lattice–gas and Ising models and some implications are derived in

T.D. Lee and C.N. Yang, *Phys. Rev.* 87 (1952) 410.

The derivation of the relation between classical fluids at criticality and the ϕ^4 field theory has been inspired by

F.W. Wiegel, *Phys. Rep.* 16C (1975) 57,

where one can find a review of the application of functional integral methods to classical and quantum Bose fluids. 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 and L.P. Pitaevskii, *Zh. Eksperim. i Teor. Fiz.* 34 (1958) 1240, English transl. *Soviet Phys: JETP* 34 (1958) 858.

The transition between superfluid transition and Bose–Einstein is described from the point of view of RG in

G. Baym, J.-P. Blaizot, and J. Zinn-Justin, *Euro. Phys. Lett.* 49 (2000) 150.

29 CALCULATION OF UNIVERSAL QUANTITIES

Many models with second order phase transitions have been investigated, the effective field theories identified and then various universal quantities calculated by field theoretical methods. We can of course here report only a small number of significant results. Therefore, in this chapter, we consider only the important example of the N-vector model, that is, the $O(N)$ symmetric $(\phi^2)^2$ field theory,

$$\mathcal{H}(\phi) = \int \left\{ \frac{1}{2} \left[\partial_\mu \phi(x)\right]^2 + \frac{1}{2} (r_c + t)\, \phi^2(x) + \frac{1}{4!} g \Lambda^\varepsilon \left(\phi^2(x)\right)^2 \right\} \mathrm{d}^d x, \tag{29.1}$$

in which ϕ is a N-vector.

We present results for critical exponents, the equation of state and some amplitude ratios. We discuss more thoroughly critical exponents because they allow the most detailed and accurate comparison between Field Theory, lattice models and experiments.

Universal quantities have been calculated in Field Theory by two methods: the ε-expansion, which we have systematically discussed in previous chapters, and perturbation theory at fixed dimension (see Section 26.1). 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 42. There exist methods to deal with divergent series (see Sections 42.5–42.7), all of which rely on some additional knowledge about the analytic properties of the function which has been expanded. In the case of the ϕ^4 field theory, the Borel summation has been extensively used because the information drawn from the large order behaviour analysis can easily be incorporated. In the case of the series at fixed dimension Borel summability has actually been rigorously proven. For the ε-expansion Borel summability has been assumed. Finally, to sum the series efficiently, it is necessary to make some plausible assumptions about the analytic properties of the Borel transform itself. Therefore, 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 provides an internal consistency check of field theory methods. The agreement with series estimates coming from high temperature expansions on lattices is then a verification of the concept of universality in theoretical models. Finally, of course, the ultimate test comes from the confrontation with experimental results.

The results of the N-vector model do not apply only to ferromagnetic systems. In Chapter 28, we have explained why the superfluid Helium transition corresponds to $N = 2$, how the $N = 0$ limit is related to the statistical properties of polymers and how the Ising-like $N = 1$ model describes the physics of the liquid–vapour transition.

29.1 The ε-Expansion

Many quantities have been calculated at three-loop order within the ε-expansion. Moreover, a series of very ingenious tricks and the use of symbolic manipulations on computers, have allowed the determination of critical exponents in the N-vector model, up to order ε^5. The higher order calculations have been done using dimensional regularization and the minimal subtraction scheme (see Chapter 11).

29 CALCULATION OF UNIVERSAL QUANTITIES

Many models with second order phase transitions have been investigated, the effective field theories identified and then various universal quantities calculated by field theoretical methods. We can of course here report only a small number of significant results. Therefore, in this chapter, we consider only the important example of the N-vector model, that is, the $O(N)$ symmetric $(\boldsymbol{\phi}^2)^2$ field theory,

$$\mathcal{H}(\boldsymbol{\phi}) = \int \left\{ \frac{1}{2} \left[\partial_\mu \boldsymbol{\phi}(x)\right]^2 + \frac{1}{2} \left(r_c + t\right) \boldsymbol{\phi}^2(x) + \frac{1}{4!} g \Lambda^{\varepsilon} \left(\boldsymbol{\phi}^2(x)\right)^2 \right\} \mathrm{d}^d x , \tag{29.1}$$

in which $\boldsymbol{\phi}$ is a N-vector.

We present results for critical exponents, the equation of state and some amplitude ratios. We discuss more thoroughly critical exponents because they allow the most detailed and accurate comparison between Field Theory, lattice models and experiments.

Universal quantities have been calculated in Field Theory by two methods: the ε-expansion, which we have systematically discussed in previous chapters, and perturbation theory at fixed dimension (see Section 26.1). 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 42. There exist methods to deal with divergent series (see Sections 42.5–42.7), all of which rely on some additional knowledge about the analytic properties of the function which has been expanded. In the case of the ϕ^4 field theory, the Borel summation has been extensively used because the information drawn from the large order behaviour analysis can easily be incorporated. In the case of the series at fixed dimension Borel summability has actually been rigorously proven. For the ε-expansion Borel summability has been assumed. Finally, to sum the series efficiently, it is necessary to make some plausible assumptions about the analytic properties of the Borel transform itself. Therefore, 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 provides an internal consistency check of field theory methods. The agreement with series estimates coming from high temperature expansions on lattices is then a verification of the concept of universality in theoretical models. Finally, of course, the ultimate test comes from the confrontation with experimental results.

The results of the N-vector model do not apply only to ferromagnetic systems. In Chapter 28, we have explained why the superfluid Helium transition corresponds to $N = 2$, how the $N = 0$ limit is related to the statistical properties of polymers and how the Ising-like $N = 1$ model describes the physics of the liquid–vapour transition.

29.1 The ε-Expansion

Many quantities have been calculated at three-loop order within the ε-expansion. Moreover, a series of very ingenious tricks and the use of symbolic manipulations on computers, have allowed the determination of critical exponents in the N-vector model, up to order ε^5. The higher order calculations have been done using dimensional regularization and the minimal subtraction scheme (see Chapter 11).

29.1.1 Critical exponents

Although the renormalization group (RG) functions of the $(\phi^2)^2$ theory and, therefore, the critical exponents are known up to five-loop order, we give here only three successive terms in the expansion, referring to the literature for higher order results. In terms of the variable

$$\tilde{g} = N_d\, g\,, \qquad N_d = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}\,, \tag{29.2}$$

the RG functions $\beta(\tilde{g})$ and $\eta_2(\tilde{g})$ at three-loop order, $\eta(\tilde{g})$ at four-loop order are

$$\begin{aligned}\beta(\tilde{g}) &= -\varepsilon\tilde{g} + \frac{(N+8)}{6}\tilde{g}^2 - \frac{(3N+14)}{12}\tilde{g}^3 \\ &\quad + \frac{[33N^2+922N+2960+96(5N+22)\zeta(3)]}{12^3}\tilde{g}^4 + O\left(\tilde{g}^5\right), \end{aligned}\tag{29.3}$$

$$\eta(\tilde{g}) = \frac{(N+2)}{72}\tilde{g}^2\left[1 - \frac{(N+8)}{24}\tilde{g} + \frac{5(-N^2+18N+100)}{576}\tilde{g}^2\right] + O\left(\tilde{g}^5\right), \tag{29.4}$$

$$\eta_2(\tilde{g}) = -\frac{(N+2)}{6}\tilde{g}\left[1 - \frac{5}{12}\tilde{g} + \frac{(5N+37)}{48}\tilde{g}^2\right] + O\left(\tilde{g}^4\right), \tag{29.5}$$

in which $\zeta(s)$ is the Riemann ζ-function:

$$\zeta(3) = 1.20205690315\ldots\,.$$

The zero $\tilde{g}^*(\varepsilon)$ of the β-function then is

$$\begin{aligned}\tilde{g}^*(\varepsilon) &= \frac{6\varepsilon}{(N+8)}\left[1 + \frac{3(3N+14)}{(N+8)^2}\varepsilon + \left(\frac{18(3N+14)^2}{(N+8)^4}\right.\right. \\ &\quad \left.\left. - \frac{33N^2+922N+2960+96(5N+22)\zeta(3)}{8(N+8)^3}\right)\varepsilon^2\right] + O\left(\varepsilon^4\right). \end{aligned}\tag{29.6}$$

The values of the critical exponents η, γ and ω,

$$\eta = \eta(\tilde{g}^*), \quad \gamma = \frac{2-\eta}{2+\eta_2(\tilde{g}^*)}, \quad \omega = \beta'(\tilde{g}^*),$$

follow

$$\begin{aligned}\eta &= \frac{\varepsilon^2(N+2)}{2(N+8)^2}\left\{1 + \frac{(-N^2+56N+272)}{4(N+8)^2}\varepsilon + \frac{1}{16(N+8)^4}\left[-5N^4 - 230N^3\right.\right. \\ &\quad \left.\left. + 1124N^2 + 17920N + 46144 - 384(5N+22)(N+8)\zeta(3)\right]\varepsilon^2\right\} + O\left(\varepsilon^5\right), \end{aligned}\tag{29.7}$$

$$\begin{aligned}\gamma &= 1 + \frac{(N+2)}{2(N+8)}\varepsilon + \frac{(N+2)}{4(N+8)^3}\left(N^2+22N+52\right)\varepsilon^2 + \frac{(N+2)}{8(N+8)^5} \\ &\quad \times \left[N^4 + 44N^3 + 664N^2 + 2496N + 3104 - 48(5N+22)(N+8)\zeta(3)\right]\varepsilon^3 \\ &\quad + O\left(\varepsilon^4\right), \end{aligned}\tag{29.8}$$

$$\begin{aligned}\omega &= \varepsilon - \frac{3(3N+14)}{(N+8)^2}\varepsilon^2 + \frac{[33N^2+922N+2960+96(5N+22)\zeta(3)]}{4(N+8)^3}\varepsilon^3 \\ &\quad - 18\frac{(3N+14)^2}{(N+8)^4}\varepsilon^3 + O\left(\varepsilon^4\right). \end{aligned}\tag{29.9}$$

Table 29.1

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

All other exponents can be obtained from the scaling relations derived in Chapter 26. Note that the results presented above involve $\zeta(3)$. At higher orders $\zeta(5)$ and $\zeta(7)$ successively appear. In table 29.1, 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 42 that the ε-expansion is divergent for all values of ε. Divergent series can be used for small values of the argument. However, only a finite number of terms of the series can be taken into account. The last term added gives an indication of the size of the irreducible error. For the exponents γ and η we, therefore, conclude from the series

$$\gamma = 1.244 \pm 0.050\,, \qquad \eta = 0.037 \pm 0.008\,,$$

where the errors are only indicative of the uncertainty about the result.

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

$$\beta(g) = g\bigl(-\varepsilon + s(g)\bigr),$$

then

$$g(s) = as \exp\left[\int_0^s \mathrm{d}s' \left(\frac{1}{\omega(s')} - \frac{1}{s'}\right)\right],$$

where a is an arbitrary normalization of the coupling constant.

29.1.2 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, and order ε^3 for $N = 1$. We set

$$H = M^{\delta} f(x = t/M^{1/\beta}), \tag{29.10}$$

in which the normalizations of x and the function $f(x)$ are such that

$$f(0) = 1\,, \qquad f(-1) = 0\,. \tag{29.11}$$

It is also convenient to set

$$y = x + 1\,, \qquad z = x + 3\,, \qquad \rho = z/4y\,. \tag{29.12}$$

The expansion up to order ε^2 is then

$$f(x) = 1 + x + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + O\left(\varepsilon^3\right) \tag{29.13}$$

Table 29.1

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

All other exponents can be obtained from the scaling relations derived in Chapter 26. Note that the results presented above involve $\zeta(3)$. At higher orders $\zeta(5)$ and $\zeta(7)$ successively appear. In table 29.1, 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 42 that the ε-expansion is divergent for all values of ε. Divergent series can be used for small values of the argument. However, only a finite number of terms of the series can be taken into account. The last term added gives an indication of the size of the irreducible error. For the exponents γ and η we, therefore, conclude from the series

$$\gamma = 1.244 \pm 0.050\,, \qquad \eta = 0.037 \pm 0.008\,,$$

where the errors are only indicative of the uncertainty about the result.

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

$$\beta(g) = g\big(-\varepsilon + s(g)\big),$$

then

$$g(s) = as \exp\left[\int_0^s \mathrm{d}s' \left(\frac{1}{\omega(s')} - \frac{1}{s'}\right)\right],$$

where a is an arbitrary normalization of the coupling constant.

29.1.2 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, and order ε^3 for $N = 1$. We set

$$H = M^\delta f(x = t/M^{1/\beta}), \tag{29.10}$$

in which the normalizations of x and the function $f(x)$ are such that

$$f(0) = 1\,, \qquad f(-1) = 0\,. \tag{29.11}$$

It is also convenient to set

$$y = x + 1\,, \qquad z = x + 3\,, \qquad \rho = z/4y\,. \tag{29.12}$$

The expansion up to order ε^2 is then

$$f(x) = 1 + x + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + O\left(\varepsilon^3\right) \tag{29.13}$$

with

$$f_1(x) = \frac{1}{2(N+8)}\left[(N-1)y\ln y + 3z\ln z - 9y\ln 3 + 6x\ln 2\right], \tag{29.14a}$$

$$\begin{aligned} f_2(x) = {} & \left[\frac{1}{2(N+8)}\right]^2 \{[N-1+6\ln 2 - 9\ln 3 + (N-1)\ln y]\,[3z\ln z + (N-1)y\ln y \\ & +6x\ln 2 - 9y\ln 3] + \tfrac{1}{2}(10-N)y\left(\ln^2 z - \ln^2 3\right) + 36\left(\ln^2 z - y\ln^2 3 + x\ln^2 2\right) \\ & - 54\ln 2\,(\ln z + x\ln 2 - y\ln 3) + 3\ln\frac{27}{4}(N-1)y\ln y + \frac{-4N^2+17N+212}{N+8} \\ & \times [z\ln z + 2x\ln 2 - 3y\ln 3] + (N-1)y\ln y\ln z - \tfrac{1}{2}N(N-1)y\ln^2 y \\ & + \frac{N-1}{N+8}(19N+92)y\ln y - 2(N-1)\left[(x+6)J_1(x) - 6yJ_1(0)\right] \\ & -6(N-1)\left[J_2(x) - yJ_2(0)\right] + 4(N-1)\left[J_3(x) - yJ_3(0)\right]\}, \end{aligned} \tag{29.14b}$$

where

$$J_i(x) = I_i(\rho) \tag{29.15}$$

and

$$I_1(\rho) = \int_0^\infty \frac{\mathrm{d}u\ln u}{u(1-u)}\left[(1-u/\rho)^{1/2}\,\theta(\rho-u) - 1\right], \tag{29.16}$$

$$I_2(\rho) = \rho\frac{\mathrm{d}}{\mathrm{d}\rho}I_1(\rho), \tag{29.17}$$

$$I_3(\rho) = I_1(\rho) + 2I_2(\rho). \tag{29.18}$$

The expressions (29.14) 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 26.5). It is, therefore, convenient to introduce Josephson's parametrization which leads to a representation uniform in both limits.

29.1.3 Parametric representation of the equation of state

We set

$$x = x_0\left(1-\theta^2\right)\theta^{-1/\beta}, \quad \theta > 0, \tag{29.19}$$

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

$$\begin{aligned} M &= R^\beta\theta, \\ t &= x_0 R\left(1-\theta^2\right), \\ H &= R^{\beta\delta}h(\theta). \end{aligned} \tag{29.20}$$

Then, the function

$$h(\theta) = \theta^\delta f\left(x(\theta)\right) \tag{29.21}$$

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

$$x_0 = 3\,(3/2)^{1/2\beta - 1}, \tag{29.22}$$

the equation of state at order ε takes a rather simple form

$$h(\theta) = \theta\left(3 - 2\theta^2\right)\left[1 + \frac{\varepsilon(N-1)}{2(N+8)}\ln\left(3 - 2\theta^2\right)\right] + O\left(\varepsilon^2\right). \tag{29.23}$$

For $N = 1$, the equation is specially 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

$$h(\theta) = h_0\theta\left(b^2 - \theta^2\right)\left(1 + c\theta^2\right) + O\left(\varepsilon^4\right), \tag{29.24}$$

in which h_0 is the field normalization constant, and b, c are given by

$$b^2 = \frac{3}{2}\left(1 - \frac{\varepsilon^2}{12}\right), \qquad c = -\frac{\varepsilon^3}{18}\left(\zeta(3) + \frac{I-1}{4}\right) \tag{29.25}$$

with

$$I = \int_0^1 \mathrm{d}x \frac{\ln\left[x(1-x)\right]}{1 - x(1-x)} = \tfrac{4}{9}\pi^2 - \tfrac{2}{3}\psi'(1/3) = -2.3439072386\ldots . \tag{29.26}$$

The constant x_0 is given by

$$x_0 = b^{1/\beta}/\left(b^2 - 1\right). \tag{29.27}$$

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 presented in Chapter 31. We will 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 more details see Sections 31.2.2,31.3).

In all cases, as stated above, 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 theory prediction.

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 per unit volume. Setting

$$F(M,t) \equiv \Omega^{-1}\Gamma_{\text{sg.}}(M,t) = R^{2-\alpha}g(\theta), \tag{29.28}$$

one finds for $g(\theta)$ a differential equation

$$h(\theta)\left(1 - \theta^2 + 2\beta\theta^2\right) = 2(2-\alpha)\theta g(\theta) + \left(1 - \theta^2\right)g'(\theta). \tag{29.29}$$

The integration constant is obtained by requiring the regularity of $g(\theta)$ at $\theta = 1$. Note that if one expands

$$h(\theta)\left(1 - \theta^2 + 2\beta\theta^2\right) = X_0 + X_1(1-\theta^2) + X_2(1-\theta^2)^2 + O\left((1-\theta^2)^3\right),$$

then for $\alpha \to 0$

$$g(\theta) \sim -\frac{X_2}{2\alpha}(1-\theta^2)^2. \tag{29.30}$$

In the same way, the inverse magnetic susceptibility is given by

$$\chi^{-1} = R^{\gamma} g_2(\theta) \tag{29.31}$$

with

$$g_2(\theta)\left(1-\theta^2+2\beta\theta^2\right) = 2\beta\delta\theta h(\theta) + \left(1-\theta^2\right)h'(\theta). \tag{29.32}$$

These expressions can in particular be used to calculate various universal ratios of amplitudes.

29.1.4 Amplitude ratios

Apart from critical exponents some other simple universal numbers have been calculated: ratios of amplitudes of singularities near T_c. 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 at T_c like

$$C_H = A^{\pm}\left|t\right|^{-\alpha}, \qquad t = T - T_c \to \pm 0. \tag{29.33}$$

The ratio A^+/A^- then is universal. It is directly related to the function $g(\theta)$ defined by equation (29.29):

$$\frac{A^+}{A^-} = \left(b^2-1\right)^{2-\alpha}\frac{g(0)}{g(b)}. \tag{29.34}$$

At order ε^2 one finds

$$\begin{aligned}\frac{A^+}{A^-} = 2^{\alpha-2}N\Big\{1+\varepsilon+\big[3N^2+26N+100+(4-N)(N-1)\zeta(2)\\ -6(5N+22)\zeta(3)-9(4-N)\lambda\big]\frac{\varepsilon^2}{2(N+8)^2}\Big\}+O\left(\varepsilon^3\right)\end{aligned} \tag{29.35}$$

with

$$\zeta(2) = \pi^2/6 = 1.64493406684\ldots,$$

while λ is defined in terms of the integral I given in (29.26):

$$\lambda = -I/2 = \tfrac{1}{3}\psi'(1/3) - \tfrac{2}{9}\pi^2 = 1.17195361934\ldots.$$

The evaluation (29.30) 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)

$$\begin{aligned}\frac{A^+}{A^-} &= 2^{\alpha}(1-K\alpha/\varepsilon),\\ K &= \frac{1}{2}(N+8) + \frac{N^2+4N+28}{2(N+8)}\varepsilon + O(\varepsilon^2).\end{aligned}$$

The magnetic susceptibility. The magnetic susceptibility in zero field can also be calculated from the function $g_2(\theta)$ defined by equation (29.32). As we know, below T_c, the susceptibility diverges for systems with Goldstone modes. We restrict ourselves, therefore, to $N = 1$. Defining

$$\chi = C^{\pm} |t|^{-\gamma} , \qquad t \to \pm 0 , \tag{29.36}$$

one obtains

$$\frac{C^+}{C^-} = \frac{2\left(1+cb^2\right)\left(b^2-1\right)^{1-\gamma}}{\left[1-b^2\left(1-2\beta\right)\right]} \tag{29.37a}$$

$$= \frac{2^{\gamma+1}}{6\beta-1}\left[1+\left(\frac{2\lambda+1}{4}-\zeta(3)\right)\frac{\varepsilon^3}{12}\right]+O\left(\varepsilon^4\right). \tag{29.37b}$$

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. Let us here define the correlation length in terms of the ratio of the two first moments of the two-point correlation function:

$$\Gamma^{(2)}(p) = \Gamma^{(2)}(0)\left(1+\xi_1^2 p^2\right)+O\left(p^4\right). \tag{29.38}$$

The function ξ_1^2 has the scaling form

$$\xi_1^2(M,t) = M^{-2\nu/\beta} f_\xi(t/M^{1/\beta}). \tag{29.39}$$

It shares otherwise all the properties of the equation of state. It can be written in parametric form as

$$\xi_1^2(M,t) = R^{-2\nu} g_\xi(\theta). \tag{29.40}$$

At order ε for $N = 1$, for example, one finds

$$g_\xi(\theta) = g_\xi(0)\left(1-\frac{5}{18}\varepsilon\theta^2\right)+O\left(\varepsilon^2\right). \tag{29.41}$$

Setting in zero field

$$\xi_1 = f_1^{\pm}|t|^{-\nu}, \qquad t \to \pm 0 , \tag{29.42}$$

quantity which exists only for $N = 1$, one can use the determination of g_ξ to calculate this ratio:

$$f_1^+/f_1^- = 2^{\nu}\left[1+\frac{5}{24}\varepsilon+\frac{1}{432}\left(\frac{295}{24}+2I\right)\varepsilon^2\right]+O\left(\varepsilon^3\right), \tag{29.43}$$

in which the constant I is given by equation (29.26).

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,

$$R_\xi^+ = f_1^+\left(\alpha A^+\right)^{1/d}. \tag{29.44}$$

Indeed, from the definitions (29.28,29.42),

$$\left(R_\xi^+\right)^d = (1-\alpha)(2-\alpha)t^{\alpha-2}F(0,t)t^{\nu d}(\xi_1)^d = (1-\alpha)(2-\alpha)F(0,t)(\xi_1)^d,$$

and the last product is normalization independent. The ε-expansion of R_ξ^+ is

$$\left(R_\xi^+\right)^d = \sigma_d \frac{N}{2}\nu(1-\alpha)\left[1+\eta\left(\frac{-11}{2}+\frac{14}{3}\lambda\right)\right]+O\left(\varepsilon^3\right) \tag{29.45}$$

with

$$\sigma_d = \Gamma(1+\varepsilon/2)\Gamma(1-\varepsilon/2)N_d$$

(we recall the definition of the loop factor N_d in (29.2)).

Other universal ratios. It is of course 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

$$\xi_1 = f_1^c/H^{2/(d+2-\eta)}, \tag{29.46}$$

the magnetic susceptibility as

$$\chi = C^c/H^{1-1/\delta}; \tag{29.47}$$

the spontaneous magnetization vanishes as

$$M = B\,(-t)^\beta\,, \tag{29.48}$$

and the spin–spin correlation function in momentum space at T_c behaves as

$$\chi(p) = \left[\Gamma^{(2)}(p)\right]^{-1} = Dp^{\eta-2}. \tag{29.49}$$

One can then define the following universal ratio:

$$R_c = \alpha A^+ C^+/B^2, \tag{29.50}$$

which corresponds to the relation between exponents

$$\alpha + 2\beta + \gamma = 2\,.$$

Indeed, using this relation, we verify that R_c is proportional to $F(0,t)M^{-2}\chi$ which is normalization independent. The ε-expansion of R_c is

$$R_c = \frac{N}{N+8}2^{-2\beta-1}\varepsilon\left[1+\left(1-\frac{30}{(N+8)^2}\right)\varepsilon\right]+O\left(\varepsilon^3\right). \tag{29.51}$$

Following Fisher and Tarko, one can construct the three following universal combinations:

$$Q_1 = C^c\delta/\left(B^{\delta-1}C^+\right)^{1/\delta}, \tag{29.52}$$

$$Q_2 = \left(f_1^c/f_1^+\right)^{2-\eta}C^+/C^c, \tag{29.53}$$

$$Q_3 = D\left(f_1^+\right)^{2-\eta}/C^+, \tag{29.54}$$

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 by Aharony and Hohenberg:

$$R_\chi = Q_1^{-\delta}.$$

Their ε-expansions are

$$R_\chi = 3^{(\delta-3)/2} 2^{\gamma+(1-\delta)/2} \left[1 + \left(\frac{2\lambda+1}{4} - \zeta(3)\right)\frac{\varepsilon^3}{18}\right] + O\left(\varepsilon^4\right), \tag{29.55}$$

$$Q_2 = 1 + \frac{\varepsilon}{18} + \left(\frac{23}{9} + \frac{4}{3}\lambda\right)\frac{\varepsilon^2}{54} + O\left(\varepsilon^3\right), \tag{29.56}$$

$$Q_3 = 1 - \left(\frac{8}{3}\lambda + 5\right)\frac{\varepsilon^2}{216} + O\left(\varepsilon^3\right). \tag{29.57}$$

Numerical results are given in table 29.10 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. Let us write any physical quantity for $t = T - T_c$ small as

$$f(t) = A_f\,|t|^{-\lambda_f}\left(1 + a_f\,|t|^{\theta} + \cdots\right), \tag{29.58}$$

where the correction exponent θ (also called Δ_1) is given by

$$\theta = \omega\nu\,. \tag{29.59}$$

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. Let us give one example corresponding to the correlation length and the susceptibility above T_c:

$$\frac{a_\chi^+}{a_\xi^+} = 2\left\{1 - \frac{\varepsilon}{N+8} - \left[\frac{2\lambda}{3(N+8)} - \frac{N^2 - 15N - 124}{2(N+8)^3}\right]\varepsilon^2\right\} + O\left(\varepsilon^3\right). \tag{29.60}$$

29.2 The Perturbative Expansion at Fixed Dimension

Critical exponents and various universal quantities have also been calculated within the framework defined in Section 26.1, that is, in the massive $(\phi^2)^2$ field theory, as perturbative series in fixed dimension. For example, the RG β-function in three dimensions, for $N = 1$, has the expansion

$$\begin{aligned}\beta(\tilde g) = {} & -\tilde g + \tilde g^2 - \tfrac{308}{729}\tilde g^3 + 0.3510695978\tilde g^4 - 0.3765268283\tilde g^5 \\ & + 0.49554751\tilde g^6 - 0.749689\tilde g^7 + O\left(\tilde g^8\right)\end{aligned} \tag{29.61}$$

with the normalization

$$\tilde g = 3g/(16\pi). \tag{29.62}$$

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 by Aharony and Hohenberg:

$$R_\chi = Q_1^{-\delta}.$$

Their ε-expansions are

$$R_\chi = 3^{(\delta-3)/2} 2^{\gamma+(1-\delta)/2} \left[1 + \left(\frac{2\lambda+1}{4} - \zeta(3)\right)\frac{\varepsilon^3}{18}\right] + O\left(\varepsilon^4\right), \tag{29.55}$$

$$Q_2 = 1 + \frac{\varepsilon}{18} + \left(\frac{23}{9} + \frac{4}{3}\lambda\right)\frac{\varepsilon^2}{54} + O\left(\varepsilon^3\right), \tag{29.56}$$

$$Q_3 = 1 - \left(\frac{8}{3}\lambda + 5\right)\frac{\varepsilon^2}{216} + O\left(\varepsilon^3\right). \tag{29.57}$$

Numerical results are given in table 29.10 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. Let us write any physical quantity for $t = T - T_c$ small as

$$f(t) = A_f \left|t\right|^{-\lambda_f} \left(1 + a_f \left|t\right|^{\theta} + \cdots\right), \tag{29.58}$$

where the correction exponent θ (also called Δ_1) is given by

$$\theta = \omega\nu\,. \tag{29.59}$$

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. Let us give one example corresponding to the correlation length and the susceptibility above T_c:

$$\frac{a_\chi^+}{a_\xi^+} = 2\left\{1 - \frac{\varepsilon}{N+8} - \left[\frac{2\lambda}{3(N+8)} - \frac{N^2 - 15N - 124}{2(N+8)^3}\right]\varepsilon^2\right\} + O\left(\varepsilon^3\right). \tag{29.60}$$

29.2 The Perturbative Expansion at Fixed Dimension

Critical exponents and various universal quantities have also been calculated within the framework defined in Section 26.1, that is, in the massive $(\phi^2)^2$ field theory, as perturbative series in fixed dimension. For example, the RG β-function in three dimensions, for $N = 1$, has the expansion

$$\begin{aligned}\beta(\tilde g) = -\tilde g + \tilde g^2 - \tfrac{308}{729}\tilde g^3 + 0.3510695978\tilde g^4 - 0.3765268283\tilde g^5 \\ + 0.49554751\tilde g^6 - 0.749689\tilde g^7 + O\left(\tilde g^8\right)\end{aligned} \tag{29.61}$$

with the normalization

$$\tilde g = 3g/(16\pi). \tag{29.62}$$

To calculate exponents or other universal quantities, we have first to find the IR stable zero g^* of the function $\beta(g)$ which is given by a few terms of a divergent expansion. A first problem is the absence of a small parameter in which to expand: g^* is a number 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 the RG functions of the N-vector model up to six- and partially seven-loop order. Estimates of critical exponents are displayed in table 29.4. Universal ratios of amplitude have also been calculated, as well as the equation of state for $N = 1$. Numerical results will be presented in Section 29.7.

Note, however, that in this framework, 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. Also, the normalization of correlation functions is singular at T_c (equation (26.17)).

Let us discuss the example of Ising-like systems ($N = 1$). The free energy F has the 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 (26.15):

$$\tilde{M} = Mm^{-\eta/2}.$$

The derivative with respect to M of F yields the equation of state under the form

$$H = g^{-1/2}m^{1+d/2-\eta/2}\mathfrak{h}(g^{1/2}m^{1-d/2}\tilde{M}, g), \quad \mathfrak{h}(z,g) = \varphi'_z(z,g). \tag{29.63}$$

At one-loop order, the function $\mathfrak{h}(z,g)$ is given by

$$\begin{aligned}\mathfrak{h}(z,g) &= z + \frac{z^3}{6} + \frac{gz}{2}\frac{1}{(2\pi)^d}\int\frac{\mathrm{d}^d p}{p^2+1+z^2/2} + O\,(2\text{ loops})\\ &= z + \frac{z^3}{6} + \frac{\pi N_d}{4\sin\pi d/2}gz\left[(1+z^2/2)^{d/2-1} - 1 - \frac{1}{4}(d-2)z^2\right].\end{aligned} \tag{29.64}$$

In terms of the deviation from the critical temperature $t = m^{1/\nu}(g^*)^{1/2\beta} \sim T - T_c$, equation (29.63) takes the form

$$H = H_0 t^{\beta\delta}\mathfrak{h}(Mt^{-\beta}) \quad (H_0 \text{ being a constant})\,. \tag{29.65}$$

This expression is adequate for the description of the disordered phase when $Mt^{-\beta}$ is small but all terms in the loop expansion become singular when t goes to zero.

The ordered phase. This 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

$$z = Mt^{-\beta}, \tag{29.66}$$

picks up a phase below T_c:

$$\text{for } t = |t|\,\mathrm{e}^{i\pi}, \quad z = |z|\,\mathrm{e}^{-i\pi\beta}. \tag{29.67}$$

The scaling variable $H(-t)^{-\beta\delta}$then is given by

$$H(-t)^{-\beta\delta} = H_0\,\mathrm{e}^{i\pi\beta\delta}\,\mathfrak{h}(z) = H_0\,|\mathfrak{h}(z)|. \tag{29.68}$$

It is in particular possible to evaluate ratios of amplitudes of singularities above and below T_c: we 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. Let us take the example of the magnetic susceptibility for illustration purpose. From equation (29.63), we derive

$$C^+/C^- = \mathrm{e}^{-i\pi\gamma}\,\mathfrak{h}'(z_0(g^*),g^*)/\mathfrak{h}'(0,g^*) = \left|\mathfrak{h}'(z_0(g^*),g^*)\right|, \tag{29.69}$$

in which z_0 is the complex zero of $\mathfrak{h}(z,g)$. We thus get a series expansion for C^+/C^-.

However, this method does not allow to extrapolate the equation of state to t small. Following the lines of Section 29.1.3, it is natural to introduce the parametric representation (29.20),

$$z = x_0^{-\beta}\theta\left(1-\theta^2\right)^{-\beta},$$

and consider the function

$$h(\theta) = \left(1-\theta^2\right)^{\beta\delta}\mathfrak{h}(z(\theta)).$$

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 anymore. 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 29.3, 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, as in the ODM method explained in Section 42.7. 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. Corresponding results for the amplitude ratios are reported in table 29.10. It would be interesting to apply this method to $N \neq 1$ series.

29.3 The Series Summation

The principles and the theoretical justification of the summation method based on Borel summation and conformal mapping are explained in Section 42.7. We add here only some details about the specific implementation used in the case of the calculation of critical exponents and other universal quantities. A few examples of transformed series are displayed in table 29.2 to illustrate the convergence.

The method. Several different variants based on the Borel–Leroy transformation 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):

$$R(z) = \sum_{k=0} R_k z^k. \tag{29.70}$$

Table 29.2

Series summed by the method based on Borel transformation and mapping for the zero $\tilde{g}^*$ of the $\beta(g)$ function and the exponents γ and ν in the ϕ^4_3 field theory.

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	0.62966	0.6302	0.6302	0.6302
γ	1.2257	1.2370	1.2386	1.2398	1.2398	1.2398

One transforms the series into

$$R(z) = \sum_{k=0}^{\infty} B_k(\rho) \int_0^{\infty} t^{\rho}\,\mathrm{e}^{-t}\,[u(zt)]^k\,\mathrm{d}t\,, \tag{29.71}$$

$$u(z) = \frac{\sqrt{1+az}-1}{\sqrt{1+za}+1}\,. \tag{29.72}$$

The coefficients B_k are calculated by identifying the expansion of the r.h.s. of equation (29.71) in powers of z with the expansion (29.70). The constant a has been determined by the large order behaviour analysis,

$$a(d=3) = 0.147774232 \times \left(9/(N+8)\right), \qquad a(d=2, N=1) = 0.238659217\,, \tag{29.73}$$

and ρ is a free parameter, 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, a conformal transformation has been made on the initial function $R(z)$ in order to send away its closest singularities, and the procedure described above applied instead to the function

$$\tilde{R}(z) = R\left[z/(1-\tau z)\right], \tag{29.74}$$

in which τ is also left as an adjustable parameter 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 series in 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. It is, therefore, essential to vary the parameters in some range around the optimal values 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 errors quoted in the final results should be considered as somewhat indicative.

The $(\phi^2)^2$ field theory at fixed dimensions. The RG β-function has been determined up to six-loop order in three dimensions, while the series for the dimensions of the fields ϕ and ϕ^2 have recently been extended to seven loops. In two dimensions the series are known only 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 $\tilde{g}=\tilde{g}^*$. Examples of convergence are given in table 29.2.

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

$$\beta(\tilde{g}, \varepsilon) = \tilde{g}(1 - \varepsilon) + \beta(\tilde{g}). \tag{29.75}$$

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 mapping $\tilde{g} \mapsto \varepsilon$ introduces singularities because the apparent convergence is poorer. It is, therefore, gratifying that all variants give consistent results. A comparison of all the results is helpful for the determination of the apparent errors.

The ε-expansion. The ε-expansion has one serious advantage: it allows us to connect the results in three and two dimensions. 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 improve the three-dimensional results by imposing the exact two-dimensional values or the behaviour near two dimensions for $N > 1$. However, since the series in ε are shorter than the series at fixed dimension 3, the apparent errors are larger. Finally,, as already emphasized, the comparison between the different results is a check of the consistency of Field Theory methods combined with the summation procedures.

29.4 Numerical Estimates of Critical Exponents

Fixed dimension. We give in table 29.3, the results obtained from summed perturbation theory at fixed dimension 2 for $N = 1$ and compare them with the exact values of the Ising model. The apparent errors are large because the series are short. The agreement with the Ising model is satisfactory. Note that $\tilde{g}^*$ is known only from HT series; only analytic corrections to scaling have been found in the Ising model which makes the identification of 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 29.3

Estimates of critical exponents in the ϕ^4_2 field theory.

	$\tilde{g}^*$	ω	γ	ν	η
ϕ^4_2	1.85 ± 0.10	1.3 ± 0.2	1.79 ± 0.04	0.96 ± 0.04	0.18 ± 0.04
Ising	1.751 ± 0.005	$4/3$?	1.75	1.	0.25

Table 29.4 displays the results obtained from summed perturbation series at fixed dimension 3. The last exponent $\theta = \omega\nu$ characterizes corrections to scaling in the temperature variable (equation (29.59)).

The ε-expansion. In table 29.5, we give the results coming from the summed ε-expansion for $\varepsilon = 2$ and compare them with exact results.

Table 29.4

Estimates of critical exponents in the $O(N)$ symmetric $(\phi^2)^2_3$ 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

Table 29.5

Critical exponents in the ϕ^4_2 field theory from the ε-expansion.

	γ	ν	η	β	ω
$N = 0$	1.39 ± 0.04	0.76 ± 0.03	0.21 ± 0.05	0.065 ± 0.015	1.7 ± 0.2
Exact	1.34375	0.75	$0.2083\cdots$	$0.0781\cdots$	?
$N = 1$	1.73 ± 0.06	0.99 ± 0.04	0.26 ± 0.05	0.120 ± 0.015	1.6 ± 0.2
Ising	1.75	1.	0.25	0.125	$1.33\ldots$?

We see in this table that the agreement for $N = 0$ and $N = 1$ between field theory and lattice models is satisfactory. We feel justified, therefore, in using a summation procedure of the ε-expansion which automatically incorporates the $d = 2$, $\varepsilon = 2$ values. Note, however, that in both cases, the identification of ω remains a problem.

Table 29.6 then displays the results for $\varepsilon = 1$, both for the ε series (free) and a modified ε series where the $d = 2$ results are imposed (bc).

Discussion. We can now compare the two sets of results coming from the perturbation series at fixed dimension, and the ε-expansion. First let us emphasize that the agreement is quite spectacular, although the apparent errors of the ε-expansion are in general larger because the series are shorter. Moreover, the agreement has improved with longer series.

The best agreement is found for the exponents ν and β. On the other hand, the values of η coming from the ε-expansion are systematically larger by about 3×10^{-3}, though the error bars always overlap. The corresponding effect is observed on γ. We notice in tables 29.3, 29.4 that a similar remark applies at $d = 2$: the result at fixed dimension, $N = 1$, for η is smaller than the result coming from the ε-expansion, however, the latter result is closer to the Ising value, even when one takes into account the relative errors.

Table 29.6

Critical exponents in the $(\phi^2)^2_3$ field theory from the ε-expansion.

N	0	1	2	3
γ (free)	1.1575 ± 0.0060	1.2355 ± 0.0050	1.3110 ± 0.0070	1.3820 ± 0.0090
γ (bc)	1.1571 ± 0.0030	1.2380 ± 0.0050	1.317	1.392
ν (free)	0.5875 ± 0.0025	0.6290 ± 0.0025	0.6680 ± 0.0035	0.7045 ± 0.0055
ν (bc)	0.5878 ± 0.0011	0.6305 ± 0.0025	0.671	0.708
η (free)	0.0300 ± 0.0050	0.0360 ± 0.0050	0.0380 ± 0.0050	0.0375 ± 0.0045
η (bc)	0.0315 ± 0.0035	0.0365 ± 0.0050	0.0370	0.0355
β (free)	0.3025 ± 0.0025	0.3257 ± 0.0025	0.3465 ± 0.0035	0.3655 ± 0.0035
β (bc)	0.3032 ± 0.0014	0.3265 ± 0.0015		
ω	0.828 ± 0.023	0.814 ± 0.018	0.802 ± 0.018	0.794 ± 0.018
θ	0.486 ± 0.016	0.512 ± 0.013	0.536 ± 0.015	0.559 ± 0.017

29.5 Comparison with Lattice Model Estimates

The N-vector with nearest-neighbour interactions has been studied on various lattices. Most of the results for critical exponents come from the analysis of HT series expansion by different types of ratio methods, Padé approximants or differential approximants (see Sections 42.7 and A42.3). Some results have also been obtained from low temperature expansions, computer calculations using stochastic methods, and in low dimensions, transfer matrix methods. Table 29.7 tries to give an idea of the agreement between lattice and Field Theory results. A historical remark is here in order: the agreement between both types of theoretical results has improved as the HT series became longer which is of course encouraging. The main reason is that, in the analysis of longer series, it has become possible to take into account the influence of confluent singularities due to corrections to the leading power law behaviour, as predicted by the RG. The effect of this improvement has been specially spectacular for the exponents γ and ν of the 3D Ising model: the longer series obtained by Nickel on the BCC lattice did almost completely eliminate the disturbing small differences which had remained between HT series and RG results.

Table 29.7

Critical exponents in the N-vector model on the lattice.

N	0	1	2	3
γ	1.1575 ± 0.0006	1.2385 ± 0.0025	1.322 ± 0.005	1.400 ± 0.006
ν	0.5877 ± 0.0006	0.631 ± 0.002	0.674 ± 0.003	0.710 ± 0.006
α	0.237 ± 0.002	0.103 ± 0.005	-0.022 ± 0.009	-0.133 ± 0.018
β	0.3028 ± 0.0012	0.329 ± 0.009	0.350 ± 0.007	0.365 ± 0.012
θ	0.56 ± 0.03	0.53 ± 0.04	0.60 ± 0.08	0.54 ± 0.10

We have not listed all available results but rather only typical numbers in order to give a feeling of the consistency between RG and lattice estimates (for $N \neq 1$ α, β are obtained

Table 29.8

Critical exponents in fluids and antiferromagnets.

	γ	ν	β	α	θ
(a)	1.236 ± 0.008	0.625 ± 0.010	0.325 ± 0.005	0.112 ± 0.005	0.50 ± 0.03
(b)	1.23– 1.25	0.625 ± 0.006	0.316– 0.327	0.107 ± 0.006	0.50 ± 0.03
(c)	1.25 ± 0.01	0.64 ± 0.01	0.328 ± 0.009	0.112 ± 0.007	

by scaling). The obvious conclusion is that one observes no systematic differences. In particular, the agreement is extremely good in the case of the Ising model where the HT series are the most accurate. To the best of our knowledge, the N-vector lattice models and the $(\phi^2)^2$ field theory belong to the same universality class.

29.6 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 non-magnetic, but belong to the universality class of the N-vector model. The case $N = 0$ describes the statistical properties of long polymers, that is, long not intersecting chains or self-avoiding walks (see Section 28.1). The case $N = 1$ (Ising-like systems) describes liquid–vapour transitions in classical fluids, critical binary fluids and uniaxial ntiferromagnets. The helium superfluid transition corresponds to $N = 2$. Finally, for $N = 3$, the experimental information comes from ferromagnetic systems.

Critical exponents and polymers. In the case of polymers, only the exponent ν is easily accessible. The best results are

$$\nu = 0.586 \pm 0.004 ,$$

in excellent agreement with the RG result.

Ising-like systems $N = 1$. Table 29.8 gives a survey of the experimental situation for critical binary fluids (a), liquid–vapour transition in classical fluids (b), and antiferromagnets (c). For binary mixtures, we quote a weighted world average. In the case of the liquid–vapour transition, we quote a range of experimental results rather than statistical errors for all exponents but ν, the reason being that the values depend much on the method of analysis of the experimental data. The agreement with RG results is clearly impressive.

Helium superfluid transition, $N = 2$. The helium transition allows extremely precise measurements very close to T_c and this explains the accuracy of the determination of critical exponents. The order parameter, however, is not directly accessible and, therefore, only ν and α have been determined. Most recent reported values are

$$\nu = 0.6705 \pm 0.0006 , \quad \nu = 0.6708 \pm 0.0004 \quad \text{and} \quad \alpha = -0.01285 \pm 0.00038 .$$

The agreement with RG values is quite remarkable but the precision of ν becomes a challenge to field theory.

Table 29.9

Ferromagnetic systems.

γ	ν	β	α	θ
1.40 ± 0.03	0.700–0.725	0.35 ± 0.03	-0.09– -0.012	0.54 ± 0.10

Ferromagnetic systems, $N = 3$. Finally, table 29.9 displays some results concerning magnetic systems.

29.7 Amplitude Ratios

Table 29.10 contains a comparison of amplitude ratios as obtained from renormalization group for $N = 1$, lattice calculations for Ising-like models and experiments on binary mixtures. RG values for amplitudes are less accurate than for exponents because the series are shorter. Note that a RG determination of the equation of state is also available.

Table 29.10

Amplitude ratios: models and binary critical fluids.

	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	

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 shows, with indeed large errors, a satisfactory agreement with RG predictions.

Finally, let us 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, we consider also the specific heat amplitude a_C and the coexistence curve magnetization amplitude a_M, we can form three independent ratios. The results for $N = 1$ are given in table 29.11.

Table 29.11

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

Table 29.12

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 29.12 the examples of A^+/A^-, R_ξ^+ and R_c.

If one takes into account all data (critical exponents, amplitude ratios,...) one is forced to conclude that the RG predictions are remarkably consistent with the whole experimental information available. Considering the variety of experimental situations, this is a spectacular confirmation of the RG ideas and the concept of universality.

Bibliographical Notes

The calculation of exponents as an ε-expansion were initiated by

K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* 28 (1972) 240; K.G. Wilson, *Phys. Rev. Lett.* 28 (1972) 548.

The results at next order ε^3 obtained by QFT techniques were reported in

E. Brézin, J.C. Le Guillou, J. Zinn-Justin and B.G. Nickel, *Phys. Lett.* 44A (1973) 227.

Exponents at order ε^5 have been obtained by

A.A. Vladimirov, D.I. Kazakov and O.V. Tarasov, *Zh. Eksp. Teor. Fiz.* 77 (1979) 1035 (*Sov. Phys. JETP* 50 (1979) 521); K.G. Chetyrkin, A.L. Kataev and F.V. Tkachov, *Phys. Lett.* B99 (1981) 147; B101 (1981) 457(E); K.G. Chetyrkin and F.V. Tkachov, *Nucl. Phys.* B192 (1981) 159; K.G. Chetyrkin, S.G. Gorishny, S.A. Larin and F.V. Tkachov, *Phys. Lett.* 132B (1983) 351; D.I. Kazakov, *Phys. Lett.* 133B (1983) 406; S.G. Gorishny, S.A. Larin and F.V. Tkachov, *Phys. Lett.* 101A (1984) 120; H. Kleinert, J.

Table 29.11

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

Table 29.12

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 29.12 the examples of A^+/A^-, R_ξ^+ and R_c.

If one takes into account all data (critical exponents, amplitude ratios,...) one is forced to conclude that the RG predictions are remarkably consistent with the whole experimental information available. Considering the variety of experimental situations, this is a spectacular confirmation of the RG ideas and the concept of universality.

Bibliographical Notes

The calculation of exponents as an ε-expansion were initiated by

K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* 28 (1972) 240; K.G. Wilson, *Phys. Rev. Lett.* 28 (1972) 548.

The results at next order ε^3 obtained by QFT techniques were reported in

E. Brézin, J.C. Le Guillou, J. Zinn-Justin and B.G. Nickel, *Phys. Lett.* 44A (1973) 227.

Exponents at order ε^5 have been obtained by

A.A. Vladimirov, D.I. Kazakov and O.V. Tarasov, *Zh. Eksp. Teor. Fiz.* 77 (1979) 1035 (*Sov. Phys. JETP* 50 (1979) 521); K.G. Chetyrkin, A.L. Kataev and F.V. Tkachov, *Phys. Lett.* B99 (1981) 147; B101 (1981) 457(E); K.G. Chetyrkin and F.V. Tkachov, *Nucl. Phys.* B192 (1981) 159; K.G. Chetyrkin, S.G. Gorishny, S.A. Larin and F.V. Tkachov, *Phys. Lett.* 132B (1983) 351; D.I. Kazakov, *Phys. Lett.* 133B (1983) 406; S.G. Gorishny, S.A. Larin and F.V. Tkachov, *Phys. Lett.* 101A (1984) 120; H. Kleinert, J.

Neu, V. Schulte-Frohlinde, K.G. Chetyrkin and S.A. Larin, *Phys. Lett.* B272 (1991) 39, Erratum B319 (1993) 545.

The Borel summation of Section 29.3 has been first applied on the ε-expansion in

J.C. Le Guillou and J. Zinn-Justin, *J. Physique Lett. (Paris)* 46 (1985) L137; *J. Physique (Paris)* 48 (1987) 19; *ibidem* 50 (1989) 1365.

The use of perturbation series at fixed dimension has been advocated by

G. Parisi, *Cargèse Lectures 1973*, published in *J. Stat. Phys.* 23 (1980) 49.

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 and D.I. Meiron, *Phys. Rev. Lett.* 36 (1976) 1351; B.G. Nickel, D.I. Meiron, G.B. Baker, *Univ. of Guelph Report* 1977.

First precise estimates based on the method of Section 29.3 were published in

J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. Lett.* 39 (1977) 95; *Phys. Rev.* B21 (1980) 3976.

The updated 3D results presented in Section 29.4 are taken from

R. Guida and J. Zinn-Justin, *J. Phys. A* 31 (1998) 8103, cond-mat/9803240,

and obtained from corrected ε series, and seven-loop 3D terms for γ and η reported in

D.B. Murray and B.G. Nickel, unpublished Guelph University report (1991).

Six-loop series have been published for generic values of N

S.A. Antonenko and A.I. Sokolov, *Phys. Rev.* E51 (1995) 1894.

The scaling equation of state is known to order ε^2 for all N and order ε^3 for $N = 1$

G.M. Avdeeva and A.A. Migdal, *JETP Lett.* 16 (1972) 178; E. Brézin, D.J. Wallace and K.G. Wilson, *Phys. Rev. Lett.* 29 (1972) 591; *Phys. Rev.* B7 (1973) 232, D.J. Wallace and R.P.K. Zia, *J. Phys. C* 7 (1974) 3480.

For an early review see

D.J. Wallace, in *Phase Transitions and Critical Phenomena* vol. 6, (references page x).

The parametric representation has been introduced in

P. Schofield, J.D. Litster and J.T. Ho, *Phys. Rev. Lett.* 23 (1969) 1098; B.D. Josephson, *J. Phys. C* 2 (1969) 1113.

Calculations of ratios of amplitudes have been reported in

E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Lett.* 47A (1974) 285; H.B. Tarko and M.E. Fisher, *Phys. Rev. Lett.* 31 (926) 1973; *Phys. Rev.* B11 (1975) 1217; M.E. Fisher and H.B. Tarko, *Phys. Rev.* B11 (1975) 1131; A. Aharony and P.C. Hohenberg, *Phys. Rev.* B13 (1976) 3081; *Physica* 86-88B (1977) 611; Y. Okabe and K. Ideura, *Prog. Theor. Phys.* 66 (1981) 1959; J.F. Nicoll and P.C. Albright, *Phys. Rev.* B31 (1985) 4576.

Series and estimates for equation of state and ratios of amplitude are reported in

C. Bagnuls and C. Bervillier, *Phys. Rev.* B24 (1981) 1226; *J. Phys. A* 19 (1986) L85; C. Bervillier, *Phys. Rev.* B34 (1986) 8141; C. Bagnuls, C. Bervillier, D.I. Meiron and B.G. Nickel, *Phys. Rev.* B35 (1987) 3585; R. Schloms and V. Dohm, *Phys. Rev.* B42 (1990) 6142; H. J. Krause, R. Schloms and V. Dohm, *Z. Physik* B79 (1990) 287; F.J. Halfkann and V. Dohm, *Z. Phys.* B89 (1992) 79; G. Münster, J. Heitger, *Nucl. Phys.* B424 [FS] (1994) 582; R. Guida and J. Zinn-Justin, *Nucl. Phys.* B489 [FS] (1997) 626; S.A. Larin, M. Moennigmann, M. Stroesser and V. Dohm, *Phys. Rev.* B58 (1998) 3394; M. Stroesser, S.A. Larin and V. Dohm, *Nucl. Phys.* B540 (1999) 654.

For a review on high temperature series see, for example,

A.J. Guttmann, *Phase Transitions and Critical Phenomena* vol. 13, C. Domb and J. Lebowitz eds. (Academic Press, New York 1989).

A number of results extracted from high temperature series can be found in

J. Zinn-Justin, *J. Physique* (Paris) 42, (1981) 783; A.J. Guttmann, *J. Phys. A* 20 (1987) 1855; A.J. Liu and M.E. Fisher, *Physica* A156 (1989) 35; A.J. Guttmann and I.G. Enting, *J. Phys. A* 26 (1993) 807, 27 (1994) 8007; G. Bhanot, M. Creutz, U. Glässner and K. Schilling, *Phys. Rev.* B49 (1994) 12909; P. Butera and M. Comi, *Phys. Rev.* E55 (1997) 6391, *Phys. Rev.* B56 (1997) 8212, 9703018.

Results from Monte Carlo simulations for Ising model and of the SAW are found in

G.S. Pawley, R.H. Swendsen, D.J. Wallace and K.G. Wilson, *Phys. Rev.* B29 (1984) 4030; C.F. Baillie, R. Gupta, K.A. Hawick and G.S. Pawley, *Phys. Rev.* B45 (1992) 10438; B. Li, N. Madras and A.D. Sokal, *J. Stat. Phys.* 80 (1995) 661; S. Caracciolo, M.S. Causo and A. Pelissetto, *Phys. Rev.* E57 (1998) 1215.

For a review on amplitude ratios see

V. Privman, P.C. Hohenberg and A. Aharony, *Phase Transitions and Critical Phenomena* vol. 14, C. Domb and J. Lebowitz eds. (Academic Press, New York 1991).

Experiment, high temperature series and field theory RG are compared, using the results then available, in the Proceedings of the Cargèse Summer Institute 1980

Phase Transitions vol. B72, M. Lévy, J.C. Le Guillou and J. Zinn-Justin eds. (Plenum, New York 1982).

The experimental determination of the exponent ν for polymers is taken from

J.P. Cotton, *J. Physique Lett. (Paris)* 41 (1980) L231.

Experimental determinations of universal quantities for binary fluids are reported in

D. Beysens, A. Bourgou and P. Calmettes, *Phys. Rev.* A26 (1982) 3589; A. Kumar, H.R. Krishnamurthy and E.S.R. Gopal, *Phys. Rep.* 98 (1983) 57.

For experimental results on Helium superfluid transition see

L.S. Goldner and G. Ahlers, *Phys. Rev.* B45 (1992) 13129; J.A. Lipa, D.R. Swanson, J. Nissen, T.C.P. Chui and U.E. Israelson, *Phys. Rev. Lett.* 76 (1996) 944.

The exact two-dimensional results have first been conjectured by

B. Nienhuis, *Phys. Rev. Lett.* 49 (1982) 1062.

Finally, for a rather exhaustive list of references, see

A. Pelissetto, E. Vicari, *Critical Phenomena and Renormalization-Group Theory*, cond-mat/0012164.

J. Zinn-Justin, *J. Physique* (Paris) 42, (1981) 783; A.J. Guttmann, *J. Phys. A* 20 (1987) 1855; A.J. Liu and M.E. Fisher, *Physica* A156 (1989) 35; A.J. Guttmann and I.G. Enting, *J. Phys. A* 26 (1993) 807, 27 (1994) 8007; G. Bhanot, M. Creutz, U. Glässner and K. Schilling, *Phys. Rev.* B49 (1994) 12909; P. Butera and M. Comi, *Phys. Rev.* E55 (1997) 6391, *Phys. Rev.* B56 (1997) 8212, 9703018.

Results from Monte Carlo simulations for Ising model and of the SAW are found in

G.S. Pawley, R.H. Swendsen, D.J. Wallace and K.G. Wilson, *Phys. Rev.* B29 (1984) 4030; C.F. Baillie, R. Gupta, K.A. Hawick and G.S. Pawley, *Phys. Rev.* B45 (1992) 10438; B. Li, N. Madras and A.D. Sokal, *J. Stat. Phys.* 80 (1995) 661; S. Caracciolo, M.S. Causo and A. Pelissetto, *Phys. Rev.* E57 (1998) 1215.

For a review on amplitude ratios see

V. Privman, P.C. Hohenberg and A. Aharony, *Phase Transitions and Critical Phenomena* vol. 14, C. Domb and J. Lebowitz eds. (Academic Press, New York 1991).

Experiment, high temperature series and field theory RG are compared, using the results then available, in the Proceedings of the Cargèse Summer Institute 1980

Phase Transitions vol. B72, M. Lévy, J.C. Le Guillou and J. Zinn-Justin eds. (Plenum, New York 1982).

The experimental determination of the exponent ν for polymers is taken from

J.P. Cotton, *J. Physique Lett. (Paris)* 41 (1980) L231.

Experimental determinations of universal quantities for binary fluids are reported in

D. Beysens, A. Bourgou and P. Calmettes, *Phys. Rev.* A26 (1982) 3589; A. Kumar, H.R. Krishnamurthy and E.S.R. Gopal, *Phys. Rep.* 98 (1983) 57.

For experimental results on Helium superfluid transition see

L.S. Goldner and G. Ahlers, *Phys. Rev.* B45 (1992) 13129; J.A. Lipa, D.R. Swanson, J. Nissen, T.C.P. Chui and U.E. Israelson, *Phys. Rev. Lett.* 76 (1996) 944.

The exact two-dimensional results have first been conjectured by

B. Nienhuis, *Phys. Rev. Lett.* 49 (1982) 1062.

Finally, for a rather exhaustive list of references, see

A. Pelissetto, E. Vicari, *Critical Phenomena and Renormalization-Group Theory*, cond-mat/0012164.

30 THE $O(N)$ VECTOR MODEL FOR N LARGE

In the preceding chapters, we have derived universal properties of critical systems within the framework of the formal $\varepsilon = 4 - d$ expansion. It is, therefore, reassuring to verify that the results obtained in this way remain valid, at least in some limiting case, even when ε is no longer infinitesimal. We show in this chapter that, in the case 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.

Large N techniques are also useful because they allow to discuss other non-perturbative questions, including issues relevant to four-dimensional physics like renormalons and triviality.

Finally, we exhibit a surprising relation between the $(\phi^2)^2$ field theory and the non-linear σ-model valid to all orders in the $1/N$ expansion.

30.1 The Large N Action

We consider the partition function of an $O(N)$ symmetric model

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)] \exp\left[-\mathcal{S}(\phi)\right], \tag{30.1}$$

where ϕ is a N-component vector field, and where the euclidean action $\mathcal{S}(\phi)$ has the form

$$\mathcal{S}(\phi) = \int \left[\frac{1}{2}\left[\partial_\mu \phi(x)\right]^2 + NU\left(\phi^2(x)/N\right)\right] \mathrm{d}^d x\,, \tag{30.2}$$

$U(\rho)$ being a general polynomial. The reason for the peculiar N dependence will become clear shortly. A cut-off Λ, consistent with the symmetry, is implied.

The $(\phi^2)^2$ field theory (25.42),

$$\mathcal{S}(\phi) = \int \left\{\frac{1}{2}\left[\partial_\mu \phi(x)\right]^2 + \frac{1}{2} r\phi^2(x) + \frac{u}{4!}\left[\phi^2(x)\right]^2\right\} \mathrm{d}^d x\,, \tag{30.3}$$

($u = \Lambda^{4-d} g$) corresponds to the polynomial

$$U(\rho) = \frac{1}{2} r\rho + \frac{uN}{4!}\rho^2. \tag{30.4}$$

The solution of the general model in the large N limit is based on an idea of mean field theory type: it is expected that for N large $O(N)$ invariant quantities self-average and, therefore, have small fluctuations. For example,

$$\left\langle \phi^2(x)\phi^2(y)\right\rangle \underset{N\to\infty}{\sim} \left\langle \phi^2(x)\right\rangle \left\langle \phi^2(y)\right\rangle .$$

This suggests to take $\phi^2(x)$ as a dynamical variable, instead 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, we use the identity

$$1 = N\int \mathrm{d}\rho\, \delta(\phi^2 - N\rho) = \frac{N}{4i\pi}\int \mathrm{d}\rho\, \mathrm{d}\lambda\, \mathrm{e}^{\lambda(\phi^2 - N\rho)/2}, \tag{30.5}$$

where the integration contour in the complex λ plane is parallel to the imaginary axis. The insertion of this identity into the functional integral (30.1) yields another representation of the partition function

$$\mathcal{Z} = \int [\mathrm{d}\boldsymbol{\phi}][\mathrm{d}\rho][\mathrm{d}\lambda] \exp\left[-\mathcal{S}(\boldsymbol{\phi}, \rho, \lambda)\right] \tag{30.6}$$

with

$$\mathcal{S}(\boldsymbol{\phi}, \rho, \lambda) = \int \left[\frac{1}{2}\left[\partial_\mu \boldsymbol{\phi}(x)\right]^2 + NU\big(\rho(x)\big) + \frac{1}{2}\lambda(x)\big(\boldsymbol{\phi}^2(x) - N\rho(x)\big)\right] \mathrm{d}^d x\,. \tag{30.7}$$

The new functional integral is gaussian in $\boldsymbol{\phi}$, the integral over the field $\boldsymbol{\phi}$ can thus be performed and the dependence on N of the partition function becomes explicit. Actually, it is convenient to separate the components of $\boldsymbol{\phi}$ 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. Then,

$$\mathcal{Z}(H) = \int [\mathrm{d}\sigma][\mathrm{d}\rho][\mathrm{d}\lambda] \exp\left[-\mathcal{S}_N(\sigma, \rho, \lambda) + \int \mathrm{d}^d x\, H(x)\sigma(x)\right] \tag{30.8}$$

with

$$\begin{aligned}\mathcal{S}_N(\sigma, \rho, \lambda) = &\int \left[\frac{1}{2}\left(\partial_\mu \sigma\right)^2 + NU\big(\rho(x)\big) + \frac{1}{2}\lambda(x)(\sigma^2(x) - N\rho(x))\right] \mathrm{d}^d x \\ &+ \frac{1}{2}(N-1)\operatorname{tr}\ln\left[-\nabla^2 + \lambda(\cdot)\right].\end{aligned} \tag{30.9}$$

We now take the large N limit with a fixed function $U(\rho)$. We assume that the expectation values of fields scale with N like $\sigma = O(N^{1/2})$, $\rho, \lambda = O(1)$. With this ansatz $\mathcal{S}_N$ is of order N for N large and the functional integral can be calculated by the steepest descent method.

$\boldsymbol{\phi}^2$ correlation functions. In this formalism, it is natural to consider correlation functions involving the ρ-field, which is proportional to the $\boldsymbol{\phi}^2$ field, that is, 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 $\boldsymbol{\phi}^2(x,t)$ by $\left\langle \boldsymbol{\phi}^2(x,t)\right\rangle$ ($\langle\cdot\rangle$ 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.

General vector field theories. The algebraic method presented here can easily be generalized to actions that depend on several vector 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.

Let us take the example of two fields ϕ_1, ϕ_2 and assume that the interaction is an arbitrary function of the three $O(N)$ invariant quantities, the scalar products $\phi_1 \cdot \phi_2$, ϕ_1^2 and ϕ_2^2. We then introduce three fields $\rho_{ij}(x)$ and implement in the functional 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 (30.5). We substitute $N\rho_{ij}(x)$ for $\phi_i(x) \cdot \phi_j(x)$ in the interaction. The functional 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 is obtained by the steepest descent method.

30.2 Large N Limit: Saddle Point Equations, Critical Domain

We look for a uniform saddle point ($\sigma(x), \rho(x), \lambda(x)$ space-independent), $\sigma(x) = \sigma$, $\rho(x) = \rho$, $\lambda(x) = m^2$ because the λ saddle point value must be positive.

The action density $\mathcal{E}$ in zero field ($H = 0$) then becomes

$$\mathcal{E}(\sigma, \rho, m^2) = NU(\rho) + \frac{1}{2}m^2(\sigma^2 - N\rho) + \frac{1}{2}N\frac{1}{(2\pi)^d}\int^{\Lambda} \mathrm{d}^d k \, \ln(k^2 + m^2). \tag{30.10}$$

Differentiating then $\mathcal{E}(\sigma, \rho, m^2)$ with respect to σ, ρ and m^2, we obtain the saddle point equations:

$$m^2\sigma = 0\,, \tag{30.11a}$$

$$U'(\rho) - \tfrac{1}{2}m^2 = 0\,, \tag{30.11b}$$

$$\sigma^2/N - \rho + \Omega_d(m) = 0\,, \tag{30.11c}$$

where we have defined a function that will appear in all large N discussions:

$$\Omega_d(m) = \frac{1}{(2\pi)^d}\int^{\Lambda} \frac{\mathrm{d}^d p}{p^2 + m^2}\,. \tag{30.12}$$

For $d > 2$ and $m \ll \Lambda$, the dominant contributions are

$$\Omega_d(m) - \Omega_d(0) = -C(d)m^{d-2} + a(d)m^2\Lambda^{d-4} + O\big(m^4\Lambda^{d-6}\big) \tag{30.13}$$

with

$$N_d = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}\,, \tag{30.14a}$$

$$C(d) = -\frac{\pi}{2\sin(\pi d/2)}N_d = \frac{1}{(4\pi)^{d/2}}\Gamma(1 - d/2)\,, \tag{30.14b}$$

where we have introduced for convenience the usual loop factor N_d. The constant $a(d)$ that for $d < 4$ characterizes the leading correction in equation (30.13), depends explicitly on the regularization, that is, the way large momenta are cut.

Discussion. We verify that equations (30.11) have solutions scaling in N consistently with the ansatz. We see from the origin of the tr ln term in expression (30.9) that, at leading order, m is the mass of the $N - 1$ $\boldsymbol{\pi}$ components. Equation (30.11a) implies either $\sigma = 0$ or $m = 0$. If $\sigma = 0$ and $m \neq 0$ the $O(N)$ symmetry is unbroken, and the

N fields have the same mass m. If instead $\sigma \neq 0$, then $m = 0$, the $O(N)$ symmetry is spontaneously broken and the π-field is massless. These $N-1$ massless modes are a manifestation of the Goldstone phenomenon. We then note from equation (30.11c) that the solution $m = 0$ can exist only for $d > 2$, because at $d = 2$, the integral is IR divergent. This property is consistent with the Mermin–Wagner–Coleman theorem: 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 $d \leq 2$.

30.2.1 The $(\phi^2)^2$ theory: Phases and exponents

The discussion from now on will be specific to the example of the $(\phi^2)^2$ theory where $U(\rho)$ is given by equation (30.4). We note that the large N limit is taken at uN fixed. Moreover, we assume $d > 2$ except if stated explicitly otherwise.

In this case, 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

$$\lambda(x) = r + \tfrac{1}{6}Nu\rho(x), \tag{30.15}$$

and equation (30.11b) becomes

$$m^2 = r + \tfrac{1}{6}Nu\rho\,. \tag{30.16}$$

After elimination of $\rho(x)$, the large N action reduces to

$$\begin{aligned} \mathcal{S}_N(\sigma,\lambda) = {}& \frac{1}{2}\int \mathrm{d}^d x \left[(\partial_\mu\sigma)^2 + \lambda(x)\sigma^2(x) - \frac{3}{u}\lambda^2(x) + \frac{6}{u}r\lambda(x)\right] \\ & + \frac{1}{2}(N-1)\,\mathrm{tr}\ln\left[-\nabla^2 + \lambda(\cdot)\right]. \end{aligned} \tag{30.17}$$

Diagrammatic interpretation. In the large N limit of the $(\phi^2)^2$ field theory, the leading perturbative contributions come from chains of “bubble” diagrams of the form displayed in figure 30.1. These diagrams form asymptotically a geometric series which is summed by the algebraic techniques of section 30.1.

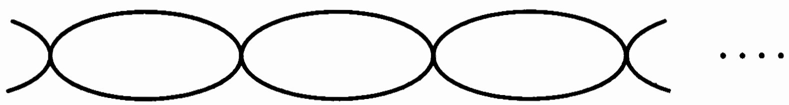


Fig. 30.1 The leading diagrams in the large N limit in the $(\phi^2)^2$ theory.

The low temperature phase. We first assume that σ, the expectation value of the field, does not vanish and thus $m = 0$. Equations (30.11c) and (30.16) then yield

$$\sigma^2/N = \rho - \rho_c = (6/u)(r_c - r)\,, \tag{30.18}$$

where we have introduced the constants ρ_c and r_c,

$$\rho_c = \Omega_d(0)\,, \quad r_c = -Nu\rho_c/6 < 0\,. \tag{30.19}$$

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 . Introducing the parameter

$$\tau = 6(r - r_c)/u\,, \tag{30.20}$$

which characterizes the deviation from the critical temperature, we find

$$\sigma = \sqrt{-\tau} \equiv (-\tau)^\beta \ \Rightarrow\ \beta = \tfrac{1}{2}\,. \tag{30.21}$$

In the N large limit, the exponent β remains mean-field like, in all dimensions.

The high temperature phase. Above T_c, σ vanishes. In expression (30.17), we see that the σ-propagator then has a free-field form

$$\Delta_\sigma(\mathbf{p}) = \frac{1}{p^2 + m^2}\,. \tag{30.22}$$

Therefore, 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 $\boldsymbol{\phi}$ field.

From equation (30.11c), we verify 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 (30.19,30.20) inside equations (30.16) and (30.11c), we then find

$$m^2 = (u/6)(N\rho - N\rho_c + \tau), \tag{30.23a}$$

$$\rho - \rho_c = -m^2 D_d(m), \tag{30.23b}$$

where $D_d(m)$ is the integral:

$$D_d(m) \equiv \frac{1}{m^2}\left(\Omega_d(0) - \Omega_d(m)\right) = \frac{1}{(2\pi)^d}\int^\Lambda \frac{\mathrm{d}^d p}{p^2\,(p^2 + m^2)}\,. \tag{30.24}$$

(i) For $d > 4$, the function $D_d(m)$ has a limit for $m = 0$. This implies $\rho - \rho_c \propto m^2$ and, therefore, at leading order:

$$m = 1/\xi \propto \sqrt{\tau} \ \Rightarrow\ \nu = \tfrac{1}{2}\,. \tag{30.25}$$

For $d > 4$, the correlation exponent ν has a mean field or gaussian value.

(ii) For $2 < d < 4$ instead, the function $D_d(m)$ has the expansion for m small (equation (30.13)):

$$D_d(m) = C(d)m^{d-4} - a(d)\Lambda^{d-4} + \left(m^2\Lambda^{d-6}\right). \tag{30.26}$$

The leading m-dependent contribution, for $m \to 0$, in equation (30.23a) now comes from $\rho - \rho_c$. Keeping only the leading term in (30.26), we obtain

$$m = \xi^{-1} \sim \tau^{1/(2-\varepsilon)}, \tag{30.27}$$

which shows that the exponent ν is no longer gaussian:

$$\nu = \frac{1}{2-\varepsilon} = \frac{1}{d-2}\,. \tag{30.28}$$

(iii) For $d = 4$, the leading m-dependent contribution in equation (30.23a) still comes from $\rho - \rho_c$:

$$D_4(m) = \frac{1}{(2\pi)^4} \int^{\Lambda} \frac{d^4 p}{p^2 (p^2 + m^2)} \underset{m \to 0}{\sim} \frac{1}{8\pi^2} \ln(\Lambda/m).$$

The correlation length no longer has a power law behaviour but instead 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.

(iv) Examining equation (30.11c) for $\sigma = 0$ and $d = 2$, we find that the correlation length becomes large only for $r \to -\infty$. This peculiar situation will be discussed in the framework of the non-linear σ-model.

Finally, in the critical limit, $\tau = 0$, m vanishes and thus from the form (30.22) of the σ-propagator, we find that the critical exponent η remains gaussian in all dimensions d:

$$\eta = 0 \ \Rightarrow \ d_\phi = \tfrac{1}{2}(d-2)\,. \tag{30.29}$$

One verifies that the exponents β, ν, η satisfy the scaling relation proven within the framework of the ε-expansion:

$$\beta = \nu d_\phi\,.$$

30.2.2 Singular free energy and scaling equation of state

In a constant magnetic field H in the σ direction, the free energy density $\mathcal{W}(H)/V_d$ (V_d is the space volume) is given by

$$\mathcal{W}(H)/V_d = \ln \mathcal{Z}/V_d = -NU(\rho) - \frac{1}{2} m^2(\sigma^2 - N\rho) + H\sigma - \frac{N}{2}\frac{1}{(2\pi)^d} \int d^d p \ln[(p^2+m^2)/p^2],$$

where the saddle point values σ, ρ, m^2 are solutions of equations (30.11b) and (30.11c) and the modified saddle point equation (30.11a):

$$m^2 \sigma = H\,. \tag{30.30}$$

The magnetization M is

$$M = V_d^{-1} \frac{\partial \mathcal{W}}{\partial H} = \sigma\,,$$

because partial derivatives of $\mathcal{W}$ with respect to m^2, σ vanish as a consequence of the saddle point equations. The thermodynamic potential $\Gamma(M)$ is the Legendre transform of $\mathcal{W}(H)$. In the $(\phi^2)^2$ theory, using the form (30.17) of the action, one finds

$$\begin{aligned}\mathcal{G}(M) \equiv \Gamma(M)/V_d &= HM - \mathcal{W}(H)/V_d \\ &= -\frac{3}{2u} m^4 + \frac{3r}{u} m^2 + \frac{1}{2} m^2 M^2 + \frac{N}{2}\frac{1}{(2\pi)^d} \int d^d p \ln[(p^2 + m^2)/p^2].\end{aligned}$$

We note

$$\frac{d}{dm^2} \frac{1}{(2\pi)^d} \int d^d p \ln[(p^2 + m^2)/p^2] = \rho_c - m^2 D_d(m).$$

The integral can thus be evaluated for Λ large in terms of r_c and the expansion (30.26):

$$\frac{1}{(2\pi)^d} \int d^d p \ln[(p^2 + m^2)/p^2] = -2\frac{C(d)}{d} m^d - \frac{6r_c}{Nu} m^2 + \frac{a(d)}{2} m^4 \Lambda^{d-4} + O\left(m^6 \Lambda^{d-6}\right).$$

The thermodynamical potential density becomes

$$\mathcal{G}(M) = \frac{3}{2}\left(\frac{1}{u^*} - \frac{1}{u}\right) m^4 + \frac{3(r - r_c)}{u} m^2 + \frac{1}{2} m^2 M^2 - \frac{NC(d)}{d} m^d, \tag{30.31}$$

where we have defined

$$u^* = \frac{6}{Na(d)} \Lambda^\varepsilon . \tag{30.32}$$

Dimensions $d < 4$. Since the dimensions $d > 4$ are simple we now discuss only the 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

$$\mathcal{G}(M) = \frac{1}{2}\tau m^2 + \frac{1}{2} m^2 M^2 - \frac{NC(d)}{d} m^d, \tag{30.33}$$

where τ has been defined in (30.20).

The general property of the Legendre transformation and the saddle point equation (30.11c) imply that the derivative of Γ with respect to m vanishes, a condition which allows to calculate m directly from $\mathcal{G}$,

$$2 \frac{\partial \mathcal{G}}{\partial m^2} = \tau + M^2 - NC(d) m^{d-2} = 0,$$

and thus

$$m = \left[\frac{1}{NC(d)} \left(\tau + M^2\right)\right]^{1/(d-2)} .$$

It follows that the thermodynamic potential, at leading order in the critical domain, is given by

$$\mathcal{G}(M) \sim \frac{(d-2)}{2d} \frac{1}{\left(NC(d)\right)^{2/(d-2)}} (\tau + M^2)^{d/(d-2)} . \tag{30.34}$$

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

$$H = \frac{\partial \mathcal{G}}{\partial M} = h_0 M^\delta f(\tau/M^2), \tag{30.35}$$

in which h_0 is a normalization constant, the exponent δ is given by

$$\delta = \frac{d+2}{d-2}, \tag{30.36}$$

in agreement with the general scaling relation relation $\delta = d/d_\phi - 1$, and the function $f(x)$ by

$$f(x) = (1 + x)^{2/(d-2)} . \tag{30.37}$$

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 26.5). In particular, the equation of state can be cast into the parametric form $\{M, \tau\} \mapsto \{R, \theta\}$:

$$M = R^\beta \theta , \quad \tau = 3R\left(1 - \theta^2\right) ,$$

and, therefore,

$$H = h_0 R^{\beta\delta} \theta \left(3 - 2\theta^2\right)^{2/(d-2)} .$$

Leading corrections to scaling. The m^4 term yields the leading corrections to the scaling behaviour. It is subleading by a power of τ:

$$m^4/m^d = O(\tau^{(4-d)/(d-2)}).$$

We conclude

$$\omega\nu = (4-d)/(d-2) \;\Rightarrow\; \omega = 4-d \,. \tag{30.38}$$

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, we obtain the specific heat $\mathcal{C}$ at fixed magnetization:

$$\mathcal{C} = \frac{1}{(d-2)} \frac{1}{\left(NC(d)\right)^{2/(d-2)}} (\tau + M^2)^{(4-d)/(d-2)}. \tag{30.39}$$

For $M = 0$, we identify the specific exponent

$$\alpha = \frac{4-d}{d-2}, \tag{30.40}$$

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 details see Section 29.7):

$$(R_\xi^+)^d = \frac{4N}{(d-2)^3} \frac{\Gamma(3-d/2)}{(4\pi)^{d/2}}, \quad R_c = \frac{4-d}{(d-2)^2}. \tag{30.41}$$

30.2.3 The ϕ^2 two-point function

Differentiating twice the action (30.9) with respect to $\lambda(x)$, then replacing the field $\lambda(x)$ by its expectation value m^2, we find the λ-propagator $\Delta_\lambda(\mathbf{p})$ above T_c:

$$\Delta_\lambda(\mathbf{p}) = -\frac{2}{N} \left[\frac{6}{Nu} + B_\Lambda(\mathbf{p}, m)\right]^{-1}, \tag{30.42}$$

where $B_\Lambda(\mathbf{p}, m)$ is the bubble diagram of figure 30.2:

$$B_\Lambda(\mathbf{p}, m) = \frac{1}{(2\pi)^d} \int^\Lambda \frac{d^d q}{(\mathbf{q}^2 + m^2)\left[(\mathbf{p}-\mathbf{q})^2 + m^2\right]}. \tag{30.43}$$

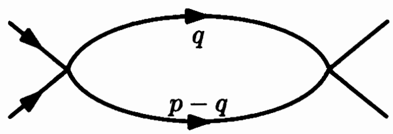


Fig. 30.2 The "bubble" diagram $B_\Lambda(\mathbf{p}, m)$.

and, therefore,

$$H = h_0 R^{\beta\delta}\theta \left(3 - 2\theta^2\right)^{2/(d-2)}.$$

Leading corrections to scaling. The m^4 term yields the leading corrections to the scaling behaviour. It is subleading by a power of τ:

$$m^4/m^d = O(\tau^{(4-d)/(d-2)}).$$

We conclude

$$\omega\nu = (4-d)/(d-2) \ \Rightarrow\ \omega = 4-d\,. \tag{30.38}$$

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, we obtain the specific heat $\mathcal{C}$ at fixed magnetization:

$$\mathcal{C} = \frac{1}{(d-2)}\frac{1}{\left(NC(d)\right)^{2/(d-2)}}(\tau + M^2)^{(4-d)/(d-2)}. \tag{30.39}$$

For $M = 0$, we identify the specific exponent

$$\alpha = \frac{4-d}{d-2}, \tag{30.40}$$

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 details see Section 29.7):

$$(R_\xi^+)^d = \frac{4N}{(d-2)^3}\frac{\Gamma(3-d/2)}{(4\pi)^{d/2}}, \qquad R_c = \frac{4-d}{(d-2)^2}. \tag{30.41}$$

30.2.3 The ϕ^2 two-point function

Differentiating twice the action (30.9) with respect to $\lambda(x)$, then replacing the field $\lambda(x)$ by its expectation value m^2, we find the λ-propagator $\Delta_\lambda(\mathbf{p})$ above T_c:

$$\Delta_\lambda(\mathbf{p}) = -\frac{2}{N}\left[\frac{6}{Nu} + B_\Lambda(\mathbf{p},m)\right]^{-1}, \tag{30.42}$$

where $B_\Lambda(\mathbf{p},m)$ is the bubble diagram of figure 30.2:

$$B_\Lambda(\mathbf{p},m) = \frac{1}{(2\pi)^d}\int^\Lambda \frac{\mathrm{d}^d q}{(\mathbf{q}^2+m^2)\left[(\mathbf{p}-\mathbf{q})^2+m^2\right]}. \tag{30.43}$$

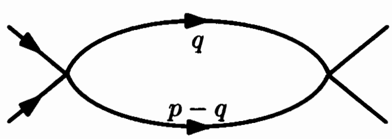


Fig. 30.2 The "bubble" diagram $B_\Lambda(\mathbf{p},m)$.

The λ-propagator is negative because the λ-field is imaginary. The λ-field is related by equation (30.15) to the ρ and thus ϕ^2 two-point function:

$$\left\langle \phi^2 \phi^2 \right\rangle = -\frac{12}{u\left[1+(Nu/6)B_\Lambda(\mathbf{p},m)\right]}. \tag{30.44}$$

At zero momentum, we recover the specific heat, up to an additive constant.

In the critical theory ($m = 0$ at this order) for $2 \le d \le 4$, the denominator of Δ_λ is dominated at low momentum by the integral

$$B_\Lambda(\mathbf{p},0) = \frac{1}{(2\pi)^d}\int^\Lambda \frac{\mathrm{d}^d q}{q^2(p-q)^2} \underset{2<d<4}{=} b(d)p^{-\varepsilon} - a(d)\Lambda^{-\varepsilon} + O\left(\Lambda^{d-6}p^2\right), \tag{30.45}$$

where

$$b(d) = -\frac{\pi}{\sin(\pi d/2)}\frac{\Gamma^2(d/2)}{\Gamma(d-1)}N_d\,, \tag{30.46}$$

and thus

$$\Delta_\lambda(\mathbf{p}) \underset{p\to 0}{\sim} -\frac{2}{Nb(d)}p^{\varepsilon}. \tag{30.47}$$

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

$$[\lambda] = \tfrac{1}{2}(d+\varepsilon) = 2\,, \tag{30.48}$$

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 $\Delta_\lambda \propto 1/\ln(\Lambda/p)$.

(ii) Note, therefore, that for $d \le 4$ the contributions generated by the term proportional to $\lambda^2(x)$ in (30.17) are always negligible in the critical domain.

30.3 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, m solution of equation (30.23) satisfies

$$\Lambda\frac{\partial m}{\partial \Lambda} + N\varepsilon a(d)\Lambda^{-\varepsilon}\frac{u^2}{6}\frac{\partial m}{\partial u} = 0\,,$$

where the constant $a(d)$, defined in (30.26), depends on the cut-off procedure but for $\varepsilon = 4-d$ small satisfies

$$a(d) \sim 1/(8\pi^2\varepsilon). \tag{30.49}$$

Then, we set (equation (30.32))

$$u = g\Lambda^{\varepsilon}, \quad g^* = u^*\Lambda^{-\varepsilon} = 6/(Na)\,. \tag{30.50}$$

In the new variables Λ, g, τ, we obtain an equation which expresses that m is RG invariant

$$\left(\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \eta_2(g)\tau\frac{\partial}{\partial\tau}\right) m(\tau,g,\Lambda) = 0 \tag{30.51}$$

with

$$\beta(g) = -\varepsilon g(1 - g/g^*), \tag{30.52}$$

$$\nu^{-1}(g) = 2 + \eta_2(g) = 2 - \varepsilon g/g^*. \tag{30.53}$$

When $a(d)$ is positive (but this not true for all regularizations, see the discussion below), one finds an IR fixed point g^*, as well as exponents $\omega = \varepsilon$, and $\nu^{-1} = d-2$, in agreement with equations (30.38,30.28). 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 two for $\varepsilon < 2$, and this extends the property for N large to all dimensions $2 \le d \le 4$.

Finally, applying the RG equations to the propagator (30.22), we find

$$\eta(g) = 0\,, \tag{30.54}$$

a result consistent with the value (30.29) 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. Let us now consider the correlation length or mass m given by equation (30.23). If we keep the leading correction to the integral for m small (equation (30.26)), we find

$$\frac{6}{u} - \frac{6}{u^*} + NC(d)m^{-\varepsilon} + O\left(m^2\Lambda^{d-6}\right) = \frac{\tau}{m^2}\,, \tag{30.55}$$

where equation (30.50) has been used. We see that the leading correction again vanishes for $u = u^*$. Actually, all correction terms suppressed by powers of order ε for $d \to 4$ 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 (30.38,30.28).

In the same way, if we keep the leading correction to the λ-propagator in the critical theory (equation (30.45)), we find

$$\Delta_\lambda(\mathbf{p}) = -\frac{2}{N}\left[\frac{6}{Nu} - \frac{6}{Nu^*} + b(d)p^{-\varepsilon}\right]^{-1}, \tag{30.56}$$

where terms of order Λ^{-2} have been neglected. The leading corrections to scaling again exactly cancel for $u = u^*$ as expected.

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 $6/Nu - a(d)\Lambda^{-\varepsilon}$ where the value of $a(d)$ depends on the cut-off procedure and thus of contributions of irrelevant operators. Let us call 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:

$$\frac{6\Lambda^\varepsilon}{Nu} - a(d) = \frac{6\Lambda^\varepsilon}{Nu'} - a'(d),$$

consistent with the special form (30.52) of the β-function.

(ii) *The sign of* $a(d)$. It is generally assumed that $a(d) > 0$. 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

$$a(d) \underset{d \to 4}{\sim} \frac{1}{8\pi^2 \varepsilon}.$$

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 artifact of the large N limit.

Another way of stating the problem is to examine directly the relation between bare and renormalized coupling constant. Calling $g_{\mathrm{r}} m^{4-d}$ the renormalized four-point function at zero momentum, we find

$$m^{4-d} g_{\mathrm{r}} = \frac{\Lambda^{4-d} g}{1 + \Lambda^{4-d} g N B_{\Lambda}(0, m)/6}. \tag{30.57}$$

In the limit $m \ll \Lambda$, the relation can be written as

$$\frac{1}{g_{\mathrm{r}}} = \frac{(d-2)NC(d)}{12} + \left(\frac{m}{\Lambda}\right)^{4-d} \left(\frac{1}{g} - \frac{Na(d)}{6}\right). \tag{30.58}$$

We see that when $a(d) < 0$, 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.

30.4 Small Coupling Constant, Large Momentum Expansions for $d < 4$

In Section 30.7, we discuss more systematically the $1/N$ expansion, but the $1/N$ correction to the two-point function can already help us to understand why the massless field theory exists but has no perturbative expansion for $d < 4$.

We have seen that in the framework at the $1/N$ expansion, we can 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 a 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 (figure 30.3), containing two $\lambda^2\sigma$ vertices. After mass renormalization and in the large cut-off limit, we find

$$\Gamma^{(2)}_{\sigma\sigma}(\mathbf{p}) = p^2 + \frac{2}{N(2\pi)^d} \int \frac{\mathrm{d}^d q}{(6/Nu) + b(d) q^{-\varepsilon}} \left(\frac{1}{(\mathbf{p}+\mathbf{q})^2} - \frac{1}{\mathbf{q}^2}\right) + O\left(\frac{1}{N^2}\right). \tag{30.59}$$

The integral cannot be expanded in a Taylor series in the coupling constant u. Instead, an analytic study reveals that it has a small u expansion of the form

$$\sum_{k \geq 1} \alpha_k u^k p^{2-k\varepsilon} + \beta_k u^{(2+2k)/\varepsilon} p^{-2k}. \tag{30.60}$$

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

$$M(s) = \int_0^\infty \mathrm{d}u\, u^{-1-s} f(u)$$

has a pole at $s = t$. Applying the transformation to the integral, and inverting q and u integrations, we find the integral

$$\int_0^\infty \mathrm{d}u \frac{u^{-1-s}}{(6/Nu) + b(d)q^{-\varepsilon}} = \frac{N}{6}\left(\frac{Nb(d)q^{-\varepsilon}}{6}\right)^{1-s} \frac{\pi}{\sin \pi s} .$$

Then, the result of the remaining q integration follows from the evaluation of the generic integral (30.93).

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.

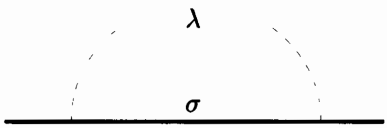


Fig. 30.3 The diagram contributing to $\Gamma^{(2)}_{\sigma\sigma}$ at order $1/N$.

30.5 Dimension 4: Triviality, Higgs Mass

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 $\langle\sigma\sigma\sigma\sigma\rangle$ at momenta of order $\mu \ll \Lambda$, is given by

$$g_\mathrm{r} = \frac{g}{1 + \frac{1}{6}NgB_\Lambda(\mu, 0)} , \tag{30.61}$$

where $B_\Lambda(\mathbf{p}, 0)$ corresponds to the bubble diagram (figure 30.2)

$$B_\Lambda(\mathbf{p}, 0) \underset{p \ll \Lambda}{\sim} \frac{1}{8\pi^2} \ln(\Lambda/p) + \text{ const.}. \tag{30.62}$$

We see that when the ratio μ/Λ goes to zero, the renormalized coupling constant g_r vanishes, for that all g (see Section 35.1.1). This is the so-called *triviality* property. In the traditional treatment of quantum field field, one insists in taking the infinite cut-off Λ limit. Here, one then finds a free field theory. This observation has the following consequence: in four dimensions, it is impossible to construct a ϕ^4 field theory consistent (in the sense of satisfying all usual physical requirements) on all length scales for non-zero coupling. In the logic of ***effective low energy field theories***, however, this is no longer an issue. The quantum field theory has a limited range of validity. The triviality property simply implies that the renormalized or effective charge decreases logarithmically when the consistent physical range increases, as indicated by equations (30.61,30.62). Note that if g is generic (not too small) and Λ/μ large, g_r is essentially independent of the initial coupling constant. Only if the bare coupling is small is the renormalized coupling an adjustable, but bounded, quantity.

If one works formally and, ignoring the problem, expresses the leading contribution to the four-point function in terms of the renormalized constant:

$$\frac{g}{1+\frac{N}{48\pi^2}g\ln(\Lambda/p)}=\frac{g_\mathrm{r}}{1+\frac{N}{48\pi^2}g_\mathrm{r}\ln(\mu/p)}\,,$$

one discovers that the function has a pole for

$$p=\mu\,\mathrm{e}^{48\pi^2/(Ng_\mathrm{r})}\,.$$

This pole is unphysical 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 (30.59)),

$$\Gamma^{(2)}_{\sigma\sigma}(\mathbf{p})=p^2+\frac{g_\mathrm{r}}{3N(2\pi)^4}\int\frac{\mathrm{d}^4q}{1+Ng_\mathrm{r}\ln(\mu/q)/48\pi^2}\left(\frac{1}{(\mathbf{p}+\mathbf{q})^2}-\frac{1}{\mathbf{q}^2}-\text{subtrac.}\right),$$

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

$$\int^\infty\frac{\mathrm{d}^4q}{q^6}\left(-\frac{Ng_\mathrm{r}}{48\pi^2}\ln(\mu/q)\right)^k\underset{k\to\infty}{\propto}\left(\frac{Ng_\mathrm{r}}{96\pi^2}\right)^k k!\,.$$

The perturbative manifestation of the Landau ghost is the appearance of (renormalon) contributions to the perturbation series which are not Borel summable (see Section 42.3.2 for details), in contrast with the semi-classical contributions corresponding to the finite momentum integration, which are Borel summable, but invisible for N large. Note finally that this UV problem is independent of the mass of the field ϕ, that we have taken zero only to simplify expressions.

The mass of the σ field in the phase of broken symmetry. The ϕ^4 theory is a piece of the Standard Model, and the field σ then represents the Higgs field. In Section 35.1.2, we show that with some reasonable assumptions it is possible to establish for finite N a semi-quantitative bound on the Higgs mass. Let us examine here what happens for N large.

In the phase of broken symmetry the action, after translation of expectation values, includes a term proportional to $\sigma\lambda$ and thus the propagators of the fields σ and λ are elements of a 2×2 matrix $\mathbf{M}$:

$$\mathbf{M}^{-1}(p)=\begin{pmatrix}p^2 & \sigma\\ \sigma & -3/u-\frac{1}{2}NB_\Lambda(p,0)\end{pmatrix},$$

coupling. In the logic of *effective low energy field theories*, however, this is no longer an issue. The quantum field theory has a limited range of validity. The triviality property simply implies that the renormalized or effective charge decreases logarithmically when the consistent physical range increases, as indicated by equations (30.61,30.62). Note that if g is generic (not too small) and Λ/μ large, g_r is essentially independent of the initial coupling constant. Only if the bare coupling is small is the renormalized coupling an adjustable, but bounded, quantity.

If one works formally and, ignoring the problem, expresses the leading contribution to the four-point function in terms of the renormalized constant:

$$\frac{g}{1+\frac{N}{48\pi^2}g\ln(\Lambda/p)}=\frac{g_{\mathrm{r}}}{1+\frac{N}{48\pi^2}g_{\mathrm{r}}\ln(\mu/p)}\,,$$

one discovers that the function has a pole for

$$p=\mu\,\mathrm{e}^{48\pi^2/(Ng_{\mathrm{r}})}\,.$$

This pole is unphysical 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 (30.59)),

$$\Gamma^{(2)}_{\sigma\sigma}(\mathbf{p})=p^2+\frac{g_{\mathrm{r}}}{3N(2\pi)^4}\int\frac{\mathrm{d}^4q}{1+Ng_{\mathrm{r}}\ln(\mu/q)/48\pi^2}\left(\frac{1}{(\mathbf{p}+\mathbf{q})^2}-\frac{1}{\mathbf{q}^2}-\text{ subtrac.}\right),$$

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

$$\int^{\infty}\frac{\mathrm{d}^4q}{q^6}\left(-\frac{Ng_{\mathrm{r}}}{48\pi^2}\ln(\mu/q)\right)^k\underset{k\to\infty}{\propto}\left(\frac{Ng_{\mathrm{r}}}{96\pi^2}\right)^k k!\,.$$

The perturbative manifestation of the Landau ghost is the appearance of (renormalon) contributions to the perturbation series which are not Borel summable (see Section 42.3.2 for details), in contrast with the semi-classical contributions corresponding to the finite momentum integration, which are Borel summable, but invisible for N large. Note finally that this UV problem is independent of the mass of the field ϕ, that we have taken zero only to simplify expressions.

The mass of the σ field in the phase of broken symmetry. The ϕ^4 theory is a piece of the Standard Model, and the field σ then represents the Higgs field. In Section 35.1.2, we show that with some reasonable assumptions it is possible to establish for finite N a semi-quantitative bound on the Higgs mass. Let us examine here what happens for N large.

In the phase of broken symmetry the action, after translation of expectation values, includes a term proportional to $\sigma\lambda$ and thus the propagators of the fields σ and λ are elements of a 2×2 matrix $\mathbf{M}$:

$$\mathbf{M}^{-1}(p)=\begin{pmatrix}p^2 & \sigma\\ \sigma & -3/u-\frac{1}{2}NB_{\Lambda}(p,0)\end{pmatrix},$$

where $\sigma = \langle \sigma(x) \rangle = |\langle \boldsymbol{\phi}(x) \rangle|$. In four dimensions, B_Λ is given by equation (30.62). It is convenient to introduce a coupling constant dependent mass scale $\Lambda(u)$ defined by

$$\frac{48\pi^2}{Nu} + 8\pi^2 B_\Lambda(p,0) \sim \ln\bigl(\Lambda(u)/p\bigr),$$

and thus

$$\Lambda(u) \propto \mathrm{e}^{48\pi^2/Nu}\,\Lambda\,.$$

The mass of the field σ at this order is a solution to the equation $\det \mathbf{M} = 0$. One finds

$$p^2 \ln\bigl(\Lambda(u)/p\bigr) = -(16\pi^2/N)\sigma^2 \ \Rightarrow\ m_\sigma^2 \ln\bigl(i\Lambda(u)/m_\sigma\bigr) = (16\pi^2/N)\sigma^2.$$

The mass m_σ solution to the equation is complex, because the particle σ can decay into massless Goldstone bosons. At σ fixed, the mass decreases when the cut-off increases or when the coupling constant goes to zero. Expressing that the mass must be smaller than the cut-off, one obtains an upper-bound on m_σ (but which slightly depends on the chosen regularization) that one can compare with the one obtained in the general discussion of Section 35.1.2.

30.6 The Non-Linear σ-Model in the Large N Limit

We now establish, within the framework of the $1/N$ expansion, a remarkable relation between the $(\boldsymbol{\phi}^2)^2$ theory and the non-linear σ model defined in Chapter 14.

We first slightly rewrite the action (30.17). We shift the field $\lambda(x)$ by its expectation value m^2, $\lambda(x) \mapsto m^2 + \lambda(x)$:

$$\begin{aligned}\mathcal{S}_N(\sigma,\lambda) = {}& \frac{1}{2}\int \mathrm{d}^d x \left[(\partial_\mu\sigma)^2 + m^2\sigma^2 + \lambda(x)\sigma^2(x) - \frac{3}{u}\lambda^2(x) - \frac{6}{u}\left(m^2 - r\right)\lambda(x)\right] \\ & + \frac{(N-1)}{2}\operatorname{tr}\ln\left[-\nabla^2 + m^2 + \lambda(\cdot)\right]. \end{aligned} \tag{30.63}$$

We have noticed that the term proportional to $\int \mathrm{d}^d x\, \lambda^2(x)$, 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 but yields only logarithmic corrections). Actually, the constant part in the inverse propagator as written in equation (30.56) plays the role of a large momentum cut-off. We thus neglect the λ^2 term. If we then work backwards, reintroduce the initial field $\boldsymbol{\phi}$ and integrate over $\lambda(x)$ we now find

$$\mathcal{Z} = \int [\mathrm{d}\boldsymbol{\phi}(x)]\,\delta\left[\boldsymbol{\phi}^2(x) - \frac{6}{u}\left(m^2 - r\right)\right] \exp\left[-\int \frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}(x)\right)^2 \mathrm{d}^d x\right]. \tag{30.64}$$

In this form, we recognize the partition function of the $O(N)$ symmetric non-linear σ-model studied in Chapter 14, in an unusual parametrization. We have, therefore, discovered a remarkable correspondence: to all orders in an $1/N$ expansion the renormalized non-linear σ-model is identical to the renormalized $(\boldsymbol{\phi}^2)^2$ field theory at the IR fixed point, that is, for generic ϕ^4 coupling.

The large N limit. In order to show the correspondence between the set of parameters used in the two models more explicitly, we also solve directly the σ-model in the large N limit (see also Section 31.2). The partition function is given by the integral

$$\mathcal{Z} = \int \left[\mathrm{d}\boldsymbol{\phi}(x)\delta(\boldsymbol{\phi}^2(x) - 1)\right] \exp\left[-\mathcal{S}(\boldsymbol{\phi})\right] \tag{30.65}$$

where $\sigma = \langle \sigma(x) \rangle = | \langle \phi(x) \rangle |$. In four dimensions, B_Λ is given by equation (30.62). It is convenient to introduce a coupling constant dependent mass scale $\Lambda(u)$ defined by

$$\frac{48\pi^2}{Nu} + 8\pi^2 B_\Lambda(p,0) \sim \ln\bigl(\Lambda(u)/p\bigr),$$

and thus

$$\Lambda(u) \propto \mathrm{e}^{48\pi^2/Nu}\,\Lambda\,.$$

The mass of the field σ at this order is a solution to the equation $\det \mathbf{M} = 0$. One finds

$$p^2 \ln\bigl(\Lambda(u)/p\bigr) = -(16\pi^2/N)\sigma^2 \;\Rightarrow\; m_\sigma^2 \ln\bigl(i\Lambda(u)/m_\sigma\bigr) = (16\pi^2/N)\sigma^2 .$$

The mass m_σ solution to the equation is complex, because the particle σ can decay into massless Goldstone bosons. At σ fixed, the mass decreases when the cut-off increases or when the coupling constant goes to zero. Expressing that the mass must be smaller than the cut-off, one obtains an upper-bound on m_σ (but which slightly depends on the chosen regularization) that one can compare with the one obtained in the general discussion of Section 35.1.2.

30.6 The Non-Linear σ-Model in the Large N Limit

We now establish, within the framework of the $1/N$ expansion, a remarkable relation between the $(\phi^2)^2$ theory and the non-linear σ model defined in Chapter 14.

We first slightly rewrite the action (30.17). We shift the field $\lambda(x)$ by its expectation value m^2, $\lambda(x) \mapsto m^2 + \lambda(x)$:

$$\begin{aligned} \mathcal{S}_N(\sigma,\lambda) = \frac{1}{2}\int \mathrm{d}^d x \left[(\partial_\mu \sigma)^2 + m^2\sigma^2 + \lambda(x)\sigma^2(x) - \frac{3}{u}\lambda^2(x) - \frac{6}{u}\left(m^2 - r\right)\lambda(x)\right] \\ + \frac{(N-1)}{2}\operatorname{tr}\ln\left[-\nabla^2 + m^2 + \lambda(\cdot)\right]. \end{aligned} \tag{30.63}$$

We have noticed that the term proportional to $\int \mathrm{d}^d x\, \lambda^2(x)$, 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 but yields only logarithmic corrections). Actually, the constant part in the inverse propagator as written in equation (30.56) plays the role of a large momentum cut-off. We thus neglect the λ^2 term. If we then work backwards, reintroduce the initial field ϕ and integrate over $\lambda(x)$ we now find

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)]\,\delta\left[\phi^2(x) - \frac{6}{u}\left(m^2 - r\right)\right] \exp\left[-\int \frac{1}{2}\left(\partial_\mu \phi(x)\right)^2 \mathrm{d}^d x\right]. \tag{30.64}$$

In this form, we recognize the partition function of the $O(N)$ symmetric non-linear σ-model studied in Chapter 14, in an unusual parametrization. We have, therefore, discovered a remarkable correspondence: to all orders in an $1/N$ 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 ϕ^4 coupling.

The large N limit. In order to show the correspondence between the set of parameters used in the two models more explicitly, we also solve directly the σ-model in the large N limit (see also Section 31.2). The partition function is given by the integral

$$\mathcal{Z} = \int \left[\mathrm{d}\phi(x)\delta(\phi^2(x) - 1)\right] \exp\left[-\mathcal{S}(\phi)\right] \tag{30.65}$$

with

$$\mathcal{S}(\phi) = \frac{1}{2T}\int \mathrm{d}^d x \left[(\partial_\mu \phi)^2\right]. \tag{30.66}$$

In this parametrization, the coupling constant T plays the role of the temperature, as the lattice regularization of the continuum σ model shows (Section 14.3.2). Therefore, this parametrization is well adapted to a study of low temperature phenomena.

We rewrite the partition function

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)\mathrm{d}\lambda(x)] \exp\left[-\mathcal{S}(\phi,\lambda)\right] \tag{30.67}$$

with

$$\mathcal{S}(\phi,\lambda) = \frac{1}{2T}\int \mathrm{d}^d x \left[(\partial_\mu \phi)^2 + \lambda\left(\phi^2 - 1\right)\right]. \tag{30.68}$$

Integrating, as we did in Section 30.1, over $N-1$ components of ϕ and calling σ the remaining component, we obtain

$$\mathcal{Z} = \int [\mathrm{d}\sigma(x)\mathrm{d}\lambda(x)] \exp\left[-\mathcal{S}_N(\sigma,\lambda)\right] \tag{30.69}$$

with

$$\mathcal{S}_N(\sigma,\lambda) = \frac{1}{2T}\int \left[(\partial_\mu \sigma)^2 + \left(\sigma^2(x) - 1\right)\lambda(x)\right] \mathrm{d}^d x + \frac{1}{2}(N-1)\operatorname{tr}\ln\left[-\nabla^2 + \lambda(\cdot)\right]. \tag{30.70}$$

The large N limit is here taken at TN fixed. The saddle point equations, analogous to equations (30.11), are

$$m^2\sigma = 0\,, \tag{30.71a}$$

$$\sigma^2 = 1 - (N-1)\Omega_d(m)T\,, \tag{30.71b}$$

where we have set $\langle\lambda(x)\rangle = m^2$, and the function $\Omega_d(m)$ has been defined by equation (30.12).

At low temperature for $d > 2$, σ is different from zero and thus m, which is the mass of the π-field, vanishes. Equation (30.71b) gives the spontaneous magnetization:

$$\sigma^2 = 1 - (N-1)\Omega_d(0)T. \tag{30.72}$$

Setting

$$\frac{1}{T_c} = (N-1)\Omega_d(0, \tag{30.73}$$

we can write equation (30.72) as

$$\sigma^2 = 1 - T/T_c\,. \tag{30.74}$$

Thus, T_c is the critical temperature where σ vanishes.

Above T_c, σ instead vanishes and m, which is now the common mass of the π- and σ-field, is for $2 < d < 4$ given in terms of the function (30.24) by

$$\frac{1}{T_c} - \frac{1}{T} = (N-1)D_d(m) \underset{d<4}{\sim} m^{d-2}(N-1)D_d(1) + O\left(m^2\Lambda^{d-4}\right). \tag{30.75}$$

We recover the scaling form 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,

$$\mathcal{G}(M) = \Gamma(M)/V_d = \frac{d-2}{2d} \frac{1}{\left(NC(d)\right)^{2/(d-2)}} (M^2 - 1 + T/T_c)^{d/(d-2)}, \tag{30.76}$$

a result which extends equation (30.34) 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 30.7.

Renormalization group. Anticipating slightly some results of Chapter 31, we calculate RG functions at leading order for N large. First, we introduce a dimensionless temperature t:

$$T = \Lambda^{2-d} t\,, \quad T_c = \Lambda^{2-d} t_c\,.$$

We then express that the mass or correlation length solutions of equation (30.75) are RG invariant:

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t}\right) [(N-1) D_d(m)] = \left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(t) \frac{\partial}{\partial t}\right) \Lambda^{d-2} \left(\frac{1}{t_c} - \frac{1}{t}\right) = 0\,.$$

It follows

$$\beta(t) = (d-2) t (1 - t/t_c). \tag{30.77}$$

Therefore, 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 will be analysed in detail in Chapter 31 we postpone further discussion.

Conversely, we can solve the equation in general and find

$$m(t) \propto \Lambda \exp\left[\int^t \frac{\mathrm{d}t'}{\beta(t')}\right]. \tag{30.78}$$

Using the β-function (30.77), we find an explicit form that coincides for $m \ll \Lambda$ with the solution of equation (30.75) only for $d < 4$. This indicates that the RG equations (31.19) are valid only for $d < 4$.

We note for later purpose that the RG invariant crossover length (31.34) is given in the large N limit for $d < 4$ by

$$\begin{aligned} \xi(t) &= \frac{t^{1/(d-2)}}{\Lambda} \exp\left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{(d-2)t'}\right) \mathrm{d}t'\right] = \frac{1}{\Lambda} \left(\frac{1}{t} - \frac{1}{t_c}\right)^{-1/(d-2)} \\ &= \left(\frac{1}{T} - \frac{1}{T_c}\right)^{-1/(d-2)}. \end{aligned} \tag{30.79}$$

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 (30.74), we obtain

$$\zeta(t) = (d-2) t/t_c\,. \tag{30.80}$$

We recover the scaling form 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,

$$\mathcal{G}(M) = \Gamma(M)/V_d = \frac{d-2}{2d}\frac{1}{\left(NC(d)\right)^{2/(d-2)}}(M^2 - 1 + T/T_c)^{d/(d-2)}, \tag{30.76}$$

a result which extends equation (30.34) 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 30.7.

Renormalization group. Anticipating slightly some results of Chapter 31, we calculate RG functions at leading order for N large. First, we introduce a dimensionless temperature t:

$$T = \Lambda^{2-d}t\,, \quad T_c = \Lambda^{2-d}t_c\,.$$

We then express that the mass or correlation length solutions of equation (30.75) are RG invariant:

$$\left(\Lambda\frac{\partial}{\partial\Lambda} + \beta(t)\frac{\partial}{\partial t}\right)[(N-1)D_d(m)] = \left(\Lambda\frac{\partial}{\partial\Lambda} + \beta(t)\frac{\partial}{\partial t}\right)\Lambda^{d-2}\left(\frac{1}{t_c} - \frac{1}{t}\right) = 0\,.$$

It follows

$$\beta(t) = (d-2)t(1 - t/t_c). \tag{30.77}$$

Therefore, 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 will be analysed in detail in Chapter 31 we postpone further discussion.

Conversely, we can solve the equation in general and find

$$m(t) \propto \Lambda \exp\left[\int^t \frac{\mathrm{d}t'}{\beta(t')}\right]. \tag{30.78}$$

Using the β-function (30.77), we find an explicit form that coincides for $m \ll \Lambda$ with the solution of equation (30.75) only for $d < 4$. This indicates that the RG equations (31.19) are valid only for $d < 4$.

We note for later purpose that the RG invariant crossover length (31.34) is given in the large N limit for $d < 4$ by

$$\begin{aligned}\xi(t) &= \frac{t^{1/(d-2)}}{\Lambda}\exp\left[\int_0^t\left(\frac{1}{\beta(t')} - \frac{1}{(d-2)t'}\right)\mathrm{d}t'\right] = \frac{1}{\Lambda}\left(\frac{1}{t} - \frac{1}{t_c}\right)^{-1/(d-2)} \\ &= \left(\frac{1}{T} - \frac{1}{T_c}\right)^{-1/(d-2)}.\end{aligned} \tag{30.79}$$

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 (30.74), we obtain

$$\zeta(t) = (d-2)t/t_c\,. \tag{30.80}$$

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 (30.71b) determines the mass m. Since $\sigma = 0$, one finds

$$m \propto \Lambda\,\mathrm{e}^{-2\pi/(Nt)}, \tag{30.81}$$

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, though the critical temperature vanishes, one can define a critical domain, that is, a range of temperature for which the correlation length is large in the microscopic scale:

$$m \ll \Lambda \ \Rightarrow\ t = T = O\bigl(1/\ln(\Lambda/m)\bigr).$$

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

$$\Gamma^{(2)}_{\sigma\sigma}(p) = p^2 + m^2\,. \tag{30.82}$$

If we substitute for m the large N expression (30.81) and expand the two-point function in powers of the coupling constant t, we see 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 is also 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 $-h\int \mathrm{d}^2x\,\sigma(x)/T$ to the action, the physical mass becomes calculable in perturbation theory,

$$\frac{h^2}{m^4} = 1 - \frac{Nt}{2\pi}\bigl(\ln(\Lambda/m) + \text{ const.}\bigr).$$

Corrections to scaling and the dimension 4. In equation (30.75), we have neglected corrections to scaling. If we take into account the leading correction we get instead

$$m^2\left(C(d)m^{d-4} - a(d)\Lambda^{d-4}\right) \propto t - t_c\,,$$

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 30.3, we discover that, although the non-linear σ-model superficially depends on one parameter less than the corresponding ϕ^4 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 (30.75) takes a different form, the dominant term in the r.h.s. 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 renormalization group the physics of the non-linear σ model it is necessary to introduce the operator $\int \mathrm{d}^dx\,\lambda^2(x)$, which irrelevant for $d<4$, becomes marginal, and to go over to the ϕ^4 field theory.

30.7 The $1/N$-Expansion: An Alternative Field Theory

Power counting. Higher order terms in the steepest descent calculation of the functional integral (30.8) generate a systematic $1/N$ expansion. We now analyse the terms in the action (30.63) 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 (30.47) of the λ-propagator, we have inferred the canonical dimension $[\lambda]$ of the field $\lambda(x)$:

$$2\,[\lambda]-\varepsilon=d\,,\qquad \text{that is,}\quad [\lambda]=2\,.$$

As noted above, λ^2 has dimension $4>d$ and thus is irrelevant. The interaction term $\int \lambda(x)\sigma^2(x)\mathrm{d}^d x$ has dimension 0. One verifies that the non-local interactions involving the λ-field, coming from the expansion of the $\operatorname{tr}\ln$, have all also canonical dimension 0:

$$\left[\operatorname{tr}\left[\lambda(x)\left(-\nabla^2+m^2\right)^{-1}\right]^k\right]=k\,[\lambda]-2k=0\,.$$

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, $2<d\le 4$. 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 power series in $1/N$. However, because the interactions are non-local, the results of renormalization theory do not immediately apply. We 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 limit.

An alternative field theory. To be able to use the standard results of renormalization theory, we reformulate the critical theory to deal with the non-local interactions. Neglecting corrections to scaling we start from the non-linear σ-model in the form (30.68):

$$\mathcal{Z}=\int[\mathrm{d}\lambda(x)]\,[\mathrm{d}\boldsymbol{\phi}(x)]\exp\left[-\mathcal{S}(\boldsymbol{\phi},\lambda)\right],\tag{30.83}$$

$$\mathcal{S}(\boldsymbol{\phi},\lambda)=\frac{1}{2T}\int\mathrm{d}^d x\left[\left(\partial_\mu\boldsymbol{\phi}\right)^2+\lambda\left(\boldsymbol{\phi}^2-1\right)\right].\tag{30.84}$$

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 (30.84) a term quadratic in λ that in the tree approximation of standard perturbation theory generates a λ-propagator of the form (30.47). The modified action $\mathcal{S}_v$ reads

$$\mathcal{S}_v(\boldsymbol{\phi},\lambda)=\frac{1}{2}\int\mathrm{d}^d x\left\{\frac{1}{T}\left[\left(\partial_\mu\boldsymbol{\phi}\right)^2+\lambda\left(\boldsymbol{\phi}^2-1\right)\right]-\frac{1}{v^2}\lambda(-\partial^2)^{-\varepsilon/2}\lambda\right\}.\tag{30.85}$$

In the limit where the parameter v goes to infinity, the coefficient of the additional term vanishes, and the initial action is recovered.

We below consider only the critical theory. This means that the couplings of all relevant interactions will be set to their critical values. These interactions contain a term linear in λ and a polynomial in $\boldsymbol{\phi}^2$ of degree depending on the dimension. Note that in a discrete set of dimensions some monomials become just renormalizable. We, therefore, work in generic dimensions. The quantities we calculate are regular in the dimension. The field theory with the action (30.85) 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 $\phi \mapsto \phi\sqrt{T}$, $\lambda \mapsto v\lambda$:

$$\mathcal{S}_v(\phi,\lambda) = \frac{1}{2}\int \mathrm{d}^d x \left[(\partial_\mu\phi)^2 + v\lambda\phi^2 - \lambda(-\partial^2)^{-\varepsilon/2}\lambda + \text{relevant terms}\right].$$

The renormalized critical action then reads

$$[\mathcal{S}_v]_{\text{ren}} = \frac{1}{2}\int \mathrm{d}^d x \left[Z_\phi\,(\partial_\mu\phi)^2 + v_{\mathrm{r}} Z_v \lambda\phi^2 - \lambda(-\partial^2)^{-\varepsilon/2}\lambda + \text{relevant terms}\right]. \quad (30.86)$$

It follows that the RG equations for 1PI correlation functions of l λ fields and n ϕ fields in the critical theory take the form

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta_{v^2}(v)\frac{\partial}{\partial v^2} - \frac{n}{2}\eta(v)\right]\Gamma^{(l,n)} = 0\,. \quad (30.87)$$

The solution to the RG equations (30.87) can be written as

$$\Gamma^{(l,n)}(\ell p, v, \Lambda) = Z^{-n/2}(\ell)\ell^{d-2l-n(d-2)/2}\Gamma^{(l,n)}(p, v(\ell), \Lambda) \quad (30.88)$$

with the usual definitions

$$\ell\frac{\mathrm{d}v^2}{\mathrm{d}\ell} = \beta(v(\ell))\,, \quad \ell\frac{\mathrm{d}\ln Z}{\mathrm{d}\ell} = \eta(v(\ell))\,.$$

We can then calculate the RG functions as 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

$$\Delta_\lambda(p) = -\frac{2p^{4-d}}{b(d)R(v)}\,, \quad (30.89)$$

where we have defined

$$R(v) = 2/b(d) + Nv^2.$$

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

$$\ell\frac{\mathrm{d}v^2}{\mathrm{d}\ell} = \rho v^2\,, \ \Rightarrow \ v^2(\ell) \sim \ell^\rho. \quad (30.90)$$

We then note that to each power of the λ field corresponds a power of v. It follows

$$\begin{aligned}\Gamma^{(l,n)}(\ell p, v, \Lambda) &\propto v^l(\ell)\ell^{d-2l-n(d-2+\eta)}\\ &\propto \ell^{d-(2-\rho/2)l-n(d-2+\eta)}.\end{aligned} \quad (30.91)$$

To compare with the result (25.76) obtained from the perturbative renormalization group 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,

$$2 - \rho/2 = d_{\phi^2} = d - 1/\nu\,. \quad (30.92)$$

30.8 Explicit Calculations: Critical Exponents

We explain here how $1/N$ corrections to exponents can be calculated and then give a few higher order results.

30.8.1 $1/N$ calculations

Most calculations at order $1/N$ rely on the evaluation of the generic integral

$$\frac{1}{(2\pi)^d}\int\frac{\mathrm{d}^d q}{(p+q)^{2\mu}q^{2\nu}}=p^{d-2\mu-2\nu}\frac{\Gamma(\mu+\nu-d/2)\Gamma(d/2-\mu)\Gamma(d/2-\nu)}{(4\pi)^{d/2}\Gamma(\mu)\Gamma(\nu)\Gamma(d-\mu-\nu)}\,. \tag{30.93}$$

For later purposes, it is convenient to set

$$X_1=\frac{2N_d}{b(d)}=\frac{4\Gamma(d-2)}{\Gamma(d/2)\Gamma(2-d/2)\Gamma^2(d/2-1)}=\frac{4\sin(\pi\varepsilon/2)\Gamma(2-\varepsilon)}{\pi\Gamma(1-\varepsilon/2)\Gamma(2-\varepsilon/2)}. \tag{30.94}$$

To compare with fixed dimension results note $X_1\sim 2(4-d)$ for $d\to 4$ and $X_1\sim(d-2)$ for $d\to 2$.

The calculation of the $\langle\phi\phi\rangle$ correlation function at order $1/N$ involves the evaluation of the diagram of figure 30.3. We want to determine the coefficient of $p^2\ln\Lambda/p$. Since we work at one-loop order we can instead replace the λ propagator $q^{-\varepsilon}$ by $q^{2\nu}$ and send the cut-off to infinity. We then use the result (30.93) with $\mu=1$. In the limit $2\nu\to-\varepsilon$ 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

$$\Gamma^{(2)}_{\sigma\sigma}(p)=p^2+\frac{\varepsilon}{4-\varepsilon}\frac{2N_d}{b(d)R(v)}v^2p^2\ln(\Lambda/p).$$

Expressing that the function satisfies the RG equation we obtain the function $\eta(v)$.

The second RG function can be deduced from the divergent parts of the $\langle\phi\phi\lambda\rangle$ function

$$\Gamma^{(3)}_{\sigma\sigma\lambda}=v+A_1v^3D^{-1}(v)\ln\Lambda+A_2v^5D^{-2}(v)\ln\Lambda+\text{ finite}$$

with

$$A_1=-\frac{2}{b(d)}N_d=-X_1$$
$$A_2=-\frac{4N}{b^2(d)}(d-3)b(d)N_d=-2N(d-3)X_1\,,$$

where A_1 and A_2 correspond to the diagrams of figures 30.4 and 30.5, respectively.

Applying the RG equation one finds the relation at order $1/N$:

$$\beta_{v^2}(v)=2v^2\eta(v)-2A_1v^4D^{-1}(v)-2A_2v^6D^{-2}(v). \tag{30.95}$$

We thus obtain

$$\eta(v)=\frac{\varepsilon v^2}{4-\varepsilon}X_1D^{-1}(v), \tag{30.96}$$

$$\beta_{v^2}(v)=\frac{8v^4}{4-\varepsilon}X_1D^{-1}(v)+4N(1-\varepsilon)v^6X_1D^{-2}(v), \tag{30.97}$$

where the first term in β_{v^2} comes from A_1 and η and the second from A_2.

Extracting the large v^2 behaviour, we find exponent η:

$$\eta = \frac{\varepsilon}{N(4-\varepsilon)} X_1 + O(1/N^2), \tag{30.98}$$

$$\rho = \frac{4(3-\varepsilon)(2-\varepsilon)}{N(4-\varepsilon)} X_1 > 0\,,$$

and thus

$$\frac{1}{\nu} = d - 2 + \frac{2(3-\varepsilon)(2-\varepsilon)}{N(4-\varepsilon)} X_1 + O(1/N^2). \tag{30.99}$$

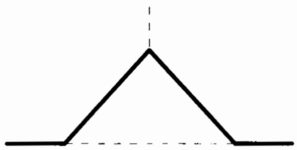


Fig. 30.4 Diagram contributing to $\Gamma^{(3)}_{\sigma\sigma\lambda}$ at order $1/N$ (A_1).

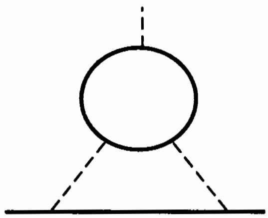


Fig. 30.5 Diagram contributing to $\Gamma^{(3)}_{\sigma\sigma\lambda}$ at order $1/N$ (A_2).

30.8.2 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 $2 \le d \le 4$, 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.

Generic dimensions. The exponents γ 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. The expansion of γ up to order $1/N$ can be directly deduced from the results of the preceding sections:

$$\gamma = \frac{1}{1-\varepsilon/2}\left(1 - \frac{3}{2N}X_1\right) + O\left(\frac{1}{N^2}\right). \tag{30.100}$$

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:

$$\omega = \varepsilon\left(1 - \frac{2(3-\varepsilon)^2}{(4-\varepsilon)N}X_1\right) + O\left(\frac{1}{N^2}\right), \tag{30.101}$$

$$\theta = \omega\nu = \frac{\varepsilon}{2-\varepsilon}\left(1 - \frac{2(3-\varepsilon)}{N}X_1\right) + O\left(\frac{1}{N^2}\right). \tag{30.102}$$

Note that the exponents are regular functions of ε up to $\varepsilon = 2$ and free of renormalon singularities at $\varepsilon = 0$.

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 refer the reader to the literature for details.

Three-dimensional results. Let us give the expansion of η in three dimensions at the order presently available:

$$\eta = \frac{\eta_1}{N} + \frac{\eta_2}{N^2} + \frac{\eta_3}{N^3} + O\left(\frac{1}{N^4}\right)$$

with

$$\eta_1 = \tfrac{8}{3\pi^2}\,, \quad \eta_2 = -\tfrac{8}{3}\eta_1^2\,, \quad \eta_3 = \eta_1^3\left[-\tfrac{797}{18} - \tfrac{61}{24}\pi^2 + \tfrac{27}{8}\psi''(1/2) + \tfrac{9}{2}\pi^2 \ln 2\right],$$

$\psi(x)$ being the logarithmic derivative of the Γ function.

The exponent γ is known only up to order $1/N^2$:

$$\gamma = 2 - \tfrac{24}{N\pi^2} + \tfrac{64}{N^2\pi^4}\left(\tfrac{44}{9} - \pi^2\right) + O\left(\tfrac{1}{N^3}\right).$$

Note that the $1/N$ expansion seems to be rapidly divergent and certainly a direct summation of these terms does not provide very reliable estimates of critical exponents in three dimensions for useful values of N.

Bibliographical Notes

As shown by Stanley the large N-limit of the classical N-vector model coincides with the spherical model solved by Berlin and Kac

T.H. Berlin and M. Kac, *Phys. Rev.* 86 (1952) 821; H.E. Stanley, *Phys. Rev.* 176 (1968) 718.

Early work on calculating critical properties includes

R. Abe, Prog. Theor. Phys. 48 (1972) 1414; 49 (1973) 113, 1074, 1877; S.K. Ma, *Phys. Rev. Lett.* 29 (1972) 1311; *Phys. Rev.* A7 (1973) 2172; M. Suzuki, *Phys. Lett.* 42A (1972) 5; *Prog. Theor. Phys.* 49 (1973) 424, 1106, 1440; R.A. Ferrel and D.J. Scalapino, *Phys. Rev. Lett.* 29 (1972) 413; K.G. Wilson, *Phys. Rev.* D7 (1973) 2911.

The contribution of order $1/N$ to the equation of state is given in

E. Brézin and D.J. Wallace, *Phys. Rev.* B7 (1973) 1967.

The spin–spin correlation in zero field is obtained in

M.E. Fisher and A. Aharony, *Phys. Rev. Lett.* 31 (1973) 1238; A. Aharony, *Phys. Rev.* B10 (1974) 2834; R. Abe and S. Hikami, *Prog. Theor. Phys.* 51 (1974) 1041.

The exponent ω has been calculated to order $1/N$ in

S.K. Ma, *Phys. Rev.* A10 (1974) 1818.

See also the contributions of S.K. Ma to

Phase Transitions and Critical Phenomena vol. 6, C. Domb and M.S. Green eds. (Academic Press, London 1976).

The consistency of the $1/N$ expansion to all orders has been proven in

I. Ya ref'eva, E.R. Nissimov and S.J. Pacheva, *Commun. Math. Phys.* 71 (1980) 213; A.N. Vasil'ev and M.Yu. Nalimov, *Teor. Mat. Fiz.* 55 (1983) 163.

At present, the longest $1/N$ series for exponents and amplitudes are found in

I. Kondor and T. Temesvari, *J. Physique Lett. (Paris)* 39 (1978) L99; Y. Okabe and M. Oku, *Prog. Theor. Phys.* 60 (1978) 1277, 1287; 61 (1979) 443; A.N. Vasil'ev, Yu.M. Pis'mak and Yu.R. Honkonen, *Teor. Mat. Fiz.* 46 (1981) 157; 50 (1982) 195.

See also

I. Kondor, T. Temesvari and L. Herenyi, *Phys. Rev.* B22 (1980) 1451.

Renormalization of operators is discussed in

K. Lang and W. Rühl, *Nucl. Phys.* B400 (1993) 597; *Z. Phys.* C61 (1994) 459.

The case of long range forces has been discussed in

S.K. Ma, *Phys. Rev.* A7 (1973) 2172.

For the Hartree–Fock point of view and QFT at finite temperature see

W.A. Bardeen and M. Moshe, *Phys. Rev.* D28 (1983) 1372.

Results concerning the β-function at order $1/N$ in the massive theory renormalized at zero momentum have been reported in

A. Pelissetto and E. Vicari, *Nucl. Phys.* B519 (1998) 626, cond-mat/9711078.

A calculation of the dimensions of composite operators to order $1/N^2$ is reported in S.E. Derkachov, A.N. Manashov *Nucl. Phys.* B522 (1998) 301, hep-th/9710015; *Phys. Rev. Lett.* 79 (1997) 1423, hep-th/9705020,

and the consequences for the stability of the fixed point of the non-linear σ model discussed.

Finally, a general review of large N techniques and applications can be found in

M. Moshe and J. Zinn-Justin, *Quantum Field Theory in the large N limit: a review*, Technion and Saclay preprint 2002, in preparation.

31 PHASE TRANSITIONS NEAR TWO DIMENSIONS

In Chapters 24–26, we have shown that universal properties of lattice spin models in the critical domain can be derived from an effective ϕ^4 euclidean field. We have determined a number of critical properties using RG methods within the context of the $\varepsilon = 4 - d$ expansion, an expansion valid near the so-called upper-critical dimension 4.

The general results obtained in this way, whose validity can only be established in an infinitesimal neighbourhood of dimension 4, have then been confirmed for all dimensions $2 < d < 4$ in the $O(N)$ symmetric spin model, in the large N limit, and more generally within the framework of the $1/N$ expansion.

A remarkable, and not totally expected, outcome of the large N analysis has been the discovery of the identity of the ϕ^4 field theory and the non-linear σ-model in the continuum limit (at least for generic ϕ^4 couplings) although from the point of view of perturbation theory these models are totally different. This property, totally mysterious at the classical level, emphasizes the essential nature of quantum fluctuations.

An additional prediction of the large N calculation concerns the behaviour of the critical domain when the dimension approaches 2 continuously: the critical temperature vanishes continuously and the 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 the lattice regularization indicates (see Section 14.3.2), 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 stressed in Chapter 23. Because, in contrast to models with discrete symmetries, they have a non-trivial long distance physics for any temperature below T_c, due to massless Goldstone modes, IR instabilities prevent a phase transition in two dimensions. Considerations based on the non-linear σ-model thus are relevant for the N-vector model only for $N \geq 2$. The main analytic tools will RG equations and $\varepsilon = d - 2$ expansion.

One property which plays an essential role is the UV asymptotic freedom of the σ-model in two dimensions. We expect some of its properties to generalize to other asymptotically free models. In two dimensions, various fermion self-interacting models share this property. The symmetry which is then broken is the chiral symmetry which prevents explicit mass terms in the action. As in the case of the non-linear σ-model, IR divergences forbid the existence of a massless phase. One model of this kind can be defined in continuous dimensions, the Gross–Neveu (GN), a simplification of the Nambu–Jona-Lasinio model, which we thus study in the second part of the chapter. It is renormalizable in two dimensions, and describes in perturbation theory only one phase, the phase with symmetry breaking. A model with the same symmetry can be identified, the Gross–Neveu–Yukawa (GNY) model which is renormalizable in four dimensions, and in which both phases can be reached already in the tree approximation. The study of these two models illustrates all the ideas and techniques developed in framework of the ϕ^4 theory and the non-linear σ-models, that is RG equations near two and four dimensions, and large N expansion.

1 PHASE TRANSITIONS NEAR TWO DIMENSIONS

In Chapters 24–26, we have shown that universal properties of lattice spin models in the critical domain can be derived from an effective ϕ^4 euclidean field. We have determined a number of critical properties using RG methods within the context of the $\varepsilon = 4 - d$ expansion, an expansion valid near the so-called upper-critical dimension 4.

The general results obtained in this way, whose validity can only be established in an infinitesimal neighbourhood of dimension 4, have then been confirmed for all dimensions $2 < d < 4$ in the $O(N)$ symmetric spin model, in the large N limit, and more generally within the framework of the $1/N$ expansion.

A remarkable, and not totally expected, outcome of the large N analysis has been the discovery of the identity of the ϕ^4 field theory and the non-linear σ-model in the continuum limit (at least for generic ϕ^4 couplings) although from the point of view of perturbation theory these models are totally different. This property, totally mysterious at the classical level, emphasizes the essential nature of quantum fluctuations.

An additional prediction of the large N calculation concerns the behaviour of the critical domain when the dimension approaches 2 continuously: the critical temperature vanishes continuously and the 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 the lattice regularization indicates (see Section 14.3.2), 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 stressed in Chapter 23. Because, in contrast to models with discrete symmetries, they have a non-trivial long distance physics for any temperature below T_c, due to massless Goldstone modes, IR instabilities prevent a phase transition in two dimensions. Considerations based on the non-linear σ-model thus are relevant for the N-vector model only for $N \geq 2$. The main analytic tools will RG equations and $\varepsilon = d - 2$ expansion.

One property which plays an essential role is the UV asymptotic freedom of the σ-model in two dimensions. We expect some of its properties to generalize to other asymptotically free models. In two dimensions, various fermion self-interacting models share this property. The symmetry which is then broken is the chiral symmetry which prevents explicit mass terms in the action. As in the case of the non-linear σ-model, IR divergences forbid the existence of a massless phase. One model of this kind can be defined in continuous dimensions, the Gross–Neveu (GN), a simplification of the Nambu–Jona-Lasinio model, which we thus study in the second part of the chapter. It is renormalizable in two dimensions, and describes in perturbation theory only one phase, the phase with symmetry breaking. A model with the same symmetry can be identified, the Gross–Neveu–Yukawa (GNY) model which is renormalizable in four dimensions, and in which both phases can be reached already in the tree approximation. The study of these two models illustrates all the ideas and techniques developed in framework of the ϕ^4 theory and the non-linear σ-models, that is RG equations near two and four dimensions, and large N expansion.

31.1 $(\phi^2)^2$ Field Theory and Non-Linear σ-Model

Before we begin our discussion of the σ-model, let us review shortly the arguments that, beyond the $1/N$ expansion, 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.

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 14.3.2. The regularized model belong to the class studied in Chapters 24–26: 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

$$\mathcal{Z} = \int \prod_i \mathrm{d}\mathbf{S}_i \, \delta\left(\mathbf{S}_i^2 - 1\right) \exp\left[-\mathcal{H}(\mathbf{S})/T\right], \tag{31.1}$$

in which the configuration energy $\mathcal{H}$ has the form

$$\mathcal{H}(\mathbf{S})/T = -\sum_{ij} V_{ij} \mathbf{S}_i \cdot \mathbf{S}_j / T \,. \tag{31.2}$$

We have shown, by different techniques, summation of the most IR terms in the mean field expansion, direct RG analysis, that the universal properties in the critical domain of such models can described in $4-\varepsilon$ dimensions by the corresponding $(\phi^2)^2$ field theory.

The $(\phi^2)^2$ in the ordered phase. Conversely, we now consider the $(\phi^2)^2$ field theory:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} r \phi^2 + \frac{1}{4!} u \left(\phi^2\right)^2\right], \tag{31.3}$$

in the low temperature phase ($r < r_c$).

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 noted, 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 functional integral:

$$\mathcal{Z} = \int [\mathrm{d}\phi] \exp[-\mathcal{S}(\phi)],$$

setting

$$\phi(x) = \rho(x) \hat{\phi}(x) \quad \text{with} \quad \hat{\phi}^2(x) = 1 \,. \tag{31.4}$$

The functional integral becomes (assuming a lattice or in the continuum a dimensional regularization):

$$\mathcal{Z} = \int \left[\rho^{N-1}(x) \mathrm{d}\rho(x)\right] \left[\mathrm{d}\hat{\phi}(x)\right] \exp\left[-\mathcal{S}(\rho, \hat{\phi})\right] \tag{31.5}$$

with

$$\mathcal{S}(\rho, \hat{\phi}) = \int \mathrm{d}^d x \left\{\frac{1}{2}\rho^2(x) \left[\partial_\mu \hat{\phi}(x)\right]^2 + \frac{1}{2}\left[\partial_\mu \rho(x)\right]^2 + \frac{1}{2} r \rho^2 + \frac{1}{4!} u \rho^4\right\}. \tag{31.6}$$

31.1 $(\phi^2)^2$ Field Theory and Non-Linear σ-Model

Before we begin our discussion of the σ-model, let us review shortly the arguments that, beyond the $1/N$ expansion, 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.

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 14.3.2. The regularized model belong to the class studied in Chapters 24–26: 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

$$\mathcal{Z} = \int \prod_i \mathrm{d}\mathbf{S}_i \, \delta\left(\mathbf{S}_i^2 - 1\right) \exp\left[-\mathcal{H}(\mathbf{S})/T\right], \tag{31.1}$$

in which the configuration energy $\mathcal{H}$ has the form

$$\mathcal{H}(\mathbf{S})/T = -\sum_{ij} V_{ij} \mathbf{S}_i \cdot \mathbf{S}_j / T \,. \tag{31.2}$$

We have shown, by different techniques, summation of the most IR terms in the mean field expansion, direct RG analysis, that the universal properties in the critical domain of such models can described in $4-\varepsilon$ dimensions by the corresponding $(\phi^2)^2$ field theory.

The $(\phi^2)^2$ in the ordered phase. Conversely, we now consider the $(\phi^2)^2$ field theory:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} r \phi^2 + \frac{1}{4!} u \left(\phi^2\right)^2\right], \tag{31.3}$$

in the low temperature phase ($r < r_c$).

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 noted, 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 functional integral:

$$\mathcal{Z} = \int [\mathrm{d}\phi] \exp[-\mathcal{S}(\phi)],$$

setting

$$\phi(x) = \rho(x) \hat{\phi}(x) \quad \text{with} \quad \hat{\phi}^2(x) = 1 \,. \tag{31.4}$$

The functional integral becomes (assuming a lattice or in the continuum a dimensional regularization):

$$\mathcal{Z} = \int \left[\rho^{N-1}(x) \mathrm{d}\rho(x)\right] \left[\mathrm{d}\hat{\phi}(x)\right] \exp\left[-\mathcal{S}(\rho, \hat{\phi})\right] \tag{31.5}$$

with

$$\mathcal{S}(\rho, \hat{\phi}) = \int \mathrm{d}^d x \left\{\frac{1}{2}\rho^2(x) \left[\partial_\mu \hat{\phi}(x)\right]^2 + \frac{1}{2}\left[\partial_\mu \rho(x)\right]^2 + \frac{1}{2} r \rho^2 + \frac{1}{4!} u \rho^4\right\}. \tag{31.6}$$

The integration over the $\rho(x)$ field generates an effective action $\mathcal{S}_{\text{eff}}(\hat{\phi})$ for the field $\hat{\phi}$:

$$\exp\left[-\mathcal{S}_{\text{eff}}(\hat{\phi})\right] = \int \left[\rho^{N-1}(x)\mathrm{d}\rho(x)\right] \exp\left[-\mathcal{S}(\rho,\hat{\phi})\right]. \tag{31.7}$$

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. The term dominant at large distance is the term with only two derivatives. Due to the $O(N)$ symmetry, it is has the form of the action of the non-linear σ-model studied in Chapter 14,

$$\mathcal{S}_{\text{eff}}(\hat{\phi}) = \frac{1}{2}\left\langle \rho^2 \right\rangle \int \mathrm{d}^d x \left[\partial_\mu \hat{\phi}(x)\right]^2 + \text{ irrelevant terms} . \tag{31.8}$$

The neglected interactions have at least four derivatives and correspond to irrelevant operators.

More explicitly since the ρ field is massive the functional integral (31.7) can thus be calculated perturbatively. Let us call R the expectation value of $\rho(x)$,

$$\rho(x) = R + \rho'(x), \tag{31.9}$$

which in the tree approximation is also the length of the spontaneous magnetization. In terms of ρ', the action (31.6) reads

$$\begin{aligned}\mathcal{S}(\rho',\hat{\phi}) = \int \mathrm{d}^d x \biggl\{ &\frac{1}{2}\left(R^2 + 2R\rho' + \rho'^2\right)\left[\partial_\mu \hat{\phi}(x)\right]^2 + \frac{1}{2}\left[\partial_\mu \rho'(x)\right]^2 \\ &+ \frac{1}{2} r (R+\rho')^2 + \frac{1}{4!} u (R+\rho')^4 \biggr\}.\end{aligned}$$

Neglecting all fluctuations of the field ρ', we obtain $\mathcal{S}_{\text{eff}}$ at leading order:

$$\mathcal{S}^{(0)}_{\text{eff}}(\hat{\phi}) = \frac{1}{2} R^2 \int \mathrm{d}^d x \left[\partial_\mu \hat{\phi}(x)\right]^2 . \tag{31.10}$$

We recognize the leading term in the action (31.8).

Loop corrections coming from the integration over ρ' renormalize the coefficient R^2 in (31.10). They also generate additional higher derivative, and, therefore, irrelevant, $\hat{\phi}$-interactions, and counter-terms for the non-linear σ-model, which in perturbation theory is more divergent that the $(\phi^2)^2$ field theory.

It is difficult to characterize precisely the domain of validity of the effective theory because the ρ integral can only be calculated perturbatively, but we expect it to include the low temperature phase T fixed, $T < T_c$. In this regime, the non-linear σ-model (31.8) 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 ϕ^4 field theory in the continuum.

At higher temperatures, problems will arise if the configurations for which ϕ vanishes become important. One example is provided by the $O(2)$ model in two dimensions (see Chapter 33).

31.2 The Non-Linear σ-Model: Symmetry Breaking, RG Equations

We now analyse the non-linear σ-model from the point of view of phase structure and renormalization group. The role of dimension 2 will emerge.

31.2.1 IR divergences, symmetry breaking and the role of dimension 2

In Chapter 14, we have shown that the perturbative phase of the non-linear σ-model is automatically a phase in which the $O(N)$ symmetry is spontaneously broken, and $(N-1)$ components of $\mathbf{S}_i$, called hereafter π_i, are massless Goldstone modes. This leads to IR divergences for $d \le 2$ which we reexamine in the context of phase transitions.

(i) We have argued in Section 23.5 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 $d > 2$ perturbation theory which also predicts spontaneous symmetry breaking (SSB), is not IR divergent. At $T < T_c$ fixed, the large distance behaviour of the theory is dominated by the massless or spin wave excitations. On the other hand nothing can be said, in perturbation theory, of a possible critical region $T \sim T_c$.

(ii) For $d \le 2$, we know from the Mermin–Wagner theorem (see also the analysis of Section 23.5) that SSB with ordering ($\langle \mathbf{S} \rangle \ne 0$) is impossible in a model with a continuous symmetry and short range forces, and this is again consistent with the appearance of IR divergences in perturbation theory. For $d \le 2$, the critical temperature T_c vanishes and perturbation theory makes sense only in the presence of an IR cut-off which breaks explicitly the symmetry and orders the spins. Therefore, nothing can be said about the long distance properties of the unbroken theory directly from perturbation theory.

To go somewhat beyond perturbation theory, we now use field theory RG methods. It is, therefore, 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. We, therefore, introduce an IR cut-off in the form of a magnetic field. We then proceed in formal analogy with the case of the $(\boldsymbol{\phi}^2)^2$ field theory, that is, study the theory in $2+\varepsilon$ dimension 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.

31.2.2 Perturbation Theory, RG equations

By three different methods, we have shown that the large distance physics of the N-vector model can be described below T_c by the non-linear σ-model. We now study this model from the point of view of renormalization and renormalization group. To generate perturbation theory, we parametrize the field $\hat{\boldsymbol{\phi}}$ of previous section as (see Section 14.1):

$$\hat{\boldsymbol{\phi}} = \begin{cases} \sigma(x) \\ \boldsymbol{\pi}(x) \end{cases}$$

and eliminate locally the field $\sigma(x)$ by

$$\sigma(x) = \left(1 - \boldsymbol{\pi}^2(x)\right)^{1/2}.$$

As we have done for the $(\boldsymbol{\phi}^2)^2$ 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 magnetic field h:

$$\mathcal{Z}(h) = \int \left[\left(1 - \boldsymbol{\pi}^2(x)\right)^{-1/2} \mathrm{d}\boldsymbol{\pi}(x)\right] \exp\left[-\mathcal{S}(\boldsymbol{\pi}, h)\right] \tag{31.11}$$

with

$$\mathcal{S}(\boldsymbol{\pi},h)=\frac{1}{T}\int \mathrm{d}^d x\left\{\frac{1}{2}\left[(\partial_\mu \boldsymbol{\pi}(x))^2+\frac{(\boldsymbol{\pi}\cdot\partial_\mu\boldsymbol{\pi}(x))^2}{1-\boldsymbol{\pi}^2(x)}\right]-h\sqrt{1-\boldsymbol{\pi}^2(x)}\right\}. \tag{31.12}$$

We immediately introduce a dimensionless coupling t:

$$T=t\,\Lambda^{2-d}. \tag{31.13}$$

The properties of this action have been studied, in slightly different notations, in Chapter 14. We here directly borrow the results. We have shown that it is renormalizable in two dimensions and characterized the form of the renormalized action. We recall that the presence of a non-vanishing magnetic field h is required, for $d\le 2$, 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. The renormalized action has the form

$$\mathcal{S}_\mathrm{r}(\boldsymbol{\pi}_\mathrm{r})=\frac{\mu^{d-2}Z}{2t_\mathrm{r}Z_t}\int \mathrm{d}^d x\left[(\partial_\mu\boldsymbol{\pi}_\mathrm{r})^2+(\partial_\mu\sigma_\mathrm{r})^2\right]-\frac{\mu^{d-2}}{t_\mathrm{r}}h_\mathrm{r}\int\sigma_\mathrm{r}(x)\mathrm{d}^d x\,, \tag{31.14}$$

in which μ is the renormalization scale, and

$$\sigma_\mathrm{r}(x)=\left[Z^{-1}-\boldsymbol{\pi}_\mathrm{r}^2\right]^{1/2}. \tag{31.15}$$

Note that the renormalization constants can be chosen h independent. This is automatically realized in the minimal subtraction scheme.

The relations

$$t=(\Lambda/\mu)^{d-2}Z_t t_\mathrm{r}\,,\quad \boldsymbol{\pi}_\mathrm{r}(x)=Z^{-1/2}\boldsymbol{\pi}(x) \tag{31.16}$$

imply

$$h=Z_h h_\mathrm{r}\,,\quad Z_h=Z_t/\sqrt{Z}\,. \tag{31.17}$$

With our conventions the coupling constant, which is proportional to the temperature, is dimensionless. The relation between the cut-off dependent and the renormalized correlation functions is

$$Z^{n/2}(\Lambda/\mu,t)\Gamma^{(n)}(p_i;t,h,\Lambda)=\Gamma_\mathrm{r}^{(n)}(p_i;t_\mathrm{r},h_\mathrm{r},\mu). \tag{31.18}$$

Differentiating with respect to Λ at renormalized parameters fixed, we obtain the RG equations:

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(t)\frac{\partial}{\partial t}-\frac{n}{2}\zeta(t)+\rho(t)h\frac{\partial}{\partial h}\right]\Gamma^{(n)}(p_i;t,h,\Lambda)=0\,. \tag{31.19}$$

We have assumed that the renormalization constants, and thus the RG functions defined by

$$\begin{aligned}\Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren. fixed}} t&=\beta(t)\,,\\ \Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren.fixed}}(-\ln Z)&=\zeta(t)\,,\\ \Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren. fixed}}\ln h&=\rho(t)\,,\end{aligned} \tag{31.20}$$

with

$$\mathcal{S}(\boldsymbol{\pi},h)=\frac{1}{T}\int \mathrm{d}^d x\left\{\frac{1}{2}\left[(\partial_\mu\boldsymbol{\pi}(x))^2+\frac{(\boldsymbol{\pi}\cdot\partial_\mu\boldsymbol{\pi}(x))^2}{1-\boldsymbol{\pi}^2(x)}\right]-h\sqrt{1-\boldsymbol{\pi}^2(x)}\right\}. \tag{31.12}$$

We immediately introduce a dimensionless coupling t:

$$T=t\,\Lambda^{2-d}. \tag{31.13}$$

The properties of this action have been studied, in slightly different notations, in Chapter 14. We here directly borrow the results. We have shown that it is renormalizable in two dimensions and characterized the form of the renormalized action. We recall that the presence of a non-vanishing magnetic field h is required, for $d\le 2$, 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. The renormalized action has the form

$$\mathcal{S}_\mathrm{r}(\boldsymbol{\pi}_\mathrm{r})=\frac{\mu^{d-2}Z}{2t_\mathrm{r}Z_t}\int \mathrm{d}^d x\left[(\partial_\mu\boldsymbol{\pi}_\mathrm{r})^2+(\partial_\mu\sigma_\mathrm{r})^2\right]-\frac{\mu^{d-2}}{t_\mathrm{r}}h_\mathrm{r}\int\sigma_\mathrm{r}(x)\mathrm{d}^d x\,, \tag{31.14}$$

in which μ is the renormalization scale, and

$$\sigma_\mathrm{r}(x)=\left[Z^{-1}-\boldsymbol{\pi}_\mathrm{r}^2\right]^{1/2}. \tag{31.15}$$

Note that the renormalization constants can be chosen h independent. This is automatically realized in the minimal subtraction scheme.

The relations

$$t=(\Lambda/\mu)^{d-2}Z_t t_\mathrm{r}\,,\quad \boldsymbol{\pi}_\mathrm{r}(x)=Z^{-1/2}\boldsymbol{\pi}(x) \tag{31.16}$$

imply

$$h=Z_h h_\mathrm{r}\,,\quad Z_h=Z_t/\sqrt{Z}\,. \tag{31.17}$$

With our conventions the coupling constant, which is proportional to the temperature, is dimensionless. The relation between the cut-off dependent and the renormalized correlation functions is

$$Z^{n/2}(\Lambda/\mu,t)\Gamma^{(n)}(p_i;t,h,\Lambda)=\Gamma_\mathrm{r}^{(n)}(p_i;t_\mathrm{r},h_\mathrm{r},\mu). \tag{31.18}$$

Differentiating with respect to Λ at renormalized parameters fixed, we obtain the RG equations:

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(t)\frac{\partial}{\partial t}-\frac{n}{2}\zeta(t)+\rho(t)h\frac{\partial}{\partial h}\right]\Gamma^{(n)}(p_i;t,h,\Lambda)=0\,. \tag{31.19}$$

We have assumed that the renormalization constants, and thus the RG functions defined by

$$\begin{aligned}\Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren. fixed}} t&=\beta(t)\,,\\ \Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren.fixed}}(-\ln Z)&=\zeta(t)\,,\\ \Lambda\frac{\partial}{\partial\Lambda}\bigg|_{\text{ren. fixed}}\ln h&=\rho(t)\,,\end{aligned} \tag{31.20}$$

have been chosen h independent. The coefficient of $\partial/\partial h$ can be derived from equation (31.17) which implies (taking the logarithm of both members)

$$0 = h^{-1}\Lambda\frac{\partial}{\partial\Lambda}h + d - 2 - \frac{1}{2}\zeta(t) - \frac{\beta(t)}{t}, \qquad (31.21)$$

and, therefore,

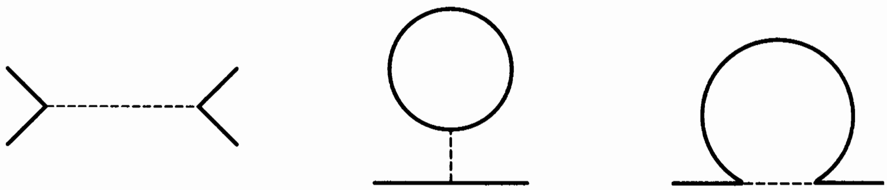
$$\rho(t) = 2 - d + \frac{1}{2}\zeta(t) + \frac{\beta(t)}{t}. \qquad (31.22)$$

To be able to discuss correlation functions involving the σ-field, we also need the RG equations satisfied by the connected correlation functions $W^{(n)}$:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(t)\frac{\partial}{\partial t} + \frac{n}{2}\zeta(t) + \left(\frac{1}{2}\zeta(t) + \frac{\beta(t)}{t} - \varepsilon\right)h\frac{\partial}{\partial h}\right] W^{(n)}(p_i; t, h, \Lambda) = 0\,, \qquad (31.23)$$

in which we now have set

$$d = 2 + \varepsilon\,. \qquad (31.24)$$



$V^{(4)} = \frac{1}{8}\delta_{i_1 i_2}\delta_{i_3 i_4}[(p_1+p_2)^2 + h]$ $\quad\frac{1}{2}(N-1)h\Omega_d(\sqrt{h})$ $\quad p^2\Omega_d(\sqrt{h})$

Fig. 31.1 One-loop diagrams: the dotted lines do not correspond to propagators but are used only to represent faithfully the flow of group indices.

In Section 15.6, the β-function has been calculated, at leading order, for all symmetric spaces among which the σ-model is one of the simplest examples. In this model, the two RG functions can be obtained at one-loop order from a calculation of the two-point function $\Gamma^{(2)}$ (figure 31.1):

$$\Gamma^{(2)}(p) = \frac{\Lambda^{\varepsilon}}{t}(p^2 + h) + \left[p^2 + \tfrac{1}{2}(N-1)h\right]\Omega_d(\sqrt{h}) + O(t)\,, \qquad (31.25)$$

where we have introduced (definition (30.12))

$$\Omega_d(\sqrt{h}) = \frac{1}{(2\pi)^d}\int^{\Lambda}\frac{\mathrm{d}^d q}{q^2 + h} \underset{\substack{\Lambda\to\infty \\ d\to 2_+}}{=} \Omega_d(0) + \frac{1}{(4\pi)^{d/2}}\Gamma(1 - d/2)h^{\varepsilon/2} + O\left(h\Lambda^{\varepsilon-2}\right) \qquad (31.26)$$

with

$$\Omega_d(0)\Lambda^{-\varepsilon} \sim -1/(2\pi\varepsilon).$$

Applying the RG equation (31.19) to $\Gamma^{(2)}$ and identifying the coefficients of p^2 and h, we derive two equations which determine $\beta(t)$ and $\zeta(t)$ at one-loop order:

$$\beta(t) = \varepsilon t - \frac{(N-2)}{2\pi}t^2 + O\left(t^3, t^2\varepsilon\right), \qquad (31.27a)$$

$$\zeta(t) = \frac{(N-1)}{2\pi}t + O\left(t^2, t\varepsilon\right). \qquad (31.27b)$$

31.3 RG Equations: Discussion

From the expression of $\beta(t)$ of equation (31.27a), we immediately conclude:

For $d \le 2$ ($\varepsilon \le 0$), $t = 0$ 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. Note that 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). Section 31.5 contains a short discussion of the physics in two dimensions for $N > 2$. The abelian case $N = 2$ is special and will be examined in Chapter 33.

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 first non-trivial zero of $\beta(t)$. Therefore, the large distance properties of the model can be obtained from the low temperature expansion and renormalization group, replacing the perturbative parameters by effective parameters obtained by solving the RG equations.

The critical temperature. Finally, we observe that, at least for ε positive and small, and $N > 2$, the RG function $\beta(t)$ has a non-trivial zero t_c:

$$t_c = \frac{2\pi\varepsilon}{N-2} + O\left(\varepsilon^2\right) \Rightarrow \beta(t_c) = 0\,, \text{ and } \quad \beta'(t_c) = -\varepsilon + O\left(\varepsilon^2\right). \tag{31.28}$$

Since t_c is an unstable IR fixed point, it is by definition a critical temperature. Consequences of this property are studied below. Let us only immediately note that t_c is also a UV fixed point, that is, 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.

We now discuss more precisely the solutions of the RG equations.

31.3.1 Integration of the RG equations: $d > 2$, $t < t_c$

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 (31.19) by introducing a scaling parameter λ and looking for a solution of the form

$$\Gamma^{(n)}(p_i, t, h, \Lambda) = Z^{-n/2}(\lambda)\Gamma^{(n)}\left(p_i, t(\lambda), h(\lambda), \lambda\Lambda\right). \tag{31.29}$$

Compatibility with equation (31.19) implies

$$\ln\lambda = \int_t^{t(\lambda)} \frac{\mathrm{d}t'}{\beta(t')}\,, \tag{31.30a}$$

$$\ln Z(\lambda) = \int_t^{t(\lambda)} \mathrm{d}t' \frac{\zeta(t')}{\beta(t')}\,, \tag{31.30b}$$

$$h(\lambda) = \lambda^{2-d} Z^{1/2}(\lambda)\frac{t(\lambda)}{t} h\,. \tag{31.30c}$$

With our conventions, $\Gamma^{(n)}$ has the dimension d and h dimension 2. Taking into account the dimensional analysis, we then rewrite relation (31.29) as

$$\Gamma^{(n)}(p_i, t, h, \Lambda) = Z^{-n/2}(\lambda)(\lambda\Lambda)^d \Gamma^{(n)}\bigl(p_i/\Lambda\lambda, t(\lambda), h(\lambda)/(\lambda\Lambda)^2, 1\bigr). \tag{31.31}$$

For $h \ll \Lambda^2$, perturbation theory has IR singularities. By choosing λ solution of the equation

$$h(\lambda) = (\lambda\Lambda)^2, \tag{31.32}$$

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 $\lambda \to 0$. We then introduce three functions of the temperature $M_0(t)$, $\xi(t)$ and $K(t)$:

$$M_0(t) = \exp\left[-\frac{1}{2}\int_0^t \frac{\zeta(t')}{\beta(t')}\mathrm{d}t'\right], \tag{31.33}$$

$$\xi(t) = \Lambda^{-1} t^{1/\varepsilon} \exp\left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{\varepsilon t'}\right)\mathrm{d}t'\right]. \tag{31.34}$$

$$K(t) = M_0(t)\left[\Lambda\xi(t)\right]^{d-2}/t = 1 + O(t). \tag{31.35}$$

Solving then (31.32), we find

$$\lambda \sim K^{1/2}(t) h^{1/2} \Lambda^{-1}. \tag{31.36}$$

Because $t = 0$ is an IR fixed point, the scale-dependent temperature $t(\lambda) \to 0$ and thus the leading terms in the small h and small momenta limit can be calculated perturbatively. Using equations (31.30a, b), we obtain the behaviours of $t(\lambda)$ and $Z(\lambda)$:

$$t(\lambda) \sim \lambda^{d-2} t K(t) M_0^{-1} \sim t\left[K(t)\right]^{d/2} M_0^{-1} h^{(d-2)/2} \Lambda^{2-d}, \tag{31.37}$$

$$Z(\lambda) \sim M_0^2(t). \tag{31.38}$$

It follows that

$$\Gamma^{(n)}(p_i, t, h, \Lambda) \sim M_0^{-n}(t)[K(t)h]^{d/2} \times \Gamma^{(n)}\left(\frac{p_i}{[K(t)h]^{1/2}}, \frac{t\left[K(t)\right]^{d/2}}{M_0(t)}\left(\frac{h}{\Lambda^2}\right)^{(d-2)/2}, 1, 1\right). \tag{31.39}$$

Actually, it is easy to verify directly, using dimensional analysis in the form

$$\left(\Lambda\frac{\partial}{\partial\Lambda} + 2h\frac{\partial}{\partial h} + p_i\frac{\partial}{\partial p_i}\right)\Gamma^{(n)} = d\Gamma^{(n)},$$

that equation (31.39) gives the general solution of equation (31.19).

Similarly, one finds

$$W^{(n)}(p_i; t, h, \Lambda) = [K(t)h]^{(1-n)d/2} M_0^n(t) \times W^{(n)}\left(\frac{p_i}{[K(t)h]^{1/2}}, \frac{t\left[K(t)\right]^{d/2}}{M_0(t)}\left(\frac{h}{\Lambda^2}\right)^{(d-2)/2}, 1, 1\right). \tag{31.40}$$

The coexistence curve. Let us apply this result 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 is given by

$$M(t,h,\Lambda) \equiv \langle \sigma(x) \rangle = W^{(1)}(t,h,\Lambda) = M_0(t) M\big(t(\lambda),1,1\big). \tag{31.41}$$

At one-loop order in a field it is given by

$$M = 1 - \frac{N-1}{2}\Lambda^{-\varepsilon} t \frac{1}{(2\pi)^d} \int^{\Lambda} \frac{\mathrm{d}^d q}{q^2+h} + O\left(t^2\right).$$

Thus, using the relation (31.41), we find

$$M(t,h,\Lambda=1) = M_0(t) - \frac{N-1}{2} t\,[K(t)]^{d/2}\, h^{(d-2)/2} \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} + O\left(h, h^{d-2}\right).$$

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 (29.23) and to the singularity for $x=-1$ found in equation (30.37).

31.3.2 Scaling form of correlation functions, the critical domain

It is convenient to rewrite equation (31.39) as

$$\Gamma^{(n)}(p_i;t,h,\Lambda) = \xi^{-d}(t) M_0^{-n}(t) F^{(n)}\left(p_i \xi(t), h/h_0(t)\right) \tag{31.42}$$

with

$$h_0(t) = t M_0^{-1}(t) \xi^{-d}(t) \Lambda^{2-d}. \tag{31.43}$$

For the connected correlation functions, the same analysis leads to

$$W^{(n)}(p_i;t,h,\Lambda) = \xi^{d(n-1)}(t) M_0^{n}(t) G^{(n)}\big(p_i \xi(t), h/h_0(t)\big). \tag{31.44}$$

The induced magnetization is

$$M = W^{(1)}(t,h,\Lambda) = M_0(t) G^{(1)}\big(h/h_0(t)\big). \tag{31.45}$$

Inversion of this relation yields the scaling form of the equation of state:

$$h = h_0(t) f\big(M/M_0(t)\big), \tag{31.46}$$

and the 1PI correlation functions can thus be written in terms of the magnetization as

$$\Gamma^{(n)}(p_i,t,M,\Lambda) = \xi^{-d}(t) M_0^{-n}(t) F^{(n)}\big(p_i \xi(t), M/M_0(t)\big). \tag{31.47}$$

Equations (31.46,31.47) are consistent with equations (26.63,26.72): the appearance of two different functions $\xi(t)$ and $M_0(t)$ corresponds to 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. There is, however, 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 (31.47) 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. We are, therefore, led to a remarkable conclusion: 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, and thus more generally at leading order in the critical domain. This conclusion is supported by the analysis of the scaling behaviour performed within the $1/N$ expansion (see equation (30.64)). 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. non too small) in the cut-off scale.

The critical domain: critical exponents. We now study more precisely the behaviour of functions for $t \to t_c$ (for $N > 2$). The function $\xi(t)$ diverges as

$$\xi(t) \sim \Lambda^{-1} \left(t_c - t\right)^{1/\beta'(t_c)} . \tag{31.48}$$

Comparing with the scaling form (26.35), we conclude that the correlation length exponent ν is given by

$$\nu = -\frac{1}{\beta'(t_c)} . \tag{31.49}$$

For d close to 2, the exponent ν thus behaves like

$$\nu \sim 1/\varepsilon . \tag{31.50}$$

The function $M_0(t)$ vanishes at t_c:

$$\ln M_0(t) = -\frac{1}{2}\frac{\zeta(t_c)}{\beta'(t_c)} \ln(t_c - t) + \text{ const.} \tag{31.51}$$

This yields the exponent β and thus also η through the scaling relation $\beta = \frac{1}{2}\nu(d-2+\eta)$:

$$\eta = \zeta(t_c) - \varepsilon . \tag{31.52}$$

A leading order, we find

$$\beta = \frac{N-1}{2(N-2)} + O(\varepsilon), \quad \eta = \frac{\varepsilon}{N-2} + O\left(\varepsilon^2\right). \tag{31.53}$$

We, finally, note that the singularity of $\Gamma^{(n)}$ coming from the prefactor $\xi^{-d} M_0^{-n}$ indeed agrees near t_c with the result of equation (26.37).

The nature of the correlation length $\xi(t)$. Equation (31.42) 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, long distance physics is governed by massless Goldstone modes (spin wave excitations) and can thus be deduced 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. There then exist distances 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 30.6). Moreover, the large N analysis indicates that RG equations in the form (31.19) 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 also the critical behaviour near two dimensions in the non-abelian case.

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. For $N > 2$, the sphere has instead 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 30.6). This gives us some confidence that the previous considerations are valid for the N-vector model at least for N sufficiently large. On the other hand, the physics for $N = 2$ is not well reproduced (see Chapter 33). 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 in their analysis.

31.4 Results Beyond One-Loop

For explicit calculations of the RG functions beyond one-loop, it is convenient to use the renormalized action (31.14), dimensional regularization and minimal subtraction. As an exercise, we calculate explicitly the RG functions at two-loop order.

Two-loop calculation. Note that the field renormalization constant can be easily obtained from the free energy or the σ-field expectation value. We again will obtain both RG functions from the two-point function. We now need the one-loop function with the renormalization constants:

$$\Gamma^{(2)}(p) = Z\left[\frac{\mu^{\varepsilon}}{tZ_t}(p^2 + hZ_h) + [p^2 + \frac{1}{2}(N-1)hZ_h]I(hZ_h)\right] + O(t) \tag{31.54}$$

(we have omitted the subscript r on all quantities). Here (equations (30.12,31.26)),

$$I(h) = \Omega_d(\sqrt{h}) = N_d\left(\frac{1}{2-d} - \frac{1}{2}\ln h\right) + O((d-2))\,, \tag{31.55}$$

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 30.6). Moreover, the large N analysis indicates that RG equations in the form (31.19) 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 also the critical behaviour near two dimensions in the non-abelian case.

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. For $N > 2$, the sphere has instead 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 30.6). This gives us some confidence that the previous considerations are valid for the N-vector model at least for N sufficiently large. On the other hand, the physics for $N = 2$ is not well reproduced (see Chapter 33). 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 in their analysis.

31.4 Results Beyond One-Loop

For explicit calculations of the RG functions beyond one-loop, it is convenient to use the renormalized action (31.14), dimensional regularization and minimal subtraction. As an exercise, we calculate explicitly the RG functions at two-loop order.

Two-loop calculation. Note that the field renormalization constant can be easily obtained from the free energy or the σ-field expectation value. We again will obtain both RG functions from the two-point function. We now need the one-loop function with the renormalization constants:

$$\Gamma^{(2)}(p) = Z\left[\frac{\mu^\varepsilon}{tZ_t}(p^2 + hZ_h) + [p^2 + \frac{1}{2}(N-1)hZ_h]I(hZ_h)\right] + O(t) \qquad (31.54)$$

(we have omitted the subscript r on all quantities). Here (equations (30.12,31.26)),

$$I(h) = \Omega_d(\sqrt{h}) = N_d\left(\frac{1}{2-d} - \frac{1}{2}\ln h\right) + O\big((d-2)\big)\,, \qquad (31.55)$$

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 the one-loop order we recover, as expected, the expressions (31.27).

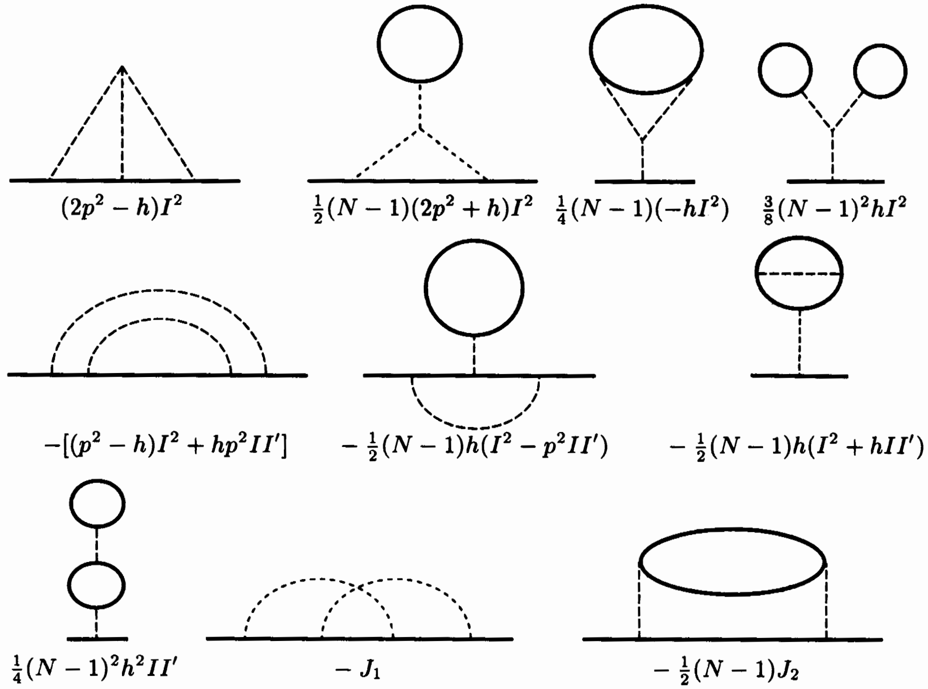


Fig. 31.2 Two-loop diagrams: faithful representation and contribution.

Figure 31.2 displays the two-loop diagrams with their natural weight factors. The notation I' means derivative with respect to h.

Note that we now also need the six-point vertex

$$V^{(6)} = \tfrac{1}{16}[(p_1+p_2)^2 + h]\delta_{i_1i_2}\delta_{i_3i_4}\delta_{i_5i_6}\,.$$

It generates four faithful two-loop diagrams, displayed in the first line of figure 31.2.

We now consider the diagrams involving two four-point vertices. One topology corresponds 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:

$$J_1 = \int \frac{\mathrm{d}^d q_1 \mathrm{d}^d q_2}{(2\pi)^{2d}} \frac{[(p+q_1)^2+h][(p+q_2)^2+h]}{(q_1^2+h)(q_2^2+h)[(p+q_1+q_2)^2+h]}\,,$$
$$J_2 = \int \frac{\mathrm{d}^d q_1 \mathrm{d}^d q_2}{(2\pi)^{2d}} \frac{[(p+q_1)^2+h]^2}{(q_1^2+h)(q_2^2+h)[(p+q_1+q_2)^2+h]}\,.$$

Although these contributions can no longer be reduced to one-loop diagrams, the divergent part can easily be extracted:

$$J_1 = 2(1-1/d)p^2I^2(h) + O(1), \quad J_2 = [(4/d-1)p^2 - h]I^2(h) + O(1).$$

It follows that the sum of the divergent terms of the two-loop contribution can be written as

$$\frac{N_d^2}{\varepsilon^2}\left[\left(\left(\tfrac{1}{8}N^2-\tfrac{1}{2}N+\tfrac{3}{8}\right)\varepsilon+\tfrac{3}{8}N^2-N+\tfrac{5}{8}\right)h+\left(\left(-\tfrac{7}{4}+\tfrac{3}{4}N\right)\varepsilon-\tfrac{1}{2}+\tfrac{1}{2}N\right)p^2\right].$$

After some straightforward algebra one finds in the minimal subtraction scheme ($d = 2+\varepsilon$):

$$Z = 1+(N-1)(\tilde{t}/\varepsilon)+(N-1)(N-\tfrac{3}{2})\left(\tilde{t}/\varepsilon\right)^2+O(\tilde{t}^3). \tag{31.56}$$

$$Z_t = 1+(N-2)(\tilde{t}/\varepsilon)+(N-2)(N-2+\tfrac{1}{2}\varepsilon)\left(\tilde{t}/\varepsilon\right)^2+O(\tilde{t}^3). \tag{31.57}$$

For convenience, we have rescaled the coupling constant t:

$$\tilde{t} = tN_d\,. \tag{31.58}$$

The RG functions in the renormalized theory follow from

$$\begin{aligned}\beta(t) &= \varepsilon t\left(1+t\frac{\mathrm{d}\ln Z_t}{\mathrm{d}t}\right)^{-1}, \\ \zeta(t) &= \beta(t)\frac{\mathrm{d}\ln Z}{\mathrm{d}t}.\end{aligned} \tag{31.59}$$

Higher orders. We now give the two RG functions and the critical exponents at the order presently available, that is, four loops, without presenting the details of the calculation. At this order $\tilde{\beta}$, the corresponding β-function, reads

$$\tilde{\beta}(\tilde{t}) = \varepsilon\tilde{t}-(N-2)\tilde{t}^2\left[1+\tilde{t}+\tfrac{1}{4}(N+2)\tilde{t}^2+b\tilde{t}^3\right]+O\left(\tilde{t}^6\right), \tag{31.60}$$

in which the quantity b is a constant:

$$b = -\tfrac{1}{12}(N^2-22N+34)+\tfrac{3}{2}\zeta(3)(N-3). \tag{31.61}$$

We recall that the value of the numerical constant $\zeta(3)$ (which should not be confused with the function $\zeta(t)$ of equation (31.27*b*)) is $\zeta(3) = 1.2020569\ldots$.

The anomalous dimensions of the composite operator of spin l (see Section 14.8) is given by

$$\zeta_l(\tilde{t}) = a\tilde{t}\left\{1+(N-2)\tilde{t}^2\left[\tfrac{3}{4}+\left(\tfrac{1}{3}(5-N)+\tfrac{1}{2}(1+a)\zeta(3)\right)\tilde{t}\right]\right\}+O\left(\tilde{t}^5\right) \tag{31.62}$$

with

$$a = -\tfrac{1}{2}(N+l-2)l\,. \tag{31.63}$$

The case $l=1$ corresponds to the field itself. The function $\zeta(t)$ given at one-loop order by equation (31.27*b*) 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

$$\tilde{\varepsilon} = \varepsilon/(N-2)\,, \tag{31.64}$$

the expansions read

$$\eta = \tilde{\varepsilon}+(N-1)\tilde{\varepsilon}^2\left\{-1+\frac{N}{2}\tilde{\varepsilon}+\left[-b+(N-2)\left(\frac{2-N}{3}+\frac{3-N}{4}\zeta(3)\right)\right]\tilde{\varepsilon}^2\right\}+O\left(\tilde{\varepsilon}^5\right),$$

$$\nu(d-2) = 1-\tilde{\varepsilon}+\frac{(4-N)}{2}\tilde{\varepsilon}^2+\frac{1}{4}\left[N^2-10N+18+18(3-N)\zeta(3)\right]\tilde{\varepsilon}^3+O\left(\tilde{\varepsilon}^4\right).$$

31.5 The Dimension 2

Dimension 2 is of pedagogical interest from the point of view of particle physics. The RG function $\beta(t)$ is, then,

$$\beta(t) = -\frac{(N-2)}{2\pi}t^2 + O\left(t^3\right). \tag{31.65}$$

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, the large momentum behaviour of correlation functions is entirely calculable from perturbation theory and RG arguments. There is, however, a counterpart, 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 QCD in four dimensions, the theory of strong interactions (see Chapters 34,35).

If we now define a function $\xi(t)$ by

$$\xi(t) = \mu^{-1}\exp\left[\int^t \frac{\mathrm{d}t'}{\beta(t')}\right], \tag{31.66}$$

we can integrate the RG equations in the same way and we find that $\xi(t)$ is the correlation length in zero field. In addition, we can use the explicit expression of the β-function (equation (31.60)) to calculate the correlation length or the physical mass for small t:

$$\xi^{-1}(t) = m(t) = K\mu t^{-1/(N-2)}\,\mathrm{e}^{-2\pi/[(N-2)t]}\left(1 + O(t)\right). \tag{31.67}$$

However, the exact value of the integration constant K, which gives the physical mass in the RG scale, can be calculated only by non-perturbative techniques.

Finally, the scaling forms (31.42,31.44) imply that the perturbative expansion at fixed magnetic field is valid, at low momenta or large distances, and for $h/h_0(t)$ large.

Elitzur's conjecture. The $O(N)$ symmetric action (31.8) ($h = 0$) has a sphere of degenerate classical minima. To define perturbation theory, we have been forced to add a linear symmetry breaking term to the action, which selects one particular classical minimum. The perturbative expansion of general correlation functions has IR divergences when the parameter h goes to zero for $d \le 2$, as one verifies on the explicit expression (31.26). This property is consistent with the absence of SSB for $d \le 2$. However, to calculate in perturbation theory another option is available, which will be used systematically in the case of instanton calculations (see Chapters 39–43): one does not introduce a symmetry breaking term but instead a set of *collective coordinates* which parametrizes the set of classical minima. One then expands in perturbation theory around one fixed minimum but treats perturbatively only the modes of the field that do not correspond to a global rotation. One finally sums over all classical minima. Clearly, after this last summation, only $O(N)$ invariant correlation functions survive. As already mentioned in Section 14.4, it has been conjectured by Elitzur and proven by David that in two dimensions, the $O(N)$ invariant correlation functions obtained by this procedure have a regular low temperature expansion: this means that if we calculate $O(N)$ invariant correlation functions by perturbation theory with a non-vanishing field and take the limit $h = 0$, this limit is IR finite. The subtlety of this problem when compared to the instanton case is that in the infinite volume the zero momentum singularity due to the choice of one classical minimum is not an isolated pole (see also Section 3.4 and Chapter 39).

31.6 Generalizations

We have shown in Chapter 15 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 (see Section 15.6), 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 is in general 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 non-infinitesimal. Let us mention that the β-function has been calculated up to four loops for a large class of symmetric spaces.

Symmetric spaces corresponding to non-compact groups. Previous considerations can be formally extended to symmetric spaces G/H in which the group G is non-compact, but the subgroup H remains compact. Symmetry groups obtained by complexification of compact group generate such spaces. A simple example is $O(M,N)/O(M)/O(N)$, in which $O(M,N)$ is a pseudo-orthogonal group. 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.

31.7 The Gross–Neveu Model

The GN model is described in terms of a $U(\tilde{N})$ symmetric action for a set of $\tilde{N}$ massless, self-interacting, Dirac fermions $\{\psi^i,\bar{\psi}^i\}$:

$$\mathcal{S}(\bar{\psi},\psi)=-\int \mathrm{d}^d x\left[\bar{\psi}\cdot\not{\partial}\psi+\tfrac{1}{2}G\left(\bar{\psi}\cdot\psi\right)^2\right].$$

The GN model has in even dimensions a discrete chiral symmetry:

$$\psi\mapsto\gamma_S\psi,\quad \bar{\psi}\mapsto-\bar{\psi}\gamma_S\,, \tag{31.68}$$

which prevents the addition of a fermion mass term while in odd dimensions a mass term breaks space parity. Actually, the two symmetry operations can be written in a form

$$\mathbf{x}=\{x_1,x_2,\ldots,x_d\}\mapsto\tilde{\mathbf{x}}=\{-x_1,x_2,\ldots,x_d\},\quad \begin{cases}\psi(x)\mapsto\gamma_1\psi(\tilde{x}),\\ \bar{\psi}(x)\mapsto-\bar{\psi}(\tilde{x})\gamma_1\end{cases},$$

valid in all dimensions.

The model illustrates the physics of spontaneous fermion mass generation and, in even dimensions, chiral symmetry breaking. It is renormalizable and asymptotically free in two dimensions. However, as in the case of 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 interchanged. Indeed, it is always the massless phase which is unstable in low dimensions.

31.6 Generalizations

We have shown in Chapter 15 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 (see Section 15.6), 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 is in general 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 non-infinitesimal. Let us mention that the β-function has been calculated up to four loops for a large class of symmetric spaces.

Symmetric spaces corresponding to non-compact groups. Previous considerations can be formally extended to symmetric spaces G/H in which the group G is non-compact, but the subgroup H remains compact. Symmetry groups obtained by complexification of compact group generate such spaces. A simple example is $O(M,N)/O(M)/O(N)$, in which $O(M,N)$ is a pseudo-orthogonal group. 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.

31.7 The Gross–Neveu Model

The GN model is described in terms of a $U(\tilde{N})$ symmetric action for a set of $\tilde{N}$ massless, self-interacting, Dirac fermions $\{\psi^i, \bar{\psi}^i\}$:

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}^d x \left[\bar{\psi}\cdot\not{\partial}\psi + \tfrac{1}{2}G\left(\bar{\psi}\cdot\psi\right)^2\right].$$

The GN model has in even dimensions a discrete chiral symmetry:

$$\psi \mapsto \gamma_S\psi, \quad \bar{\psi} \mapsto -\bar{\psi}\gamma_S\,, \tag{31.68}$$

which prevents the addition of a fermion mass term while in odd dimensions a mass term breaks space parity. Actually, the two symmetry operations can be written in a form

$$\mathbf{x} = \{x_1, x_2, \ldots, x_d\} \mapsto \tilde{\mathbf{x}} = \{-x_1, x_2, \ldots, x_d\}, \quad \begin{cases} \psi(x) \mapsto \gamma_1\psi(\tilde{x}), \\ \bar{\psi}(x) \mapsto -\bar{\psi}(\tilde{x})\gamma_1 \end{cases},$$

valid in all dimensions.

The model illustrates the physics of spontaneous fermion mass generation and, in even dimensions, chiral symmetry breaking. It is renormalizable and asymptotically free in two dimensions. However, as in the case of 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 interchanged. Indeed, it is always the massless phase which is unstable in low dimensions.

Since the symmetry breaking mechanism is non-perturbative it will eventually be instructive to compare the GN model with a different model with the same symmetries: the Gross–Neveu–Yukawa model.

Notation. In what follows we denote by N the total number of fermion components:

$$N = \tilde{N} \operatorname{tr} \mathbf{1}, \tag{31.69}$$

where $\mathbf{1}$ is the unit matrix in the representation of γ matrices.

31.7.1 RG equations near and in two dimensions

The GN model is renormalizable in two dimensions, and in perturbation theory describes only 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 breaks softly the chiral symmetry. It is possible to use dimensional regularization in practical calculations. Note that in two dimensions, the symmetry group is really $O(N)$ ($N = 2\tilde{N}$ 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 coupling below 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 appear at leading order and does not affect the analysis, we disregard here this subtlety.

It is convenient to introduce here a dimensionless coupling constant

$$u = G\Lambda^{2-d}. \tag{31.70}$$

As a function of the cut-off Λ, the bare correlation functions satisfy the RG equations:

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(u)\frac{\partial}{\partial u} - \frac{n}{2}\eta_\psi(u) - \eta_{\mathcal{M}}(u)\mathcal{M}\frac{\partial}{\partial \mathcal{M}}\right]\Gamma^{(n)}(p_i; u, \mathcal{M}, \Lambda) = 0\,. \tag{31.71}$$

A direct calculation of the β-function in $d = 2+\varepsilon$ dimension yields

$$\beta(u) = \varepsilon u - (N-2)\frac{u^2}{2\pi} + (N-2)\frac{u^3}{4\pi^2} + \frac{(N-2)(N-7)}{32\pi^3}u^4 + O\left(u^5\right), \tag{31.72}$$

Note that $N = 2\tilde{N}$ for $d = 2$.

The special case $N = 2$, for which the β-function vanishes identically in two dimensions, corresponds to the Thirring model (because for $N = 2$ $(\bar{\psi}\gamma_\mu\psi)^2 = -2(\bar{\psi}\psi)^2$). The latter model is to the equivalent the sine-Gordon or the $O(2)$ vector model.

Finally, the field and mass RG functions, at the order presently available, are

$$\begin{aligned} \eta_\psi(u) &= \frac{N-1}{8\pi^2}u^2 - \frac{(N-1)(N-2)}{32\pi^3}u^3 + \frac{(N-1)(N^2-7N+7)}{128\pi^4}u^4, \\ \eta_{\mathcal{M}}(u) &= \frac{N-1}{2\pi}u - \frac{N-1}{8\pi^2}u^2 - \frac{(2N-3)(N-2)}{32\pi^3}u^3 + O(u^4). \end{aligned} \tag{31.73}$$

As in the example of the non-linear σ-model, the solution of the RG equations (31.71) involves a length scale ξ of the type of a correlation length which is a RG invariant:

$$\xi^{-1}(u) \equiv \Lambda(u) \propto \Lambda \exp\left[-\int^u \frac{\mathrm{d}u'}{\beta(u')}\right]. \tag{31.74}$$

31.7.2 Discussion

Two dimensions. For $d = 2$, the model is asymptotically free. In the chiral theory ($\mathcal{M} = 0$) the spectrum, then, 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 excitation (as in a superfluid or superconductor). All masses are proportional to the mass parameter $\Lambda(u)$ which is a RG invariant. Its dependence in the coupling constant is given by equation (31.74):

$$\Lambda(u) \propto \Lambda u^{1/(N-2)}\, \mathrm{e}^{-2\pi/(N-2)u}\bigl(1 + O(u)\bigr). \tag{31.75}$$

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

$$m_n = \Lambda(u)\frac{N-2}{\pi}\sin\left(\frac{n\pi}{N-2}\right), \quad n = 1, 2 \ldots < N\,,\ N > 4\,,$$

To each mass value corresponds a representation of the $O(N)$ 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 particular, the ratio of the masses of the fundamental fermion and the lowest lying boson is

$$\frac{m_\sigma}{m_\psi} = 2\cos\left(\frac{\pi}{N-2}\right) = 2 + O(1/N^2). \tag{31.76}$$

The large N limit will be recovered in Section 31.9.1.

Note that the two first values of N are special, the model $N = 4$ is conjectured to be equivalent to two decoupled sine-Gordon models.

Dimension $d = 2 + \varepsilon$. As in the case of the σ-model, asymptotic freedom implies the existence of a non-trivial UV fixed point u_c, in $2 + \varepsilon$ dimension:

$$u_c = \frac{2\pi}{N-2}\varepsilon\left(1 - \frac{\varepsilon}{N-2}\right) + O\left(\varepsilon^3\right).$$

u_c is also the critical coupling constant 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 finds the correlation length exponent ν:

$$\nu^{-1} = -\beta'(u_c) = \varepsilon - \frac{\varepsilon^2}{N-2} + O\left(\varepsilon^3\right). \tag{31.77}$$

The fermion field dimension $[\psi]$ is

$$2[\psi] = d - 1 + \eta_\psi(u_c) = 1 + \varepsilon + \frac{N-1}{2(N-2)^2}\varepsilon^2 + O\left(\varepsilon^3\right). \tag{31.78}$$

The dimension of the composite field $\sigma = \bar\psi\psi$ is given by

$$[\sigma] = d - 1 - \eta_{\mathcal{M}}(u_c) = 1 - \frac{\varepsilon}{N-2}\,.$$

31.7.2 Discussion

Two dimensions. For $d = 2$, the model is asymptotically free. In the chiral theory ($\mathcal{M} = 0$) the spectrum, then, 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 excitation (as in a superfluid or superconductor). All masses are proportional to the mass parameter $\Lambda(u)$ which is a RG invariant. Its dependence in the coupling constant is given by equation (31.74):

$$\Lambda(u) \propto \Lambda u^{1/(N-2)} \, \mathrm{e}^{-2\pi/(N-2)u} \big(1 + O(u)\big) . \qquad (31.75)$$

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

$$m_n = \Lambda(u) \frac{N-2}{\pi} \sin\left(\frac{n\pi}{N-2}\right), \qquad n = 1, 2 \ldots < N \,, \ N > 4 \,,$$

To each mass value corresponds a representation of the $O(N)$ 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 particular, the ratio of the masses of the fundamental fermion and the lowest lying boson is

$$\frac{m_\sigma}{m_\psi} = 2\cos\left(\frac{\pi}{N-2}\right) = 2 + O(1/N^2). \qquad (31.76)$$

The large N limit will be recovered in Section 31.9.1.

Note that the two first values of N are special, the model $N = 4$ is conjectured to be equivalent to two decoupled sine-Gordon models.

Dimension $d = 2 + \varepsilon$. As in the case of the σ-model, asymptotic freedom implies the existence of a non-trivial UV fixed point u_c, in $2 + \varepsilon$ dimension:

$$u_c = \frac{2\pi}{N-2} \varepsilon \left(1 - \frac{\varepsilon}{N-2}\right) + O\left(\varepsilon^3\right) .$$

u_c is also the critical coupling constant 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 finds the correlation length exponent ν:

$$\nu^{-1} = -\beta'(u_c) = \varepsilon - \frac{\varepsilon^2}{N-2} + O\left(\varepsilon^3\right) . \qquad (31.77)$$

The fermion field dimension $[\psi]$ is

$$2[\psi] = d - 1 + \eta_\psi(u_c) = 1 + \varepsilon + \frac{N-1}{2(N-2)^2} \varepsilon^2 + O\left(\varepsilon^3\right) . \qquad (31.78)$$

The dimension of the composite field $\sigma = \bar\psi\psi$ is given by

$$[\sigma] = d - 1 - \eta_{\mathcal{M}}(u_c) = 1 - \frac{\varepsilon}{N-2} .$$

As in the σ-model the existence of a non-trivial UV fixed point implies that large momentum behaviour is not given by perturbation theory above two dimensions, and this explains why the perturbative result that the model cannot be defined in higher dimensions cannot be trusted. In the massless phase, the length scale (31.74) 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 ε expansion makes sense beyond an infinitesimal neighbourhood of dimension 2, other methods are required, like the $1/N$ expansion which will be considered in Section 31.9.1.

31.8 The Gross–Neveu–Yukawa Model

The GNY model has the same chiral and $U(\tilde{N})$ symmetries as the GN model. The action is ($\varepsilon = 4 - d$)

$$\mathcal{S}(\bar{\psi}, \psi, \sigma) = \int \mathrm{d}^d x \left[-\bar{\psi} \cdot \left(\not{\partial} + g\Lambda^{\varepsilon/2}\sigma \right) \psi + \frac{1}{2} (\partial_\mu \sigma)^2 + \frac{1}{2} m^2 \sigma^2 + \frac{\lambda}{4!} \Lambda^\varepsilon \sigma^4 \right], \quad (31.79)$$

where σ is an additional scalar field, Λ the momentum cut-off, and g, λ dimensionless "bare", that is, effective coupling constants at large momentum scale Λ.

The action still has a reflection symmetry, σ transforming into $-\sigma$ when the fermions transform by (31.68). In contrast with the GN model, however, the chiral transition can here be discussed by perturbative methods. An analogous situation has already been encountered when comparing the $(\phi^2)^2$ field theory with the non-linear σ-model. Even more, the GN model is renormalizable in dimension 2 and the GNY model in dimension 4.

The phase transition. Examining the action (31.79), we see that in the tree approximation when m^2 is negative the chiral symmetry is spontaneously broken. The σ expectation value gives a mass to the fermions, a mechanism reminiscent of the Standard Model of weak-electromagnetic interactions:

$$m_\psi = g \langle \sigma \rangle , \quad (31.80)$$

while the σ mass then is

$$m_\sigma^2 = \frac{\lambda}{3g^2} m_\psi^2 . \quad (31.81)$$

As a result of interactions, the transition value m_c^2 of the parameter m^2 will be modified. In what follows we set

$$m^2 = m_c^2 + t , \quad (31.82)$$

where the new parameter t, in the language of phase transitions, plays the role of the deviation from the critical temperature.

To study the model beyond the tree approximation, we now discuss RG equations near four dimensions.

As in the σ-model the existence of a non-trivial UV fixed point implies that large momentum behaviour is not given by perturbation theory above two dimensions, and this explains why the perturbative result that the model cannot be defined in higher dimensions cannot be trusted. In the massless phase, the length scale (31.74) 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 ε expansion makes sense beyond an infinitesimal neighbourhood of dimension 2, other methods are required, like the $1/N$ expansion which will be considered in Section 31.9.1.

31.8 The Gross–Neveu–Yukawa Model

The GNY model has the same chiral and $U(\tilde{N})$ symmetries as the GN model. The action is ($\varepsilon = 4 - d$)

$$\mathcal{S}(\bar{\psi}, \psi, \sigma) = \int \mathrm{d}^d x \left[-\bar{\psi} \cdot \left(\not{\partial} + g\Lambda^{\varepsilon/2}\sigma \right) \psi + \frac{1}{2} \left(\partial_\mu \sigma \right)^2 + \frac{1}{2} m^2 \sigma^2 + \frac{\lambda}{4!} \Lambda^\varepsilon \sigma^4 \right], \quad (31.79)$$

where σ is an additional scalar field, Λ the momentum cut-off, and g, λ dimensionless "bare", that is, effective coupling constants at large momentum scale Λ.

The action still has a reflection symmetry, σ transforming into $-\sigma$ when the fermions transform by (31.68). In contrast with the GN model, however, the chiral transition can here be discussed by perturbative methods. An analogous situation has already been encountered when comparing the $(\phi^2)^2$ field theory with the non-linear σ-model. Even more, the GN model is renormalizable in dimension 2 and the GNY model in dimension 4.

The phase transition. Examining the action (31.79), we see that in the tree approximation when m^2 is negative the chiral symmetry is spontaneously broken. The σ expectation value gives a mass to the fermions, a mechanism reminiscent of the Standard Model of weak-electromagnetic interactions:

$$m_\psi = g \left\langle \sigma \right\rangle , \quad (31.80)$$

while the σ mass then is

$$m_\sigma^2 = \frac{\lambda}{3g^2} m_\psi^2 \, . \quad (31.81)$$

As a result of interactions, the transition value m_c^2 of the parameter m^2 will be modified. In what follows we set

$$m^2 = m_c^2 + t \, , \quad (31.82)$$

where the new parameter t, in the language of phase transitions, plays the role of the deviation from the critical temperature.

To study the model beyond the tree approximation, we now discuss RG equations near four dimensions.

31.8.1 RG equations near four dimensions

The model (31.79) is trivial above four dimensions, renormalizable in four dimensions and can thus be studied near dimension 4 by RG techniques. Five renormalization constants are required, corresponding to the two field renormalizations, the σ mass, and the two coupling constants. The RG equations thus involve five RG functions. The 1PI correlation functions $\Gamma^{(l,n)}$, for l ψ and n σ fields, then satisfy

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+\beta_{g^2}\frac{\partial}{\partial g^2}+\beta_\lambda\frac{\partial}{\partial\lambda}-\frac{1}{2}l\eta_\psi-\frac{1}{2}n\eta_\sigma-\eta_m t\frac{\partial}{\partial t}\right)\Gamma^{(l,n)}=0\,. \tag{31.83}$$

The RG functions at one-loop order have been calculated in Section 11.7:

$$\beta_\lambda=-\varepsilon\lambda+\frac{1}{8\pi^2}\left(\frac{3}{2}\lambda^2+N\lambda g^2-6Ng^4\right), \tag{31.84}$$

$$\beta_{g^2}=-\varepsilon g^2+\frac{N+6}{16\pi^2}g^4. \tag{31.85}$$

In four dimensions, $\operatorname{tr}\mathbf{1}=4$ and, therefore, $N=4\tilde{N}$.

Solution in four dimensions. In four dimensions, the origin $\lambda=g^2=0$ is IR stable. Indeed, the second equation implies that g goes to zero, and the first then that λ also goes to zero. As a consequence, if the bare coupling constants are generic, that is, if the effective couplings at cut-off scale are of order 1, the effective couplings at scale $\mu\ll\Lambda$ go to zero and in a way asymptotically independent from the bare couplings. One finds

$$g^2(\mu)\sim\frac{16\pi^2}{(N+6)\ln(\Lambda/\mu)}\,,\qquad \lambda(\mu)\sim\frac{8\pi^2\tilde{\lambda}^*}{\ln(\Lambda/\mu)}\,,$$

where we have defined

$$\tilde{\lambda}_*=\frac{48N}{(N+6)\left[(N-6)+2\sqrt{N^2+132N+36}\right]}. \tag{31.86}$$

This result allows us to use renormalized perturbation theory to calculate physical observables. For example, we can evaluate the ratio between the masses of the scalar and fermion fields. It is then optimal to take for μ a value of order $\langle\sigma\rangle$. A remarkable consequence follows: the ratio (31.81) of scalar and fermion masses is fixed:

$$\frac{m_\sigma^2}{m_\psi^2}=\frac{\lambda_*}{3g_*^2}=\frac{8N}{N-6+\sqrt{N^2+132N+36}}, \tag{31.87}$$

while in the classical limit it is arbitrary.

Of course, if the bare couplings are "unnaturally" small, the same will apply to the renormalized couplings at scale μ and the ratio will be modified.

Dimension $d=4-\varepsilon$. One then finds a non-trivial IR fixed point:

$$g_*^2=\frac{16\pi^2\varepsilon}{N+6},\quad \lambda_*=8\pi^2\varepsilon\tilde{\lambda}_*\,. \tag{31.88}$$

The matrix of derivatives of the β-functions has two eigenvalues ω,ω',

$$\omega_1=\varepsilon,\quad \omega_2=\varepsilon\sqrt{N^2+132N+36}/(N+6), \tag{31.89}$$

and thus the fixed point is IR stable. The first eigenvalue is always the smallest.

The field renormalization RG functions are at the same order:

$$\eta_\sigma = \frac{N}{16\pi^2} g^2, \qquad \eta_\psi = \frac{1}{16\pi^2} g^2. \tag{31.90}$$

At the fixed point, one finds

$$\eta_\sigma = \frac{N\varepsilon}{N+6}, \quad \eta_\psi = \frac{\varepsilon}{(N+6)}, \tag{31.91}$$

and thus the dimensions d_ψ and d_σ of the fields

$$d_\psi = \frac{3}{2} - \frac{N+4}{2(N+6)}\varepsilon, \qquad d_\sigma = 1 - \frac{3}{N+6}\varepsilon. \tag{31.92}$$

The RG function η_m corresponding to the mass operator is at one-loop order:

$$\eta_m = -\frac{\lambda}{16\pi^2} - \eta_\sigma,$$

and thus the exponent ν,

$$\frac{1}{\nu} = 2 + \eta_m = 2 - \frac{\varepsilon}{2}\tilde{\lambda}_* - \frac{N\varepsilon}{N+6} = 2 - \varepsilon \frac{5N + 6 + \sqrt{N^2 + 132N + 36}}{6(N+6)}. \tag{31.93}$$

Finally, the ratio of masses (31.81) at the fixed point keeps its four-dimensional value (31.87) at this leading order.

In $d = 4$ and $d = 4 - \varepsilon$, the existence of an IR fixed point has the same consequence: if we assume 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.*

31.9 GNY and GN Models in the Large N Limit

We now show that the GN model plays with respect to the GNY model (31.79), the role the non-linear σ-model plays with respect to the $(\phi^2)^2$ field theory. We consider below the interesting dimensions $2 \le d \le 4$.

31.9.1 The GNY model in the large N limit

We start from the action (31.79) of the GNY model and integrate over $\tilde{N} - 1$ fermion fields. We also rescale for convenience $\Lambda^{(4-d)/2} g\sigma$ into σ, and then get the large N action:

$$\begin{aligned} \mathcal{S}_N(\bar\psi, \psi, \sigma) = \int \mathrm{d}^d x \left\{ -\bar\psi\left(\not{\partial} + \sigma\right)\psi + \Lambda^{d-4}\left[\frac{1}{2g^2}\left(\partial_\mu \sigma\right)^2 + \frac{m^2}{2g^2}\sigma^2 + \frac{\lambda}{4! g^4}\sigma^4\right]\right\} \\ - (\tilde{N} - 1)\operatorname{tr}\ln\left(\not{\partial} + \sigma\right). \end{aligned} \tag{31.94}$$

To take the large N limit, we assume σ finite and $g^2, \lambda = O(1/N)$.

The action density $\mathcal{E}(M)$ for constant field $\sigma(x) = M$ and vanishing fermion fields follows

$$\begin{aligned}\mathcal{E}(M) &= \Lambda^{d-4}\left(\frac{m^2}{2g^2}M^2 + \frac{\lambda}{4!g^4}M^4\right) - \tilde{N}\,\mathrm{tr}\ln(\not{\partial} + M) \\ &= \Lambda^{d-4}\left(\frac{m^2}{2g^2}M^2 + \frac{\lambda}{4!g^4}M^4\right) - \frac{N}{2}\int^{\Lambda}\frac{\mathrm{d}^d q}{(2\pi)^d}\ln(q^2 + M^2). \end{aligned} \tag{31.95}$$

The expectation value of σ for N large is given by a *gap* equation:

$$\mathcal{E}'(M)\Lambda^{4-d} = \frac{m^2}{g^2}M + \frac{\lambda}{6g^4}M^3 - N\Lambda^{4-d}M\Omega_d(M) = 0\,, \tag{31.96}$$

where Ω_d is the one-loop diagram (30.12,30.13):

$$\Omega_d(M) = \frac{1}{(2\pi)^d}\int^{\Lambda}\frac{\mathrm{d}^d p}{p^2 + M^2} \underset{d>2}{=} \Omega_d(0) - C(d)m^{d-2} + a(d)m^2\Lambda^{d-4} + O\left(m^d/\Lambda^2, m^4\Lambda^{d-6}\right).$$

It is also useful to calculate the second derivative to check stability of the extrema

$$\mathcal{E}''(M)\Lambda^{4-d} = \frac{m^2}{g^2} + \frac{\lambda}{2g^4}M^2 + N\Lambda^{4-d}\int^{\Lambda}\frac{\mathrm{d}^d q}{(2\pi)^d}\frac{M^2 - q^2}{(q^2 + M^2)^2}.$$

The solution $M = 0$ is stable provided

$$\mathcal{E}''(0) > 0 \;\Leftrightarrow\; m^2/g^2 > N\Lambda^{4-d}\Omega_d(0).$$

Instead, the non-trivial solution to the gap equation exists only for

$$m^2/g^2 > N\Lambda^{4-d}\Omega_d(0),$$

but then it is stable. We conclude that the critical temperature or critical bare mass is given by

$$m_c^2/g^2 = N\Lambda^{4-d}\Omega_d(0), \tag{31.97}$$

which shows that the fermions favour the chiral transition. In particular, when d approaches 2, we observe that $m_c^2 \to +\infty$ which implies that the chiral symmetry is always broken in two dimensions. Using equation (31.97), and setting

$$r = \Lambda^{d-4}(m^2 - m_c^2)/g^2, \tag{31.98}$$

we can write the equation for the non-trivial solution as

$$r + \Lambda^{d-4}\lambda M^2/6g^4 - N\left(\Omega_d(M) - \Omega_d(0)\right) = 0\,.$$

Keeping only the leading terms for $r \to 0$, we obtain for $d < 4$ the scaling behaviour

$$M \sim (-r/NC)^{1/(d-2)}. \tag{31.99}$$

Since, at leading order, the fermion mass $m_\psi = M$, it immediately follows that the exponent ν is also given by

$$\nu \sim \beta \sim 1/(d-2) \;\Rightarrow\; \eta_\sigma = 4 - d\,. \tag{31.100}$$

At leading order, for $N \to \infty$, ν has the same value as in the non-linear σ-model.

At leading order in the scaling limit the thermodynamic (or effective) potential $\mathcal{E}(M)$ then becomes

$$\mathcal{E}(M) = \tfrac{1}{2} r M^2 + (N/d) C(d) |M|^d . \tag{31.101}$$

We note that, 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.

We read from the large N action, that at this order $\eta_\psi = 0$.

Finally from the large N action, we can calculate the σ-propagator at leading order. Quite generally, using the saddle point equation, one finds for the inverse σ-propagator in the massive phase:

$$\begin{aligned}\Delta_\sigma^{-1}(p) = \Lambda^{d-4}&\left(\frac{p^2}{g^2} + \frac{\lambda}{3g^4} M^2\right) \\ &+ \frac{N}{2(2\pi)^d}\left(p^2 + 4M^2\right)\int^\Lambda \frac{\mathrm{d}^d q}{(q^2+M^2)\left[(p+q)^2 + M^2\right]} .\end{aligned} \tag{31.102}$$

We see that in the scaling limit $p, M \to 0$, the integral yields the leading contribution. Neglecting corrections to scaling, we find that the propagator vanishes for $p^2 = -4M^2$ which is just the $\bar\psi\psi$ threshold. Thus, in this limit, $m_\sigma = 2m_\psi$ in all dimensions, a result consistent with $d = 2$ exact value.

At the transition, the propagator reduces to

$$\Delta_\sigma \sim \frac{2}{N b(d) p^{d-2}} \tag{31.103}$$

with (equation (30.46))

$$b(d) = -\frac{\pi}{\sin(\pi d/2)} \frac{\Gamma^2(d/2)}{\Gamma(d-1)} N_d . \tag{31.104}$$

The result is consistent with the value of η_σ found above.

Let us finally note that the behaviour of the propagator at the critical point, $\Delta_\sigma(p) \propto p^{2-d}$, implies for the field σ the canonical dimension $[\sigma]$ in the large N expansion, for $2 \le d \le 4$:

$$[\sigma] = 1 . \tag{31.105}$$

Corrections to scaling and the IR fixed point. The IR fixed point is determined by demanding the cancellation of the leading corrections to scaling. Let us thus consider the effective potential $\mathcal{E}(M)$. The leading correction to scaling is proportional to

$$\left(\frac{\lambda}{4!g^4} - \frac{Na(d)}{4}\right) M^4 ,$$

($a(\varepsilon) \sim 1/8\pi^2\varepsilon$). Demanding the cancellation of the coefficient of M^2, we obtain a relation between λ and g^2:

$$g_*^4 = \frac{\lambda_*}{6Na(d)} = \frac{4\lambda_*\varepsilon\pi^2}{3N} + O\left(\varepsilon^2\right),$$

a result consistent with the results of the ε-expansion.

At leading order, for $N \to \infty$, ν has the same value as in the non-linear σ-model.

At leading order in the scaling limit the thermodynamic (or effective) potential $\mathcal{E}(M)$ then becomes

$$\mathcal{E}(M) = \tfrac{1}{2} r M^2 + (N/d) C(d) |M|^d . \tag{31.101}$$

We note that, 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.

We read from the large N action, that at this order $\eta_\psi = 0$.

Finally from the large N action, we can calculate the σ-propagator at leading order. Quite generally, using the saddle point equation, one finds for the inverse σ-propagator in the massive phase:

$$\begin{aligned}\Delta_\sigma^{-1}(p) = \Lambda^{d-4}\left(\frac{p^2}{g^2} + \frac{\lambda}{3g^4} M^2\right) \\ + \frac{N}{2(2\pi)^d}\left(p^2 + 4M^2\right)\int^\Lambda \frac{\mathrm{d}^d q}{(q^2+M^2)\left[(p+q)^2+M^2\right]} .\end{aligned} \tag{31.102}$$

We see that in the scaling limit $p, M \to 0$, the integral yields the leading contribution. Neglecting corrections to scaling, we find that the propagator vanishes for $p^2 = -4M^2$ which is just the $\bar\psi\psi$ threshold. Thus, in this limit, $m_\sigma = 2m_\psi$ in all dimensions, a result consistent with $d = 2$ exact value.

At the transition, the propagator reduces to

$$\Delta_\sigma \sim \frac{2}{N b(d) p^{d-2}} \tag{31.103}$$

with (equation (30.46))

$$b(d) = -\frac{\pi}{\sin(\pi d/2)} \frac{\Gamma^2(d/2)}{\Gamma(d-1)} N_d . \tag{31.104}$$

The result is consistent with the value of η_σ found above.

Let us finally note that the behaviour of the propagator at the critical point, $\Delta_\sigma(p) \propto p^{2-d}$, implies for the field σ the canonical dimension $[\sigma]$ in the large N expansion, for $2 \le d \le 4$:

$$[\sigma] = 1 . \tag{31.105}$$

Corrections to scaling and the IR fixed point. The IR fixed point is determined by demanding the cancellation of the leading corrections to scaling. Let us thus consider the effective potential $\mathcal{E}(M)$. The leading correction to scaling is proportional to

$$\left(\frac{\lambda}{4! g^4} - \frac{N a(d)}{4}\right) M^4 ,$$

($a(\varepsilon) \sim 1/8\pi^2\varepsilon$). Demanding the cancellation of the coefficient of M^2, we obtain a relation between λ and g^2:

$$g_*^4 = \frac{\lambda_*}{6 N a(d)} = \frac{4\lambda_* \varepsilon \pi^2}{3N} + O\left(\varepsilon^2\right),$$

a result consistent with the results of the ε-expansion.

In the same way, it is possible to calculate the leading correction to the σ-propagator (31.102). Demanding the cancellation of the leading correction, we obtain

$$\frac{p^2}{g_*^2} + \frac{\lambda_*}{3g_*^4}M^2 - \frac{1}{2}N\left(p^2 + 4M^2\right)a(d) = 0\,.$$

The coefficient of M^2 cancels from the previous relation and the cancellation the coefficient of p^2 yields

$$g_*^2 = \frac{2}{Na(d)} = \frac{16\pi^2\varepsilon}{N} + O\left(\varepsilon^2\right)$$

in agreement with the ε-expansion for N large.

31.9.2 GN and GNY models

We have seen that the terms $(\partial_\mu\sigma)^2$ and σ^4 of the large N action which have a canonical dimension 4, are irrelevant in the IR critical region for $d \le 4$. We recognize a situation already encountered in the $(\phi^2)^2$ field theory in the large N limit. In the scaling region, it is possible to omit them and one then finds the action:

$$\mathcal{S}_N(\bar\psi,\psi,\sigma) = \int \mathrm{d}^d x \left[-\bar\psi\cdot(\not\partial+\sigma)\,\psi + \Lambda^{d-4}\frac{m^2}{2g^2}\sigma^2\right]. \qquad (31.106)$$

The integral over the σ field can explicitly be performed and yields the action of the GN model:

$$\mathcal{S}_N(\bar\psi,\psi) = -\int \mathrm{d}^d x \left[\bar\psi\cdot\not\partial\psi + \frac{\Lambda^{4-d}}{2m^2}g^2\left(\bar\psi\cdot\psi\right)^2\right].$$

The GN and GNY models are thus equivalent for the large distance physics. In the GN model, in the large N limit, the σ particle appears as a $\bar\psi\psi$ boundstate at threshold.

Conversely, it would seem that the GN model depends on a smaller number of parameters than its renormalizable extension. Again, this problem is interesting only in four dimensions where corrections to scaling, that is, to free field theory, are important. However, if we examine the divergences of the term $\operatorname{tr}\ln(\not\partial+\sigma)$ in the effective action (31.94) relevant for the large N limit, we find a local polynomial in σ of the form

$$\int \mathrm{d}^4 x\left[A\sigma^2(x) + B\left(\partial_\mu\sigma\right)^2 + C\sigma^4(x)\right].$$

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, we change the amplitude of higher dimension operators. These bare operators in the IR limit have a component on all lower dimensional renormalized operators.

Note finally 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 of the ferromagnets.

In the same way, it is possible to calculate the leading correction to the σ-propagator (31.102). Demanding the cancellation of the leading correction, we obtain

$$\frac{p^2}{g_*^2} + \frac{\lambda_*}{3g_*^4}M^2 - \frac{1}{2}N\left(p^2 + 4M^2\right)a(d) = 0\,.$$

The coefficient of M^2 cancels from the previous relation and the cancellation the coefficient of p^2 yields

$$g_*^2 = \frac{2}{Na(d)} = \frac{16\pi^2\varepsilon}{N} + O\left(\varepsilon^2\right)$$

in agreement with the ε-expansion for N large.

31.9.2 GN and GNY models

We have seen that the terms $(\partial_\mu\sigma)^2$ and σ^4 of the large N action which have a canonical dimension 4, are irrelevant in the IR critical region for $d \leq 4$. We recognize a situation already encountered in the $(\phi^2)^2$ field theory in the large N limit. In the scaling region, it is possible to omit them and one then finds the action:

$$\mathcal{S}_N(\bar{\psi},\psi,\sigma) = \int \mathrm{d}^d x \left[-\bar{\psi}\cdot(\not{\partial} + \sigma)\,\psi + \Lambda^{d-4}\frac{m^2}{2g^2}\sigma^2\right]. \tag{31.106}$$

The integral over the σ field can explicitly be performed and yields the action of the GN model:

$$\mathcal{S}_N(\bar{\psi},\psi) = -\int \mathrm{d}^d x \left[\bar{\psi}\cdot\not{\partial}\psi + \frac{\Lambda^{4-d}}{2m^2}g^2\left(\bar{\psi}\cdot\psi\right)^2\right].$$

The GN and GNY models are thus equivalent for the large distance physics. In the GN model, in the large N limit, the σ particle appears as a $\bar{\psi}\psi$ boundstate at threshold.

Conversely, it would seem that the GN model depends on a smaller number of parameters than its renormalizable extension. Again, this problem is interesting only in four dimensions where corrections to scaling, that is, to free field theory, are important. However, if we examine the divergences of the term $\operatorname{tr}\ln(\not{\partial} + \sigma)$ in the effective action (31.94) relevant for the large N limit, we find a local polynomial in σ of the form

$$\int \mathrm{d}^4 x \left[A\sigma^2(x) + B\left(\partial_\mu\sigma\right)^2 + C\sigma^4(x)\right].$$

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, we change the amplitude of higher dimension operators. These bare operators in the IR limit have a component on all lower dimensional renormalized operators.

Note finally 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 of the ferromagnets.

31.10 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 30.7.

Alternative theory. To prove that the large N expansion is renormalizable one proceeds as in the case of the scalar theory in Section 30.7. One starts from a critical action with an additional term quadratic in σ which generates the large N σ-propagator already in perturbation theory:

$$\mathcal{S}(\psi,\bar{\psi},\sigma)=\int \mathrm{d}^d x\left[-\bar{\psi}(\not{\partial}+\sigma)\psi+\frac{1}{2v^2}\sigma(-\partial^2)^{d/2-1}\sigma\right]. \tag{31.107}$$

The initial theory is recovered in the limit $v\to\infty$. One then rescales σ in $v\sigma$. The model is renormalizable without σ field renormalization because divergences generate only local counter-terms:

$$\mathcal{S}_{\mathrm{r}}(\psi,\bar{\psi},\sigma)=\int \mathrm{d}^d x\left[-Z_\psi\bar{\psi}(\not{\partial}+v_{\mathrm{r}}Z_v\sigma)\psi+\frac{1}{2}\sigma(-\partial^2)^{d/2-1}\sigma\right]. \tag{31.108}$$

RG equations follow

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta_{v^2}(v)\frac{\partial}{\partial v^2}-\frac{n}{2}\eta_\psi(v)\right]\Gamma^{(l,n)}=0\,. \tag{31.109}$$

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 large N σ propagator reads

$$\langle\sigma\sigma\rangle=\frac{2}{b(d)R(v)p^{d-2}}. \tag{31.110}$$

The solution to the RG equations can be written as

$$\Gamma^{(l,n)}(\ell p,v,\Lambda)=Z^{-n/2}(\ell)\ell^{d-l-n(d-2)/2}\Gamma^{(l,n)}(p,v(\ell),\Lambda) \tag{31.111}$$

with the usual definitions

$$\ell\frac{\mathrm{d}v^2}{\mathrm{d}\ell}=\beta(v(\ell))\,,\quad \ell\frac{\mathrm{d}\ln Z}{\mathrm{d}\ell}=\eta_\psi(v(\ell))\,.$$

We are interested in the neighbourhood of the fixed point $v^2=\infty$. Then, the RG function $\eta(v)$ approaches the exponent η. The flow equation for the coupling constant becomes

$$\ell\frac{\mathrm{d}v^2}{\mathrm{d}\ell}=\rho v^2,\ \Rightarrow\ v^2(\ell)\sim\ell^\rho.$$

We again note that a correlation function with l σ fields becomes proportional to v^l. Therefore,

$$\Gamma^{(l,n)}(\ell p,v,\Lambda)\propto\ell^{d-(1-\rho/2)l-n(d-2+\eta_\psi)/2}. \tag{31.112}$$

We conclude

$$d_\sigma = \tfrac{1}{2}(d-2+\eta_\sigma) = 1 - \tfrac{1}{2}\rho \ \Leftrightarrow\ \eta_\sigma = 4-d-\rho\,. \tag{31.113}$$

RG functions at order $1/N$. A new generic integral is useful here:

$$\frac{1}{(2\pi)^d}\int \frac{\mathrm{d}^d q(\not p + \not q)}{(p+q)^{2\mu}q^{2\nu}} = \not p p^{d-2\mu-2\nu}\frac{\Gamma(\mu+\nu-d/2)\Gamma(d/2-\mu+1)\Gamma(d/2-\nu)}{(4\pi)^{d/2}\Gamma(\mu)\Gamma(\nu)\Gamma(d-\mu-\nu+1)}\,. \tag{31.114}$$

We first calculate the $1/N$ contribution to the fermion two-point function at the critical point

$$\Gamma^{(2)}_{\bar\psi\psi}(p) = i\not p + \frac{2iv^2}{b(d)R(v)(2\pi)^d}\int^\Lambda \frac{\mathrm{d}^d q(\not p + \not q)}{q^{d-2}(p+q)^2}\,.$$

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

$$\Gamma^{(2)}_{\bar\psi\psi}(p) = i\not p + \frac{2iv^2}{b(d)R(v)}N_d\left(\frac{d-2}{d}\right)\not p \ln(\Lambda/p)\,, \tag{31.115}$$

where N_d is the loop factor (30.14a). Expressing that the $\langle\bar\psi\psi\rangle$ function satisfies RG equations, we immediately obtain the RG function

$$\eta_\psi(v) = \frac{v^2}{R(v)}\frac{(d-2)}{d}X_1\,, \tag{31.116}$$

where X_1 is given by equation (30.94). We then calculate the function $\langle\sigma\bar\psi\psi\rangle$ at order $1/N$:

$$\Gamma^{(3)}_{\sigma\bar\psi\psi}(p) = v + A_1 D^{-1}(v)v^3\ln\Lambda$$

with

$$A_1 = -\frac{2}{b(d)}N_d = -X_1\,.$$

The last diagram vanishes for symmetry reasons.

The β-function follows

$$\beta_{v^2}(v) = \frac{4(d-1)v^4}{d}X_1 D^{-1}(v), \tag{31.117}$$

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

$$\eta_\psi = \frac{(d-2)}{d}\frac{X_1}{N} = \frac{(d-2)^2}{d}\frac{\Gamma(d-1)}{\Gamma^3(d/2)\Gamma(2-d/2)N}\,. \tag{31.118}$$

$$2d_\psi = d-1-\frac{2(d-2)}{d}\frac{X_1}{N}\,. \tag{31.119}$$

For $d = 4-\varepsilon$, we find $\eta_\psi \sim \varepsilon/N$, result consistent with (31.91) for N large. For $d = 2+\varepsilon$ instead one finds $\eta_\psi \sim \varepsilon^2/2N$, consistent with (31.78). The dimension d_σ of the field σ is

$$d_\sigma = \tfrac{1}{2}(d-2+\eta_\sigma) = 1 - \frac{2(d-1)}{dN}X_1 + O(1/N^2). \tag{31.120}$$

A similar evaluation of the $\langle\sigma^2\sigma\sigma\rangle$ function enables us to determine the exponent ν to order $1/N$:

$$\frac{1}{\nu} = d-2 - \frac{2(d-1)(d-2)}{dN}X_1 \,. \tag{31.121}$$

Actually, all exponents are known to order $1/N^2$ except η_ψ which is known to order $1/N^3$.

Bibliographical Notes

The renormalization group properties of the non-linear σ-model have been discussed in
A.M. Polyakov, *Phys. Lett.* 59B (1975) 79; E. Brézin and J. Zinn-Justin, *Phys. Rev. Lett.* 36 (1976) 691; *Phys. Rev.* B14 (1976) 3110; W.A. Bardeen, B.W. Lee and R.E. Shrock, *Phys. Rev.* D14 (1976) 985.

Higher order calculations of critical exponents are due to
S. Hikami, *Nucl. Phys.* B215[FS7] (1983) 555; W. Bernreuther and F.J. Wegner, *Phys. Rev. Lett.* 57 (1986) 1383; F. Wegner, *Nucl. Phys.* B316 (1989) 663.

Speculations about a RG flow near dimension 2 consistent with the $N = 2$ Kosterlitz–Thouless transition and the non-linear σ-model results for $N > 2$ can be found in
J.L. Cardy and H.W. Hamber, *Phys. Rev. Lett.* 45 (1980) 499.

Long range forces have been discussed in
E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *J. Phys. A: Math. Gen.* 9 (1976) L119.

For an introduction to the problem of localization in electron systems see, for example,
E. Fradkin, G. Parisi in *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, R. Stora and J.-B. Zuber eds. (North Holland, Amsterdam 1984); F. Wegner, *Lecture Notes in Physics* 216 (1985) 141.

Renormalization of composite operators is discussed in
F.J. Wegner, *Z. Phys.* B78 (1990) 33; G.E. Castilla and S. Chakravarty, *Phys. Rev. Lett.* 71 (1993) 384.

For the Gross–Neveu model, see
K.G. Wilson, *Phys. Rev.* D7 (1973) 2911; D.J. Gross and A. Neveu, *Phys. Rev.* D10 (1974) 3235; A.B. Zamolodchikov and Al.B. Zamolodchikov, *Phys. Lett.* 72B (1978) 481; W. Wentzel, *Phys. Lett.* 153B (1985) 297; N.A. Kivel, A.S. Stepanenko and A.N. Vasil'ev, *Nucl. Phys.* B424 [FS] (1994) 619.

For a recent work with RG calculations and many references, see
D.B. Ali and J.A. Gracey, *Four loop wave function renormalization in the non-abelian Thirring model*, hep-th/0104207.

The relation with the GNY model is discussed in
A. Hasenfratz, P. Hasenfratz, K. Jansen, J. Kuti and Y. Shen, *Nucl. Phys.* B365 (1991) 79; J. Zinn-Justin, *Nucl. Phys.* B367 (1991) 105.

Finally, all these techniques have also been applied to the supersymmetric extension of the non-linear σ-model
J.A. Gracey, *Nucl. Phys.* B352 (1991) 183.

32 TWO-DIMENSIONAL MODELS AND BOSONIZATION METHOD

In Chapter 31, we have discussed the generic $O(N)$ non-linear σ-model. We have noticed that the abelian case $N = 2$ is special because the RG β-function vanishes in two dimensions. The corresponding $O(2)$ invariant spin model is specially interesting: it provides an example of the celebrated Kosterlitz–Thouless phase transition and will be examined in Chapter 33. However, a thorough discussion of this model requires a new technique: we have to establish relations, special to two dimensions, between fermion and boson local field theories, by a method called bosonization. The derivation involves several steps which we illustrate in this chapter with the help of various other two-dimensional models which are physically interesting in their own right.

We first study the free massless boson and fermion fields. We evaluate the determinant of the covariant fermion derivative in the presence of an external gauge field. The result is at the basis of the bosonization technique. We then discuss the sine-Gordon (SG) model. We solve the Schwinger model, QED with massless fermions. Finally, we demonstrate the equivalence between the SG model and two fermion models with current–current interaction: the Thirring model and another model with two species of fermions.

Massless (unitary) field theories are conformal invariant. In two dimensions, this is a very powerful property because the conformal group is infinite dimensional. A sophisticated technology has been developed to study them. A general discussion of conformal field theory (CFT) goes much beyond the scope of this work and the interested reader is referred to the abundant literature. Here, we limit ourselves to simple considerations. In the spirit of this work functional integral techniques will be used everywhere though in some cases the operator formulation may lead to more elegant derivations.

The appendix contains some additional remarks concerning the $SU(N)$ Thirring model and solitons in the SG model.

Finally, note that unlike in most other chapters, we here work with correlation functions expressed in terms of space variables rather their Fourier transform. Renormalization then takes the form of defining singular distributions (in the mathematical sense).

32.1 The Free Massless Scalar Field

We first examine the peculiar properties of massless, free, scalar fields in two dimensions. We consider the action $\mathcal{S}$ for a free massive scalar field $\varphi(x)$:

$$\mathcal{S}(\varphi) = \frac{1}{2}\int \mathrm{d}^2x \left[\left(\partial_\mu\varphi(x)\right)^2 + m^2\varphi^2(x)\right]. \tag{32.1}$$

In the massless limit $m = 0$, the action is invariant under constant translations of the field

$$\varphi(x) = \varphi'(x) + \theta\,. \tag{32.2}$$

To this symmetry corresponds a current J^V_μ which is conserved classically:

$$J^V_\mu(x) = \partial_\mu\varphi(x). \tag{32.3}$$

Note a specific property of dimension 2: there exists another current that is trivially conserved:

$$J^A_\mu(x) = \epsilon_{\mu\nu}\partial_\mu\varphi(x) = \epsilon_{\mu\nu}J^V_\nu(x), \tag{32.4}$$

where $\epsilon_{\mu\nu}$ is the antisymmetric tensor with $\epsilon_{12}=1$.

The propagator. The field propagator $\Delta(x)$ is

$$\Delta(x,m) = \frac{1}{(2\pi)^2}\int \mathrm{d}^2p \frac{\mathrm{e}^{ipx}}{p^2+m^2}\,. \tag{32.5}$$

We see that in dimension 2 the propagator of the massless scalar field is IR divergent. The fluctuations which translate the field $\varphi(x)$ by an almost constant function are not damped enough by the action (32.1) and are responsible for this divergence.

The form of the IR divergence can be obtained by expanding the massive propagator (32.5) for $m\to 0$:

$$\Delta(x,m) = -\frac{1}{4\pi}\left(\ln(m^2x^2/4)+2\gamma\right)+O(m)\,, \tag{32.6}$$

where $\gamma=-\psi(1)$ is Euler's constant. At $x=0$, the massive propagator is UV divergent. Introducing a UV cut-off Λ, we find

$$\Delta(0,m) \underset{\Lambda\to\infty}{=} \frac{1}{2\pi}\ln(\Lambda/m)+K+o\left(\Lambda^{-1}\right),$$

where K is a constant which depends on the cut-off procedure. For convenience, we absorb it hereafter in the definition of Λ, setting

$$\Delta(0,m) = \frac{1}{2\pi}\left[\ln(2\Lambda/m)-\gamma\right]+o\left(\Lambda^{-1}\right). \tag{32.7}$$

IR finite correlation functions. We now show that, although the massless 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

$$\left\langle\prod_{i=1}^n \mathrm{e}^{i\kappa_i\varphi(x_i)}\right\rangle = \int[\mathrm{d}\varphi]\exp\left\{-\frac{1}{2}\int\mathrm{d}^2x\left[\left(\partial_\mu\varphi\right)^2+m^2\varphi^2\right]+i\sum_i\kappa_i\varphi(x_i)\right\}, \tag{32.8}$$

where the κ_i's are arbitrary coefficients. We set

$$J(x) = i\sum_i\kappa_i\,\delta(x-x_i). \tag{32.9}$$

The integral in equation (32.8) takes the form

$$\int[\mathrm{d}\varphi]\exp\left\{-\int\mathrm{d}^2x\left[\tfrac{1}{2}\left(\partial_\mu\varphi\right)^2+\tfrac{1}{2}m^2\varphi^2(x)-J(x)\varphi(x)\right]\right\}.$$

We recognize the basic gaussian functional integral

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\kappa_i \varphi(x_i)} \right\rangle = \exp\left[\frac{1}{2}\int \mathrm{d}^2x\,\mathrm{d}^2y\, J(x)\Delta(x-y,m)J(y)\right]$$
$$= \exp\left[-\frac{1}{2}\sum_{i,j}\kappa_i\kappa_j\Delta(x_i-x_j,m)\right]. \tag{32.10}$$

Replacing the propagator by its small m expansion (equation (32.6)) and collecting all IR divergent terms, we find

$$\sum_{i,j}\kappa_i\kappa_j\Delta(x_i-x_j,m) = -\frac{1}{2\pi}\left[\left(\sum_i \kappa_i\right)^2(\ln m - \ln 2 + \gamma) - \sum_i \kappa_i^2 \ln\Lambda\right.$$
$$\left. + \sum_{i\neq j}\kappa_i\kappa_j \ln|x_i - x_j|\right] + O(m)\,. \tag{32.11}$$

Therefore, when m goes to zero only the products such that $\sum_i \kappa_i = 0$ survive. This result is related to the simple property that in two dimensions the Coulomb 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 33.3.1).

From equation (32.11), we also learn that the UV divergences can be removed by a multiplicative renormalization ζ_i of the composite fields $\mathrm{e}^{i\kappa_i\varphi(x)}$:

$$\mathrm{e}^{i\kappa_i\varphi(x)} = \zeta_i \left[\mathrm{e}^{i\kappa_i\varphi(x)}\right]_{\text{ren.}}, \quad \zeta_i = (\Lambda/\mu)^{-\kappa_i^2/(4\pi)}\,, \tag{32.12}$$

where μ is a renormalization scale. Thus, finally in the massless limit $m=0$:

$$\left.\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\kappa_i\varphi(x_i)} \right\rangle\right|_{\text{ren.}} = \prod_{i<j}(\mu|x_i-x_j|)^{\kappa_i\kappa_j/2\pi}\,. \tag{32.13}$$

Note that the correlation functions of exponentials of the field decay algebraically at large distance.

Remark. In a translation (32.2) of the φ-field, the field $\mathrm{e}^{i\kappa_i\varphi}$ becomes

$$\mathrm{e}^{i\kappa_i\varphi(x)} \mapsto \mathrm{e}^{i\kappa_i\theta}\,\mathrm{e}^{i\kappa_i\varphi(x)}\,.$$

All non-vanishing correlation functions are invariant under this transformation, a result in agreement with the analysis in previous chapters of spontaneous breaking of continuous symmetries in two dimensions.

The two charge example. A particular case of equation (32.13) is of special interest, when the κ_i's take only two values $\pm\kappa$. Then, only products with an equal number of $\pm$ signs do not vanish

$$\left.\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\kappa[\varphi(x_i)-\varphi(y_i)]} \right\rangle\right|_{\text{ren.}} = \frac{\prod_{i<j}(\mu|x_i-x_j|)^{\kappa^2/2\pi}(\mu|y_i-y_j|)^{\kappa^2/2\pi}}{\prod_{i,j}(\mu|x_i-y_j|)^{\kappa^2/2\pi}}. \tag{32.14}$$

The two-point function is thus given by

$$\left\langle e^{i\kappa\varphi(x)}\, e^{-i\kappa\varphi(0)}\right\rangle \propto x^{-\kappa^2/2\pi}. \tag{32.15}$$

The dimension of the operator $e^{i\kappa\varphi(x)}$ follows

$$[e^{i\kappa\varphi(x)}] = \kappa^2/4\pi\,. \tag{32.16}$$

Note that for $\kappa^2/2\pi \ge 2$ the two-point function is singular at short distance in the sense of distributions. It has to be renormalized as the UV divergences of its Fourier transform show

$$\int d^2x \frac{e^{ipx}}{x^{\kappa^2/2\pi}} = \pi \frac{\Gamma(1-\kappa^2/4\pi)}{\Gamma(\kappa^2/4\pi)} \left(\frac{p^2}{4}\right)^{\kappa^2/4\pi-1}.$$

This expression is valid for $\kappa^2 < 4\pi$. At $\kappa^2 = 4\pi$, a divergent constant has to be subtracted. For $4\pi \le \kappa^2 < 8\pi$, it is thus defined only up an arbitrary additive renormalization. For $\kappa^2 = 8\pi$, a second additive renormalization proportional to p^2 is required. We shall again meet these two special values when we discuss the SG model.

The currents. The currents (32.3,32.4) provide other examples of fields with IR finite correlation functions. For instance, in Fourier representation, the two-point functions are

$$\left\langle \tilde J^V_\mu(k)\tilde J^V_\nu(-k)\right\rangle = k_\mu k_\nu/k^2\,, \tag{32.17a}$$

$$\left\langle \tilde J^A_\mu(k)\tilde J^A_\nu(-k)\right\rangle = \delta_{\mu\nu} - k_\mu k_\nu/k^2\,. \tag{32.17b}$$

After Fourier transformation, the two-point functions become

$$\langle J^V_\mu(x)J^V_\nu(0)\rangle = \frac{1}{2\pi x^2}\left(\delta_{\mu\nu} - 2\frac{x_\mu x_\nu}{x^2}\right), \tag{32.18a}$$

$$\langle J^A_\mu(x)J^A_\nu(0)\rangle = -\frac{1}{2\pi x^2}\left(\delta_{\mu\nu} - 2\frac{x_\mu x_\nu}{x^2}\right). \tag{32.18b}$$

We find proportional results (as implied by current conservation), although in Fourier space the functions are different: the r.h.s. involves a singular distribution, ambiguous at $x = 0$, defined up to the addition of a $\delta^{(2)}(x)$ function.

It is also easy 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. We find

$$\begin{aligned}\left\langle J^V_\mu(x)\prod_{i=1}^n e^{i\kappa_i\varphi(x_i)}\right\rangle &= i\sum_i \kappa_i \partial^x_\mu \Delta(x-x_i)\left\langle \prod_{i=1}^n e^{i\kappa_i\varphi(x_i)}\right\rangle \\ &= -\frac{i}{2\pi}\sum_i \kappa_i \frac{(x-x_i)_\mu}{(x-x_i)^2}\left\langle \prod_{i=1}^n e^{i\kappa_i\varphi(x_i)}\right\rangle.\end{aligned}$$

Complex coordinates. It is a peculiarity of the dimension 2 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 $x = \{\xi_1, \xi_2\}$ plane in terms of complex coordinates

$$z = \xi_1 + i\xi_2\,, \qquad \bar z = \xi_1 - i\xi_2\,, \tag{32.19}$$

which leads to the action

$$\mathcal{S}(\varphi) = \int \mathrm{d}z\, \mathrm{d}\bar{z}\, \partial_z \varphi\, \partial_{\bar{z}} \varphi\,, \tag{32.20}$$

and to a decomposition of the field into analytic and anti-analytic components

$$\varphi(\xi_1, \xi_2) = \varphi_+(z) + \varphi_-(\bar{z})$$

with the propagators

$$\Delta(\xi_1, \xi_2) = \langle \varphi_+(z)\varphi_+(0)\rangle + \langle \varphi_-(\bar{z})\varphi_-(0)\rangle\,,$$
$$\langle \varphi_+(z)\varphi_+(0)\rangle = -\frac{1}{4\pi}\ln z + \text{const.}\,, \quad \langle \varphi_-(\bar{z})\varphi_-(0)\rangle = -\frac{1}{4\pi}\ln \bar{z} + \text{const.}\,.$$

These complex variables are particularly well suited for exploring the consequences of conformal symmetry: the action (32.20) is obviously invariant in the change $z = f(z'), \bar{z} = \bar{f}(z')$. They will appear naturally in the discussion of massless fermions.

32.2 The Free Massless Dirac Fermion

We now consider the massless Dirac fermion action

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^2 x\, \bar{\psi}(x) \not{\partial} \psi(x)\,. \tag{32.21}$$

The massless classical action has two $U(1)$ symmetries corresponding to phase and chiral phase transformations

$$\begin{aligned} \psi(x) &= \mathrm{e}^{i(\theta_S \gamma_S + i\theta)}\, \psi'(x), \\ \bar{\psi}(x) &= \bar{\psi}'(x)\, \mathrm{e}^{i(\gamma_S \theta_S - i\theta)}\,. \end{aligned} \tag{32.22}$$

A peculiarity of dimension 2 is that the corresponding vector and axial currents

$$J_\mu(x) = \bar{\psi}(x)\gamma_\mu \psi(x)\,, \qquad J_\mu^S(x) = i\bar{\psi}(x)\gamma_S \gamma_\mu \psi(x), \tag{32.23}$$

are related since equation ($A8.23$) implies

$$i\gamma_S \gamma_\mu = -\epsilon_{\mu\nu}\gamma_\nu\,, \tag{32.24}$$
$$\Rightarrow \quad J_\mu^S = -\epsilon_{\mu\nu} J_\nu\,. \tag{32.25}$$

The free massless fermion propagator is (equation (8.36))

$$[\Delta_\psi]_{\alpha\beta}(x) \equiv \langle \bar{\psi}'_\alpha(x) \psi'_\beta(0)\rangle = \frac{1}{4\pi^2}\int \mathrm{d}^2 p\, \mathrm{e}^{-ipx} [i\not{p}]^{-1}_{\beta\alpha} = -\frac{1}{2\pi}[\not{x}]^{-1}_{\beta\alpha}. \tag{32.26}$$

We now decompose the massless fermion into left and right moving components $\psi_\pm = \frac{1}{2}(1 \pm \gamma_S)\psi$, $\bar{\psi}_\pm = \frac{1}{2}\bar{\psi}(1 \mp \gamma_S)$ (equation ($A8.36$)). In the representation in which γ_S is diagonal, we write

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}\,, \quad \bar{\psi} = \begin{pmatrix} \bar{\psi}_- \\ \bar{\psi}_+ \end{pmatrix}\,. \tag{32.27}$$

With this parametrization, the vector current becomes

$$\bar{\psi}(x)\gamma_1 \psi(x) = \bar{\psi}_+ \psi_+ + \bar{\psi}_- \psi_-\,, \quad \bar{\psi}(x)\gamma_2 \psi(x) = i\left(\bar{\psi}_+ \psi_+ - \bar{\psi}_- \psi_-\right)\,. \tag{32.28}$$

In the complex notation (32.19), the action (32.21) becomes

$$\mathcal{S}(\bar{\psi},\psi) = -\int \mathrm{d}z\,\mathrm{d}\bar{z}\,\left(\bar{\psi}_+\partial_{\bar{z}}\psi_+ + \bar{\psi}_-\partial_z\psi_-\right). \tag{32.29}$$

In the euclidean formalism, the left and right movers become analytic and anti-analytic fields. The corresponding propagators Δ_ψ^+ and Δ_ψ^- then read

$$\Delta_\psi^+ \equiv \left\langle \bar{\psi}_+(\bar{z},z)\psi_+(0,0)\right\rangle = -\frac{1}{2\pi z}, \tag{32.30}$$

$$\Delta_\psi^- \equiv \left\langle \bar{\psi}_-(\bar{z},z)\psi_-(0,0)\right\rangle = -\frac{1}{2\pi \bar{z}}. \tag{32.31}$$

The $\psi_\pm$ correlation functions. Using Wick's theorem for Grassmann variables (equation (1.77)), we can calculate the $2n$-point correlation functions of $\psi_\pm, \bar{\psi}_\pm$:

$$\left\langle \prod_{i=1}^{n} \bar{\psi}_+(x_i)\psi_+(x_i')\right\rangle = \det \Delta_\psi^+(z_i - z_j') = \left(\frac{-1}{2\pi}\right)^n \det \frac{1}{z_i - z_j'}, \tag{32.32a}$$

$$\left\langle \prod_{i=1}^{n} \bar{\psi}_-(x_i)\psi_-(x_i')\right\rangle = \det \Delta_\psi^-(\bar{z}_i - \bar{z}_j') = \left(\frac{-1}{2\pi}\right)^n \det \frac{1}{\bar{z}_i - \bar{z}_j'}, \tag{32.32b}$$

where here x_i stands for the pair $\{z_i, \bar{z}_i\}$.

In the r.h.s., we find Cauchy determinants which satisfy the simple identity

$$(-1)^{n+1}\det\frac{1}{z_i - z_j'} = \frac{\prod_{i<j}(z_i - z_j)(z_i' - z_j')}{\prod_{i,j}(z_i - z_j')}, \tag{32.33}$$

and its complex conjugated. This identity enables us to relate boson and fermion quantum field theories in two dimensions.

32.2.1 Bilinear operator correlations and boson–fermion correspondence

We now consider the two composite operators

$$\sigma_+(x) = \bar{\psi}_-(x)\psi_+(x)\,, \qquad \sigma_-(x) = \bar{\psi}_+(x)\psi_-(x), \tag{32.34}$$

linear combinations of the scalar and pseudoscalar operators

$$\bar{\psi}(x)\psi(x) = \sigma_+(x) + \sigma_-(x)\,, \qquad \bar{\psi}(x)\gamma_S\psi(x) = \sigma_+(x) - \sigma_-(x)\,. \tag{32.35}$$

From the form of the action, it is clear 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

$$\left\langle \prod_{i=1}^{n}\sigma_+(x_i)\sigma_-(x_i')\right\rangle = (-1)^n \left\langle \prod_{i=1}^{n}\bar{\psi}_+(x_i')\psi_+(x_i)\right\rangle \left\langle \prod_{i=1}^{n}\bar{\psi}_-(x_i)\psi_-(x_i')\right\rangle, \tag{32.36}$$

where again x_i stands for the pair $\{z_i, \bar{z}_i\}$.

We now use the identities (32.32) and (32.33), and find

$$\left\langle \prod_{i=1}^{n} \sigma_+(x_i)\sigma_-(x'_i) \right\rangle = \left(\frac{1}{2\pi}\right)^{2n} \frac{\prod_{i<j} |x_i - x_j|^2 \, |x'_i - x'_j|^2}{\prod_{i,j} |x_i - x'_j|^2}. \tag{32.37}$$

The r.h.s. has short distance divergences associated with the required additive renormalization of the $\sigma_+\sigma_-$ two-point function.

Comparing with equation (32.14), we discover an identity between averages with the free massless fermion and boson actions:

$$\left.\left\langle \prod_{i=1}^{n} \mu^2 \, \mathrm{e}^{i\kappa(\varphi(x_i)-\varphi(x'_i))} \right\rangle_{\varphi} \right|_{\text{ren.}} = (2\pi)^{2n} \left\langle \prod_{i=1}^{n} \sigma_+(x_i)\sigma_-(x'_i) \right\rangle_{\psi}$$

for $\kappa^2 = 4\pi$. This equation can be translated into a correspondence between operators

$$2\pi\sigma_\pm(x) \mapsto \mu \left[\mathrm{e}^{\pm i\sqrt{4\pi}\varphi(x)}\right]_{\text{ren.}} = \Lambda \, \mathrm{e}^{\pm i\sqrt{4\pi}\varphi(x)} \tag{32.38}$$

with the definition (32.7). 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.

32.2.2 The massive free fermion and the sine-Gordon model

We consider the partition function $\mathcal{Z}_\psi(M_+, M_-)$:

$$\mathcal{Z}_\psi(M_\pm) = \int [\mathrm{d}\psi \, \mathrm{d}\bar\psi] \exp[-\mathcal{S}(\bar\psi, \psi)] \tag{32.39}$$

with the action

$$\mathcal{S}(\bar\psi, \psi) = -\int \mathrm{d}^2 x \left[\bar\psi(x) \not{\partial} \psi(x) + M_+(x)\sigma_+(x) + M_-(x)\sigma_-(x)\right] . \tag{32.40}$$

If we expand $\mathcal{Z}_\psi$ in powers of the sources $M_\pm$, the term of degree n in $M_\pm$ in the expansion is an expectation value with the action (32.21) of the form

$$\frac{1}{n!} \left\langle \left[\int \mathrm{d}^2 x \left(M_+(x)\sigma_+(x) + M_-(x)\sigma_-(x)\right)\right]^n \right\rangle . \tag{32.41}$$

Evaluating the expectation values, we obtain an expansion of the partition function $\mathcal{Z}_\psi(M)$ of the form

$$\mathcal{Z}_\psi(M) = \sum_n \frac{1}{(n!)^2} \int \prod_i \mathrm{d}^2 x_i \, \mathrm{d}^2 y_i \, \frac{M_+(x_i) M_-(y_i)}{(2\pi)^2} \frac{\prod_{i<j} |x_i - x_j|^2 \, |y_i - y_j|^2}{\prod_{i,j} |x_i - y_j|^2}. \tag{32.42}$$

We can also directly use the correspondence (32.38). The term (32.41) is replaced by

$$\frac{1}{n!} \left(\frac{\mu}{2\pi\zeta}\right)^n \left\langle \left[\int \mathrm{d}^2 x \left(M_+(x) \, \mathrm{e}^{i\sqrt{4\pi}\varphi(x)} + M_-(x) \, \mathrm{e}^{-i\sqrt{4\pi}\varphi(x)}\right)\right]^n \right\rangle_{\varphi} , \tag{32.43}$$

where ζ is the renormalization constant (32.12) for $\kappa^2 = 4\pi$:

$$\zeta = \mu/\Lambda \Rightarrow \mu/\zeta = \Lambda \tag{32.44}$$

with the definition (32.7).

We now sum this 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 with the action

$$\mathcal{S}(\varphi) = \int \mathrm{d}^2 x \left\{ \frac{1}{2} \left[\partial_\mu \varphi(x)\right]^2 - \frac{\Lambda}{2\pi} \left[M_+(x)\, \mathrm{e}^{i\sqrt{4\pi}\varphi(x)} + M_-(x)\, \mathrm{e}^{-i\sqrt{4\pi}\varphi(x)} \right] \right\}. \tag{32.45}$$

If we take the constant $M_+ = M_- = M$ limit, we find another remarkable correspondence: the free massive fermion theory is equivalent to a particular case of the SG model. This relation will be discussed in Section 32.3 and generalized in Section 32.5.

32.2.3 The gauge invariant fermion determinant

We now consider the fermion action in an external gauge field B_μ:

$$\mathcal{S}(\bar\psi, \psi, B) = -\int \mathrm{d}^2 x\, \bar\psi(x) \left[\not\partial + i \not B(x) \right] \psi(x). \tag{32.46}$$

Because the vector and axial current are related, the classical action (32.46) is invariant not only under phase gauge transformations but also under chiral gauge transformations:

$$\psi(x) = \mathrm{e}^{i\gamma_S \varphi(x)}\, \psi'(x), \qquad \bar\psi(x) = \bar\psi'(x)\, \mathrm{e}^{i\gamma_S\varphi(x)}. \tag{32.47}$$

Indeed, this transformation generates the term

$$\bar\psi'(x) i \not\partial \varphi(x) \gamma_S \psi'(x) = \bar\psi'(x) \epsilon_{\mu\nu} \gamma_\nu \partial_\mu \varphi(x) \gamma_S \psi'(x), \tag{32.48}$$

which is cancelled by the transformation

$$B_\mu(x) = B'_\mu(x) - i\epsilon_{\mu\nu} \partial_\nu \varphi(x). \tag{32.49}$$

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 by gauge transformations from the classical action, which is thus equivalent to a free field action. Indeed, let us parametrize B_μ as

$$B_\mu(x) = -\left[\partial_\mu \chi(x) + i\epsilon_{\mu\nu} \partial_\nu \varphi(x) \right], \tag{32.50}$$

which implies

$$i\epsilon_{\mu\nu} \partial^2 \varphi(x) = \partial_\mu B_\nu - \partial_\nu B_\mu\,, \quad \text{or} \quad \partial^2 \varphi(x) = -i\epsilon_{\mu\nu} \partial_\mu B_\nu\,, \tag{32.51}$$

and set

$$\begin{aligned} \psi(x) &= \mathrm{e}^{i(\chi(x) + \gamma_S \varphi(x))}\, \psi'(x), \\ \bar\psi(x) &= \bar\psi'(x)\, \mathrm{e}^{i(-\chi(x) + \gamma_S\varphi(x))}. \end{aligned} \tag{32.52}$$

The action (32.46) then becomes

$$\mathcal{S}(\bar\psi, \psi, B) = -\int \mathrm{d}^2 x\, \bar\psi'(x) \not\partial \psi'(x). \tag{32.53}$$

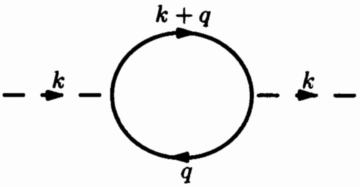


Fig. 32.1 The one-loop contribution to the current two-point function.

Quantum theory: the anomaly. As we have more generally discussed in Section 20.3 the chiral gauge symmetry is broken at the quantum level by the chiral anomaly. Here, we will recover the anomaly by a direct calculation. In Appendix A32.1.1, we show the consistency of the result obtained below with the general form derived in Section 20.3 (equation (20.101)).

The integral over fermion fields generates a determinant which has an expansion in terms of one-loop diagrams. In two dimensions, only the term quadratic in B_μ shown in figure 32.1 is divergent.

To regularize the determinant, we introduce a regulator field in the form of a massive spinor boson ϕ:

$$\mathcal{S}(\bar{\psi}, \psi, B_\mu, \phi) = -\int \mathrm{d}^2x \left[\bar{\psi}(x) \left(\not{\partial} + i\not{B}\right) \psi(x) - \bar{\phi}(x) \left(\not{\partial} + i\not{B} + M\right) \phi(x)\right]. \quad (32.54)$$

With this addition, the theory is finite but the ϕ mass breaks chiral gauge invariance explicitly (not global chiral symmetry). Let us parametrize the two-component gauge field in terms of two scalar fields as in equation (32.50) and perform the corresponding gauge transformations (32.52) both on $\bar{\psi}, \psi$ and $\bar{\phi}, \phi$. The action becomes

$$\mathcal{S}(\bar{\psi}', \psi', B_\mu, \phi') = -\int \mathrm{d}^2x \left[\bar{\psi}'(x)\not{\partial}\psi'(x) - \bar{\phi}'(x) \left(\not{\partial} + M\, \mathrm{e}^{2i\gamma_S \varphi(x)}\right) \phi'(x)\right]. \quad (32.55)$$

We now integrate over the fields $\bar{\phi}', \phi'$. The result is a functional $\mathcal{D}(\varphi)$,

$$\mathcal{D}^{-1}(\varphi) = \det\left(\not{\partial} + M\, \mathrm{e}^{2i\gamma_S \varphi(x)}\right), \quad (32.56)$$

which has to be evaluated for M large. The ***bosonization*** technique is useful in this respect. Indeed, $\mathcal{D}^{-1}(\varphi)$ can be written as a fermion integral

$$\frac{1}{\mathcal{D}(\varphi)} = \int [\mathrm{d}\psi\, \mathrm{d}\bar{\psi}] \exp \int \mathrm{d}^2x \left[\bar{\psi}(x)\not{\partial}\psi(x) + \mathrm{e}^{2i\varphi(x)}\, \sigma_+(x) + \mathrm{e}^{-2i\varphi(x)}\, \sigma_-(x)\right]. \quad (32.57)$$

In Section 32.2.2, we have derived a fermion–boson equivalence. We can thus replace the fermion average with the action (32.40) by a boson average with the action (32.45) and $M_\pm = \mathrm{e}^{\pm 2i\varphi}$. Denoting by $\vartheta(x)$ the SG field, we find a modified SG model (32.45) with ϑ replaced by $\vartheta + \varphi/\sqrt{\pi}$ in the interaction

$$\mathcal{S}(\vartheta, \varphi) = \int \mathrm{d}^2x \left\{\frac{1}{2}\left[\partial_\mu \vartheta(x)\right]^2 - \frac{M\Lambda}{\pi} \cos\left(\sqrt{4\pi}\vartheta + 2\varphi\right)\right\}. \quad (32.58)$$

We now translate ϑ setting $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. The action becomes

$$\mathcal{S}'(\vartheta) = \int \mathrm{d}^2x \left\{\frac{1}{2}\left[\partial_\mu \vartheta(x) - \partial_\mu \varphi(x)/\sqrt{\pi}\right]^2 - \frac{M\Lambda}{\pi} \cos\sqrt{4\pi}\vartheta(x)\right\}. \quad (32.59)$$

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 find the finite result

$$\mathcal{D}^{-1}(\varphi)\propto\exp\left[-\frac{1}{2\pi}\int d^2x\left(\partial_\mu\varphi(x)\right)^2\right]. \tag{32.60}$$

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 (32.46) is thus equivalent to

$$\mathcal{S}(\bar\psi',\psi',B_\mu)=-\int d^2x\left[\bar\psi'(x)\not\partial\psi'(x)+\frac{1}{2\pi}\int d^2x\left(\partial_\mu\varphi(x)\right)^2\right]. \tag{32.61}$$

32.2.4 Current correlation functions

Note first that the expressions (32.32) are singular when $x_i\to x'_i$, therefore, they cannot be used to define the current without a short distance regularization. We can avoid this problem because we note that in the action (32.46) 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 (32.60) expanded at second order in φ and the relation (32.51) we obtain the current two-point function

$$\left\langle\tilde J_\mu(k)\tilde J_\nu(-k)\right\rangle=\frac{1}{\pi}\left(\delta_{\mu\nu}-\frac{k_\mu k_\nu}{k^2}\right),$$

$$\langle J_\mu(x)J_\nu(0)\rangle=-\frac{1}{2\pi^2x^2}\left(\delta_{\mu\nu}-2\frac{x_\mu x_\nu}{x^2}\right),$$

expressions proportional to the two-point functions (32.17*b*) and (32.18*b*) of the bosonic current. Note that the second expression is a distribution singular at $x=0$, where it is ambiguous.

Then, using equation (32.25), we obtain the axial current two-point function

$$\left\langle\tilde J^S_\mu(k)\tilde J^S_\nu(-k)\right\rangle=\frac{1}{\pi}\frac{k_\mu k_\nu}{k^2},$$

a result identical to (32.17*a*). In the space variables, both currents are proportional but their Fourier representations are different: we see that only the vector current is conserved.

Current insertion. We now calculate $\sigma_\pm$ correlations with one current $J_\mu(x)$ insertion. In the transformation (32.52), 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

$$2\epsilon_{\mu\nu}\int d^2x\,\partial_\mu B_\nu(x)\sum_i\left[\Delta(x-x_i)-\Delta(x-x'_i)\right].$$

Differentiating with respect to $B_\mu(x)$, we finally obtain

$$\left\langle J_\mu(x)\prod_{i=1}^n\sigma_+(x_i)\sigma_-(x'_i)\right\rangle=\left\langle\prod_{i=1}^n\sigma_+(x_i)\sigma_-(x'_i)\right\rangle\\ \times\sum_i\frac{i}{\pi}\epsilon_{\mu\nu}\left(\frac{(x-x_i)_\nu}{(x-x_i)^2}-\frac{(x-x'_i)_\nu}{(x-x'_i)^2}\right). \tag{32.62}$$

This expression is proportional to the boson current insertion $J^A_\mu(x)$ for $\kappa_i = \pm\sqrt{4\pi}$. In the same way, the axial current J^A_μ is proportional to J^S_μ:

$$J^A_\mu(x) \;\mapsto\; \sqrt{\pi} J_\mu(x), \quad J^V_\mu(x) \;\mapsto\; \sqrt{\pi} J^S_\mu(x). \tag{32.63}$$

The second result is not surprising: a translation of φ, $\varphi \mapsto \varphi + \theta$ multiplies $\mathrm{e}^{\pm i\sqrt{4\pi}\varphi}$ by $\mathrm{e}^{\pm i\sqrt{4\pi}\theta}$. To J^S_μ is associated the transformation $\sigma_\pm \mapsto \sigma_\pm\,\mathrm{e}^{\pm 2i\theta}$.

32.3 The Sine-Gordon Model

The SG model is defined by the action

$$\mathcal{S}(\vartheta) = \int \mathrm{d}^2 x \left\{ \frac{1}{2} \left[\partial_\mu \vartheta(x)\right]^2 - \frac{\alpha_0}{\kappa^2} \cos \kappa\vartheta(x) \right\}, \tag{32.64}$$

where $\vartheta(x)$ is a scalar boson field. Note one can choose $\alpha_0 > 0$ without loss of generality. Depending on the question we want to investigate, we will sometimes normalize the SG field differently, setting $\theta = \kappa\vartheta$ and thus

$$\mathcal{S}(\theta) = \frac{1}{t} \int \mathrm{d}^2 x \left\{ \frac{1}{2} \left[\partial_\mu \theta(x)\right]^2 - \alpha_0 \cos \theta(x) \right\} \qquad \text{with } t = \kappa^2 . \tag{32.65}$$

The model has been extensively studied. The classical field equations are completely integrable. This allows to obtain finite energy solutions of the real-time equation of motions, ***solitons***, and to infer the semi-classical spectrum. 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. The relevant techniques, however, are outside the scope of this work, and will not be presented here. We merely want to study the model for some of its algebraic and RG properties.

32.3.1 Perturbative expansion

For κ small, the field θ fluctuates around one of the minima $\theta = 2n\pi$ of $-\cos\theta$. We can choose one of them, $\theta = 0$, to expand perturbation theory because they are all equivalent and cannot be connected (see Section 41.4). Note, however, that the degeneracy of the classical minimum is responsible for the existence of solitons. Expanding $\cos\theta$ in powers of θ, we see that the θ-field is massive with a mass $\alpha^{1/2} + O(\kappa^2)$.

The model is super-renormalizable. From the discussion of Appendix A10.2, we know that divergences, in the perturbative expansion in powers of κ, arise only from the self-contractions of the interaction term $\cos\theta$. These divergences have actually been calculated in Section 32.1. Using equation (32.12), we can set

$$\mathrm{e}^{\pm i\theta(x)} = \mathrm{e}^{\pm i\kappa\vartheta(x)} = Z_\theta^{1/2} \left[\mathrm{e}^{\pm i\theta(x)}\right]_{\text{ren.}}, \tag{32.66a}$$

$$\alpha_0 = Z_\theta^{-1/2} \alpha \tag{32.66b}$$

$$\text{with} \quad Z_\theta = (\Lambda/\mu)^{-\kappa^2/(2\pi)}, \tag{32.66c}$$

in which μ is the renormalization scale.

32.3.2 RG equations

The theory is super-renormalizable and thus the β-function vanishes. The field $e^{\pm i\theta(x)}$ and the coupling constant α RG functions are given in terms of Z_θ. We find

$$\eta(t=\kappa^2)=\mu\frac{\partial}{\partial\mu}\ln Z_\theta(t,\Lambda/\mu)=\frac{t}{2\pi}. \tag{32.67}$$

The renormalized n-point correlation function $W^{(n)}$ of $e^{\pm i\theta(x)}$ thus satisfies the RG equation:

$$\left[\mu\frac{\partial}{\partial\mu}+\frac{n}{2}\eta(t)+\frac{1}{2}\eta(t)\alpha\frac{\partial}{\partial\alpha}\right]W^{(n)}(p_i,t,\alpha)=0\,. \tag{32.68}$$

Using the dimensional relation

$$\mu\frac{\partial}{\partial\mu}+2\alpha\frac{\partial}{\partial\alpha}+p_i\frac{\partial}{\partial p_i}=2(1-n)\,,$$

we obtain the scaling relation

$$W^{(n)}(p_i)=\alpha^{t/(8\pi-t)}\alpha^{1-n}F^{(n)}\left(p_i\alpha^{-4\pi/(8\pi-t)},t\right). \tag{32.69}$$

In particular, for the one-point function and the mass scale or θ-mass we find

$$\left\langle e^{\pm i\theta(x)}\right\rangle=W^{(1)}\sim\alpha^{t/(8\pi-t)},\qquad m_\theta\sim\alpha^{4\pi/(8\pi-t)}. \tag{32.70}$$

In these relations, we see that two values of $t=\kappa^2$ play a special role:

(i) For $\kappa^2=4\pi$ the propagator (32.15) 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 (32.45) with a free massive fermion model. Indeed, $m_\theta\propto m_\psi=\alpha$ and

$$\langle e^{i\theta}\rangle\propto\langle\bar\psi(x)\psi(x)\rangle=\frac{1}{2\pi}\operatorname{tr}\int d^2p\frac{1}{i\not p+\alpha}\sim\alpha\ln(\Lambda^2/\alpha^2).$$

The UV divergence is the same that we have discussed above since by differentiating with respect to α, we obtain the two-point function at zero-momentum. It leads to a logarithmic correction to the linear behaviour. We shall further discuss this relation in Section 32.5. For $\kappa^2>4\pi$, generically the leading contribution to $W^{(1)}$ comes from short distance effects and remains linear in α.

(ii) For $\kappa^2=8\pi$, the quantities $\langle e^{i\theta}\rangle$ and m_θ in equations (32.70) vanish identically for α small. We can find an interpretation of the result by calculating the dimension of the operator $\cos\theta(x)$ and more generally of $\cos l\theta(x)$:

$$[\cos l\theta]=l^2\kappa^2/4\pi\,. \tag{32.71}$$

Therefore, these operators, which are relevant for κ small and give a mass to the θ-field, become irrelevant beyond a finite value of κ:

$$l^2\kappa^2/4\pi=2\Longrightarrow\kappa^2=8\pi/l^2. \tag{32.72}$$

For $\kappa^2>8\pi$, the last operator $\cos l\theta$ also becomes irrelevant, which explains why no mass is generated. At $\kappa^2=8\pi$, the interaction $\int d^2x\cos\theta(x)$ is marginal. Translated into field theory language this means that the theory is just renormalizable; for $\kappa^2>8\pi$ the theory is no longer renormalizable. We discuss in Chapter 33 in more detail this peculiar transition. Note, however, that this analysis is valid only at leading order in α since at $\kappa^2=8\pi$ new divergences are generated at each order in α and thus the dimension of operators can be modified.

32.4 The Schwinger Model

We now consider two-dimensional QED, with one Dirac fermion coupled to an abelian gauge field. The bosonization technique will allow to solve the massless model exactly and will provide us with some interesting information about the massive model.

32.4.1 The massless model

The corresponding action reads

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int d^2x \left[\tfrac{1}{4} F_{\mu\nu}^2(x) - \bar{\psi}(x) \left(\not{\partial} + ie\not{A}\right) \psi(x)\right]. \tag{32.73}$$

This model, first discussed by Schwinger, exhibits the simplest example of a *chiral anomaly*, illustrates both *confinement* (see Chapter 34) and *spontaneous chiral symmetry breaking* in two dimensions.

The field theory is super-renormalizable by power counting in a covariant gauge. The only divergent diagram corresponds to the one-loop contribution to the gauge field two-point function (figure 32.1) and comes from the expansion of the fermion determinant that we have already discussed in Section 32.2.3. In Section 32.2.3, we have shown that the fermion part of the action with $B_\mu = eA_\mu$, after the gauge transformations (32.52) and with the parametrization (32.50), is equivalent to the free field action (32.61). We have here simply to add the contribution coming from $F_{\mu\nu}^2$:

$$\frac{1}{4} F_{\mu\nu}^2 = \frac{1}{2e^2} \partial_\mu \varphi \partial^2 \partial_\mu \varphi .$$

In terms of the fields $\bar{\psi}', \psi', \varphi$, the Schwinger model action then takes the free field form (in the gauge $\partial_\mu A_\mu = \partial^2 \chi / e = 0$):

$$\mathcal{S}(\bar{\psi}', \psi', \varphi) = \int d^2x \left[-\bar{\psi}' \not{\partial} \psi' - \frac{1}{2\pi} \left(\partial_\mu \varphi\right)^2 + \frac{1}{2e^2} \partial_\mu \varphi \partial^2 \partial_\mu \varphi\right]. \tag{32.74}$$

We infer the φ-field propagator in momentum space

$$\Delta_\varphi(p) = \pi \left(\frac{1}{p^2 + e^2/\pi} - \frac{1}{p^2}\right). \tag{32.75}$$

This expression shows that the scalar φ-field propagates two modes corresponding to a positive metric neutral massive field with mass

$$m = e/\sqrt{\pi}\,, \tag{32.76}$$

and a massless mode with negative metric. The appearance of a non-vanishing mass is a direct consequence of the chiral anomaly.

Finally, using the representation (32.50), we can calculate the gauge field transverse two-point function. We find

$$\Delta_{\mu\nu}^{(2)}(p) = \left(\delta_{\mu\nu} - p_\mu p_\nu / p^2\right) \frac{1}{p^2 + m^2}\,, \tag{32.77}$$

a result we verify in Appendix A32.1.2 by a direct one-loop calculation.

We can now also bosonize the free fermion $\bar\psi', \psi'$, introducing a free massless boson ϑ. The action becomes

$$\mathcal{S}(\vartheta,\varphi) = \int \mathrm{d}^2x \left[\frac{1}{2}(\partial_\mu\vartheta)^2 - \frac{1}{2\pi}(\partial_\mu\varphi)^2 + \frac{1}{2\pi m^2}\partial_\mu\varphi\partial^2\partial_\mu\varphi\right]. \tag{32.78}$$

We then translate ϑ, $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. The action becomes

$$\mathcal{S}(\vartheta,\varphi) = \int \mathrm{d}^2x \left[\frac{1}{2}(\partial_\mu\vartheta)^2 - \frac{1}{\sqrt{\pi}}\partial_\mu\varphi\partial_\mu\vartheta + \frac{1}{2\pi m^2}\partial_\mu\varphi\partial^2\partial_\mu\varphi\right].$$

We then integrate over φ and finally obtain

$$\mathcal{S}(\vartheta) = \frac{1}{2}\int \mathrm{d}^2x \left[(\partial_\mu\vartheta(x))^2 + m^2\vartheta^2(x)\right]. \tag{32.79}$$

It is remarkable that the fermion contribution just cancels the massless boson field, leaving only the massive boson. Starting from the representation (32.52), it is easy to verify that, in these transformations the chiral components $\sigma_\pm$ of the neutral composite field $\bar\psi\psi$ are mapped to

$$\sigma_\pm(x) = \bar\psi_\mp(x)\psi_\pm(x) \ \longmapsto\ \frac{\Lambda}{2\pi}\,\mathrm{e}^{\pm i\sqrt{4\pi}\vartheta(x)}. \tag{32.80}$$

Remark. Before integration over φ, the action is formally invariant when we translate ϑ by a constant. Therefore, the form (32.79) corresponds to a choice of boundary conditions on the inverse of the Laplace operator. A different choice leads to the replacement $\vartheta(x) \mapsto \vartheta(x) - \vartheta_\infty$. From the correspondence (32.80), 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 $\partial^2\varphi \propto \epsilon_{\mu\nu}F_{\mu\nu}$ 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.

32.4.2 Confinement and chiral symmetry breaking

We note that the gauge field (32.77) two-point function has no cut corresponding to fermion intermediate states. This is a sign of *confinement* because it means 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 34).

This observation is confirmed by the form of $\sigma_\pm(x)$ correlation functions. Indeed, they can be calculated with the massive free action (32.79) where no massless field appears.

We now calculate the expectation values of $\sigma_\pm$. From the correspondence (32.80) and equation (32.10), we obtain

$$\langle\sigma_\pm\rangle = \frac{\Lambda}{2\pi}\left\langle \mathrm{e}^{\pm i\sqrt{4\pi}\vartheta(x)}\right\rangle = \frac{\Lambda}{2\pi}\,\mathrm{e}^{-2\pi\Delta(0,m)}.$$

Taking into account the definition (32.7), we find a finite result

$$\langle\sigma_\pm\rangle = \frac{\mathrm{e}^\gamma}{4\pi}m \quad\Rightarrow\quad \left\langle\bar\psi(x)\psi(x)\right\rangle = \frac{\mathrm{e}^\gamma}{2\pi}m. \tag{32.81}$$

Since $\bar{\psi}\psi$ is a composite field which is not chiral invariant, the non-vanishing result shows that global chiral symmetry is spontaneously broken. Note that spontaneous symmetry breaking with order is here possible because the electromagnetic interaction generates long range forces.

We now calculate from equation (32.10), the two-point functions of $\sigma_\pm$:

$$\langle \sigma_{\epsilon_1}(x_1)\sigma_{\epsilon_2}(x_2)\rangle = \langle\sigma\rangle^2 \, \mathrm{e}^{-4\pi\epsilon_1\epsilon_2\Delta(x_1-x_2,m)} . \tag{32.82}$$

It follows

$$\left\langle \bar{\psi}(x)\psi(x)\bar{\psi}(0)\psi(0)\right\rangle = \left\langle\bar{\psi}\psi\right\rangle^2 \cosh\left(4\pi\Delta(x,m)\right), \tag{32.83a}$$

$$\left\langle \bar{\psi}(x)\gamma_S\psi(x)\bar{\psi}(0)\gamma_S\psi(0)\right\rangle = \left\langle\bar{\psi}\psi\right\rangle^2 \sinh\left(4\pi\Delta(x,m)\right). \tag{32.83b}$$

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 (32.83*b*) 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 (32.83*a*) 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 $\bar{\psi}\gamma_S\psi$ two-point function.

Finally, the short distance behaviour of the propagator $\Delta(m,x)$ is given by the asymptotic form (32.6), 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 zero 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.

32.4.3 The massive Schwinger model

Let us now briefly discuss the effect of the addition of a fermion mass term. Of course, chiral symmetry is no longer an issue since the mass term explicitly breaks chiral symmetry. However, there is still the problem of confinement: will charged particles appear in the spectrum?

We thus consider the action

$$\mathcal{S}(\bar{\psi},\psi,A_\mu) = \int \mathrm{d}^2 x \left[\tfrac{1}{4}F^2_{\mu\nu}(x) - \bar{\psi}(x)\left(\not{\partial} + ie\not{A} + M\right)\psi(x)\right]. \tag{32.84}$$

We now perform the transformations (32.47) and (32.50) ($B_\mu = eA_\mu$) and find

$$\mathcal{S}(\bar{\psi}',\psi',\varphi) = -\int \mathrm{d}^2 x \left[\bar{\psi}'\left(\not{\partial} + M\,\mathrm{e}^{2i\gamma_S\varphi}\right)\psi' + \frac{1}{2e^2}\partial_\mu\varphi\left(m^2-\partial^2\right)\partial_\mu\varphi\right]. \tag{32.85}$$

The model is no longer free and cannot be solved exactly. However, the action can be further transformed, as shown in Section 32.2.2. We call ϑ the boson field associated by bosonization with $\bar{\psi}', \psi'$. We then use the equivalence between the actions (32.40) and (32.45). The fermion action is replaced by the action (32.58):

$$\mathcal{S}(\vartheta, \varphi) = \int \mathrm{d}^2 x \left\{ \frac{1}{2} \left[\partial_\mu \vartheta(x)\right]^2 - \frac{M\Lambda}{\pi} \cos\left(\sqrt{4\pi}\vartheta + 2\varphi\right) - \frac{1}{2e^2} \partial_\mu \varphi \left(m^2 - \partial^2\right) \partial_\mu \varphi \right\}.$$

We now translate ϑ setting $\vartheta + \varphi/\sqrt{\pi} \mapsto \vartheta$. After translation, the action becomes quadratic in φ. We can thus integrate over φ and find a "massive" SG action for ϑ:

$$\mathcal{S}(\vartheta) = \int \mathrm{d}^2 x \left\{ \frac{1}{2} \left[\partial_\mu \vartheta(x)\right]^2 + \frac{1}{2} m^2 \vartheta^2(x) - \alpha_0 \cos \sqrt{4\pi}\vartheta(x) \right\} \tag{32.86}$$

with the correspondence (32.80) and

$$\alpha_0 = M\Lambda/\pi \,.$$

Physical consequences. We see in action (32.86) 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. Note on the other hand that 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. This potential raises, at large distance, linearly with the distance between charged particles and thus charged particles can never be separated.

32.5 The Massive Thirring Model

In Section 32.2, comparing equations (32.13,32.37), we have discovered identities between free massless fermion averages of powers of $\bar{\psi}\psi$ and averages of exponentials of free massless boson fields. This has led to the correspondence derived in Section 32.2.2 between the free massive fermion and the SG model for $\kappa^2 = 4\pi$. The same identities allow to establish a relation between the general quantum SG and the massive Thirring model. The Thirring model is described in terms of Dirac fermions with an action:

$$\mathcal{S}(\bar{\psi}, \psi) = -\int \mathrm{d}^2 x \left[\bar{\psi} \left(\not{\partial} + m_0\right) \psi - \tfrac{1}{2} g J_\mu J_\mu\right], \tag{32.87}$$

where

$$J_\mu = \bar{\psi} \gamma_\mu \psi \,. \tag{32.88}$$

For $m_0 = 0$, the Thirring model retains the $U(1)$ symmetries (32.22) 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 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:

$$1 + g/\pi = 4\pi/\kappa^2 \,, \tag{32.89a}$$

$$J_\mu \longmapsto \frac{1}{2\pi} \epsilon_{\mu\nu} \partial_\nu \theta \,, \tag{32.89b}$$

$$\bar{\psi}\psi \longmapsto \frac{\Lambda}{\pi} \cos\theta \,. \tag{32.89c}$$

The chiral invariant model, $m_0 = 0$, is mapped onto a free boson theory, as we have already discussed in Section 32.3, and thus exactly soluble.

We derive below this correspondence, using the bosonization identities established in previous sections.

32.5.1 Bosonization: the sine-Gordon model

The first step of the bosonization is to introduce a vector field A_μ and rewrite the interaction term as a gaussian integral over A_μ:

$$\tfrac{1}{2} g J_\mu J_\mu \longmapsto A_\mu^2/2g + iA_\mu J_\mu \,. \tag{32.90}$$

The fermion action becomes a quadratic action of charged fermions coupled to an abelian gauge field:

$$\mathcal{S}(A_\mu, \bar\psi, \psi) = -\int \mathrm{d}^2x \left[\bar\psi \left(\not\partial + i\not A + m_0\right)\psi - A_\mu^2/2g\right]. \tag{32.91}$$

Bosonization. The action then differs from the massive Schwinger model action (32.84) only by the kinetic gauge field term. As we discussed in Section 32.2.2, the action can thus be replaced by a purely bosonic action. Parametrizing the two-component field A_μ in terms of two scalar fields χ and φ as in (32.50), we find

$$\int \mathrm{d}^2x\, A_\mu^2(x) = \int \mathrm{d}^2x \left[\left(\partial_\mu \chi(x)\right)^2 - \left(\partial_\mu \varphi(x)\right)^2\right].$$

After the gauge transformations (32.52), the action (32.91) then becomes

$$\begin{aligned}\mathcal{S}(\chi, \varphi, \bar\psi', \psi') = {} & -\int \mathrm{d}^2x\, \bar\psi' \left(\not\partial + m_0\, \mathrm{e}^{2i\gamma_S \varphi}\right)\psi' \\ & + \frac{1}{2}\int \mathrm{d}^2x \left[\frac{1}{g}\left(\partial_\mu\chi\right)^2 - \left(\frac{1}{g} + \frac{1}{\pi}\right)\left(\partial_\mu\varphi\right)^2\right],\end{aligned} \tag{32.92}$$

where we have taken into account the anomaly.

The field χ decouples from neutral correlation functions. The kinetic term of the field φ has the wrong sign. Averages of products of $\mathrm{e}^{\pm i\varphi}$ when the numbers of + and − signs are different are thus IR divergent. Fortunately, they always appear multiplied by fermion averages which, as we have indicated in Section 32.2.1, due to chiral symmetry vanish.

The correspondence between the massive Thirring and SG models now follows directly from the analysis of Section 32.2.2. We have shown that the fermion action can be replaced by a SG action for $\kappa^2 = t = 4\pi$ of the form (32.58):

$$\int \mathrm{d}^2x\, \bar\psi' \left(\not\partial + m_0\, \mathrm{e}^{2i\gamma_S\varphi}\right)\psi' \mapsto \int \mathrm{d}^2x \left[\frac{1}{2}\left(\partial_\mu\vartheta\right)^2 - \frac{m_0\Lambda}{\pi}\cos\left(\sqrt{4\pi}\vartheta + 2\varphi\right)\right]. \tag{32.93}$$

In the special case of the massless Thirring model $m_0 = 0$, we obtain an equivalent free boson theory:

$$\mathcal{S}(\vartheta, \varphi, \chi) = \frac{1}{2}\int \mathrm{d}^2x \left[\left(\partial_\mu\vartheta\right)^2 + \frac{1}{g}\left(\partial_\mu\chi\right)^2 - \left(\frac{1}{g} + \frac{1}{\pi}\right)\left(\partial_\mu\varphi\right)^2\right]. \tag{32.94}$$

The massless Thirring model can thus be solved exactly and we calculate a few correlation functions in Section 32.5.2.

For $m_0 \neq 0$, we change variables $\varphi \mapsto \theta$, setting

$$\sqrt{4\pi}\vartheta + 2\varphi = \theta \,.$$

The integral over ϑ is then gaussian,

$$\mathcal{S}(\vartheta,\theta)=\frac{1}{2}\int \mathrm{d}^2x\left[(\partial_\mu\vartheta)^2-\frac{2m_0\Lambda}{\pi}\cos\theta-\frac{\pi+g}{4\pi g}(\partial_\mu\theta-\sqrt{4\pi}\partial_\mu\vartheta)^2\right],$$

and can be explicitly performed. We finally obtain

$$\mathcal{S}(\theta)=\int \mathrm{d}^2x\left[\frac{1}{8\pi}\left(1+\frac{g}{\pi}\right)(\partial_\mu\theta)^2-\frac{m_0\Lambda}{\pi}\cos\theta\right], \tag{32.95}$$

which establishes the correspondence between the massive Thirring model and the SG model. Comparing with the action in the normalization (32.65), we see that the relation between parameters is ($t=\kappa^2$)

$$\frac{4\pi}{t}=1+\frac{g}{\pi},$$

proving relations (32.89a,c). Since the coupling constant t is not renormalized this correspondence establishes that the coupling constant g is not renormalized either and thus the RG $\beta(g)$-function vanishes.

The field $\mathrm{e}^{\pm 2i\varphi}$ instead has to be renormalized, the fermion mass thus needs a renormalization. With the regularization and renormalization at mass scale μ of Section 32.1, we find

$$m_0=m(\mu/\Lambda)^{g/(g+\pi)}.$$

The relation between renormalized parameters is then

$$\alpha/\kappa^2=m\mu/\pi\,.$$

We now briefly discuss the special massless case.

32.5.2 The massless Thirring model

The massless Thirring model has a $U(1)$ chiral symmetry. Adding a source term B_μ for the current J_μ we find the effective action $\mathcal{S}(A_\mu)$ for the remaining A_μ field:

$$\mathcal{S}(A_\mu)=\frac{1}{2}\int \mathrm{d}^2x\left[\frac{1}{g}A_\mu^2+\frac{1}{\pi}(A_\mu-iB_\mu)\left(\delta_{\mu\nu}-\frac{\partial_\mu\partial_\nu}{\partial^2}\right)(A_\nu-iB_\nu)\right]. \tag{32.96}$$

We can then integrate over the A_μ field to obtain the current two-point correlation function. In momentum space

$$\langle J_\mu(k)J_\nu(-k)\rangle=\frac{1}{\pi+g}\left(\delta_{\mu\nu}-\frac{k_\mu k_\nu}{k^2}\right). \tag{32.97}$$

All other connected correlation functions vanish. If we instead calculate the two-point correlation function of $\epsilon_{\mu\nu}\partial_\nu\theta$ with the free action (32.95) for $m_0=0$, we find

$$\langle\epsilon_{\mu\rho}\partial_\rho\theta(k)\epsilon_{\nu\sigma}\partial_\sigma\theta(-k)\rangle=\frac{4\pi^2}{\pi+g}\left(\delta_{\mu\nu}-\frac{k_\mu k_\mu}{k^2}\right), \tag{32.98}$$

result consistent with the two relations (32.89a,b).

All ψ-field correlation functions can be calculated explicitly. In particular, the bosonized form of the $\bar\psi(x)\psi(x)$ correlation functions depends only on $\theta(x)$ and thus can be calculated with the action (32.95) for $m_0=0$. The one-point function $\bar\psi(x)\psi(x)$ vanishes. The $\bar\psi\psi$ two-point function is given by

$$\left\langle\bar\psi(x)\psi(x)\bar\psi(0)\psi(0)\right\rangle\propto\langle\cos\theta(x)\cos\theta(0)\rangle\propto x^{-\lambda}$$

with $\lambda=t/2\pi=2\pi/(g+\pi)$.

32.5.3 Discussion

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 decreases only algebraically.

As we have discussed in the case of the SG model, at the value $\kappa^2 = 8\pi$, that is, $g = \pi/2$, something new happens, the mass term becomes irrelevant. This suggests the possibility of a phase transition, a question we examine in the next chapter. The transition point, however, 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 $\kappa^2 = 8\pi$ on a free model by doubling the number of fermions. We indeed exhibit in the next section a two fermion model which is also equivalent to the SG model.

Mass spectrum. The correspondence between the Thirring and SG models allows, in particular, to calculate physical quantities in the SG model for g small, that is, for t close to 4π, or in the Thirring model for g large. Moreover, we see that for $g > 0$, that is, $t < 4\pi$, where the potential between fermions is attractive, the spectrum of the theory consists at least in one Dirac fermion and a 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

$$\begin{aligned} m_n &= 2m(\alpha,t)\sin(n\gamma/16)\,, \quad n = 1,2\ldots < 8\pi/\gamma \\ \text{with}\quad \gamma &= \frac{t}{1-t/8\pi} = \frac{8\pi}{1+2g/\pi}\,, \end{aligned}$$

where $m(\alpha,t)$ is the fermion mass (see also equation (33.21)):

$$m(\alpha,t) = f(t)\frac{8}{\gamma}\alpha^{1/2+\gamma/16\pi}, \quad f(0) = 1\,. \tag{32.99}$$

The lowest mass $n = 1$ corresponds to the SG field; the remaining boundstates can also be considered as collective excitations of the SG particle.

32.6 A Two-Fermion Model

The derivation of the correspondence between the massive Thirring and the SG models indicates that it is possible to construct a fermion model equivalent to the SG model, such that $\kappa^2 = 8\pi$ corresponds to a free field theory, by introducing two Dirac fermions and by choosing the four-point interaction appropriately. We now describe such a model.

32.6.1 The model

We consider a generalized massless Thirring model with two fermions coupled through a current–current interaction. The action is

$$\mathcal{S}(\bar{\psi}^a,\psi^a) = -\int \mathrm{d}^2x\,\left[\bar{\psi}^a \not{\partial}\psi^a - \tfrac{1}{2}g_{ab}J_\mu^a J_\mu^b - f\bar{\psi}^1\gamma_\mu\psi^2\bar{\psi}^2\gamma_\mu\psi^1\right], \tag{32.100}$$

where the matrix g_{ab} has the form

$$\mathbf{g} = \frac{1}{2}\begin{pmatrix} g' + g & g' - g \\ g' - g & g' + g \end{pmatrix}, \tag{32.101}$$

and the current j^a_μ is

$$j^1_\mu = \bar{\psi}^1\gamma_\mu\psi^1, \quad j^2_\mu = \bar{\psi}^2\gamma_\mu\psi^2. \tag{32.102}$$

This action is the most general renormalizable action with the following symmetries:
(i) A $U(1)$ chiral symmetry which prevents a fermion mass term:

$$\begin{aligned} \psi^a(x) &\mapsto \mathrm{e}^{i\gamma_S\theta}\,\psi^a(x), \\ \bar{\psi}^a(x) &\mapsto \bar{\psi}^a\,\mathrm{e}^{i\gamma_S\theta}\,. \end{aligned} \tag{32.103}$$

(ii) Separate fermion number conservation for each type of fermions and finally symmetry between fermions 1 and 2.

Note that the model has an additional $SU(2)$ invariance when $g = f$ since the interaction can also be written as

$$-\tfrac{1}{4}\left[g' J^0_\mu J^0_\mu + g J^3_\mu J^3_\mu + f\left(J^1_\mu J^1_\mu + J^2_\mu J^2_\mu\right)\right],$$

where J^0_μ and $\mathbf{J}_\mu$ are the $U(1)$ and the $SU(2)$ currents, respectively,

$$J^0_\mu = \bar{\psi}^a\gamma_\mu\psi^a, \qquad \mathbf{J}_\mu = \bar{\psi}^a\boldsymbol{\sigma}_{ab}\gamma_\mu\psi^b\,.$$

We have denoted by $\boldsymbol{\sigma}$ the three Pauli matrices acting on the $SU(2)$ indices.

Due to the $U(1)$ chiral invariance of action (32.100), the coupling g' is not renormalized, as will be shown in Section 32.6.2. The RG β-functions at one-loop order are (see Appendix A32.2.2)

$$\begin{aligned} \beta_g &= -2f^2/\pi\,, \\ \beta_f &= -2fg/\pi\,. \end{aligned} \tag{32.104}$$

The equivalence with the SG model is derived below. The correspondence is summarized by the relations:

$$8\pi/\kappa^2 = 1 + g/\pi\,, \qquad \alpha_0 = \kappa^2\frac{f\Lambda^2}{\pi^2}\,, \tag{32.105}$$

$$\cos\theta \mapsto \frac{\pi^2}{\Lambda^2}\bar{\psi}^1\gamma_\mu\psi^2\bar{\psi}^2\gamma_\mu\psi^1. \tag{32.106}$$

From these expressions, we see that for t 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 taking place at $g = f = 0$ is discussed extensively in Chapter 33, we here note only 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.

32.6.2 Derivation

The general idea is to separate the action (32.100),

$$\mathcal{S}(\bar{\psi}^a, \psi^a) = -\int \mathrm{d}^2x \left(\bar{\psi}^a \not{\partial} \psi^a - \tfrac{1}{2} g_{ab} j^a_\mu j^b_\mu - f \bar{\psi}^1 \gamma_\mu \psi^2 \bar{\psi}^2 \gamma_\mu \psi^1\right),$$

into the sum of two terms $\mathcal{S} = \mathcal{S}_0 + \mathcal{S}_1$:

$$\begin{aligned}\mathcal{S}_0(\bar{\psi}^a, \psi^a) &= -\int \mathrm{d}^2x \left(\bar{\psi}^a \not{\partial} \psi^a - \tfrac{1}{2} g_{ab} j^a_\mu j^b_\mu\right),\\ \mathcal{S}_1(\bar{\psi}^a, \psi^a) &= f \int \mathrm{d}^2x\, \bar{\psi}^1 \gamma_\mu \psi^2 \bar{\psi}^2 \gamma_\mu \psi^1 .\end{aligned}$$

The first term $\mathcal{S}_0$ contains all interaction terms which have a separate chiral invariance for each species of fermions (interaction of Thirring model type) and is treated in much the same way as the massless Thirring action: it can be transformed into a free field action.

The remainder $\mathcal{S}_1$ is only invariant with respect to the same chiral transformation for ψ_1 and ψ_2. It is expanded in perturbation theory like the mass term in the Thirring model, and eventually yields the interaction.

We thus write the interaction terms of $\mathcal{S}_0$ as generated by gaussian integrals over two vector fields $A^\pm_\mu$. The action then becomes

$$\mathcal{S}_0(A^\pm_\mu, \bar{\psi}, \psi) = -\int \mathrm{d}^2x \left[\bar{\psi}\left(\not{\partial} + i\not{A}^+ + i\sigma_3 \not{A}^-\right)\psi - \left(A^+_\mu\right)^2/g' - \left(A^-_\mu\right)^2/g\right],$$

where $[\sigma_3]_{ab}$ is a Pauli matrix and acts on the vector ψ^b. Parametrizing the vector fields as in equation (32.50):

$$A^\pm_\mu = -\partial_\mu \chi^\pm - i\epsilon_{\mu\nu}\partial_\nu \varphi^\pm , \tag{32.107}$$

and performing the corresponding gauge transformations (32.52) on both fermion fields we obtain (see equation (32.92))

$$\begin{aligned}\mathcal{S}_0 = -\int \mathrm{d}^2x \Big[&\bar{\psi}^a \not{\partial} \psi^a + \frac{\pi + g}{\pi g}\left(\partial_\mu \varphi^-\right)^2 + \frac{\pi + g'}{\pi g'}\left(\partial_\mu \varphi^+\right)^2 - \frac{1}{g}\left(\partial_\mu \chi^-\right)^2 \\ &- \frac{1}{g'}\left(\partial_\mu \chi^+\right)^2\Big]\end{aligned} \tag{32.108}$$

(we have omitted the primes on the fermions), which shows that g' indeed is not renormalized.

For $\mathcal{S}_1$, to be able to use directly the identities derived for the Thirring model, it is convenient to rewrite the f-term using a Fierz transformation (see Appendix A8.4):

$$\bar{\psi}^1 \gamma_\mu \psi^2 \bar{\psi}^2 \gamma_\mu \psi^1 = -\bar{\psi}^1 \psi^1 \bar{\psi}^2 \psi^2 + \bar{\psi}^1 \gamma_S \psi^1 \bar{\psi}^2 \gamma_S \psi^2 . \tag{32.109}$$

Introducing the chiral components of the mass operators

$$\varsigma^1_\pm(x) = \bar{\psi}^1_\mp(x)\psi^1_\pm(x), \quad \varsigma^2_\pm(x) = \bar{\psi}^2_\mp(x)\psi^2_\pm(x),$$

we can rewrite the r.h.s.:

$$-\bar{\psi}^1 \psi^1 \bar{\psi}^2 \psi^2 + \bar{\psi}^1 \gamma_S \psi^1 \bar{\psi}^2 \gamma_S \psi^2 = -2\left(\varsigma^1_+ \varsigma^2_- + \varsigma^2_+ \varsigma^1_-\right). \tag{32.110}$$

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 (32.52), the chiral components $\varsigma^a_\pm$ become

$$\varsigma^1_\pm \mapsto \varsigma^1_\pm \, \mathrm{e}^{\pm 2i(\varphi_+ + \varphi_-)}, \qquad \varsigma^2_\pm \mapsto \varsigma^2_\pm \, \mathrm{e}^{\pm 2i(\varphi_+ - \varphi_-)} .$$

Then, using identity (32.110), we can rewrite the f term:

$$f \bar{\psi}^1 \, \mathrm{e}^{2i\gamma_S \varphi^-} \, \gamma_\mu \psi^2 \bar{\psi}^2 \, \mathrm{e}^{-2i\gamma_S \varphi^-} \, \gamma_\mu \psi^1 = -2f \left(\mathrm{e}^{4i\varphi_-} \, \varsigma^1_+ \varsigma^2_- + \mathrm{e}^{-4i\varphi_-} \, \varsigma^2_+ \varsigma^1_- \right) . \tag{32.111}$$

We now associate two bosons ϑ^a with the fermions $\bar{\psi}^a, \psi^a$. In the expansion in powers of f, we see that we can simply replace the quantities $\varsigma^a_\pm$ by their bosonic counterparts. We then obtain a boson action

$$\begin{aligned} \mathcal{S}_{\text{bos}} = \int \mathrm{d}^2 x \Big\{ & \frac{1}{2} \partial_\mu \vartheta^a \partial_\mu \vartheta^a - \frac{f\Lambda^2}{\pi^2} \cos \left[\sqrt{4\pi} \left(\vartheta^1 - \vartheta^2 \right) + 4\varphi_- \right] \\ & - \frac{\pi + g}{\pi g} \left(\partial_\mu \varphi^- \right)^2 - \frac{\pi + g'}{\pi g'} \left(\partial_\mu \varphi^+ \right)^2 + \frac{1}{g} \left(\partial_\mu \chi^- \right)^2 + \frac{1}{g'} \left(\partial_\mu \chi^+ \right)^2 \Big\} . \end{aligned}$$

Changing variables $\vartheta^i \mapsto \theta_i$:

$$\theta_1 = \sqrt{4\pi} \vartheta^1 + 2(\varphi_+ + \varphi_-), \quad \theta_2 = \sqrt{4\pi} \vartheta^2 + 2(\varphi_+ - \varphi_-),$$

and integrating over the fields $\varphi^\pm, \chi_\pm$, we find a boson action

$$\begin{aligned} \mathcal{S}(\theta) = \int \mathrm{d}^2 x \Big[& \frac{1}{16\pi} \left(1 + \frac{g}{\pi} \right) \left[\partial_\mu (\theta_1 - \theta_2) \right]^2 + \frac{1}{16\pi} \left[\partial_\mu (\theta_1 + \theta_2) \right]^2 \\ & - \frac{f\Lambda^2}{\pi^2} \cos(\theta_1 - \theta_2) \Big] , \end{aligned} \tag{32.112}$$

which in terms of $\theta_\pm = \theta_1 \pm \theta_2$ is the sum of a SG action and a massless decoupled scalar action

$$\mathcal{S}(\theta) = \int \mathrm{d}^2 x \left[\frac{1}{16\pi} \left(1 + \frac{g}{\pi} \right) \left(\partial_\mu \theta_- \right)^2 - \frac{f\Lambda^2}{\pi^2} \cos \theta_- + \frac{1}{16\pi} \left(\partial_\mu \theta_+ \right)^2 \right] . \tag{32.113}$$

A comparison with the form (32.65) shows that the SG model has a parameter κ:

$$8\pi/\kappa^2 = 1 + g/\pi . \tag{32.114}$$

Therefore, the value $\kappa^2 = 8\pi$ corresponds to a free massless fermion theory.

Finally, for $\kappa^2 < 8\pi$, although the remaining spectrum is massive, chiral symmetry is not broken, simply the corresponding field θ^+ is massless and free and thus decouples.

Bibliographical Notes

Some references on the sine-Gordon model and its exact integrability are

L.D. Faddeev and L.A. Takhtajan, *Teor. Mat. Fiz.* 21 (1974) 160; R.F. Dashen, B. Hasslacher and A. Neveu, *Phys. Rev.* D11 (1975) 3424; L.D. Faddeev and V. Korepin, *Phys. Rep.* 42C (1978) 1; A.B. Zamolodchikov and Al. B. Zamolodchikov, *Ann. Phys.* 120 (1979) 253.

Some references on the Schwinger model and its relation to confinement are

J. Schwinger, *Phys. Rev.* 128 (1962) 2425; J.H. Lowenstein and J.A. Swieca, *Ann. Phys. (NY)* 68 (1971) 172; A. Casher, J. Kogut and L. Susskind, *Phys. Rev.* D10 (1974) 732; S. Coleman, R. Jackiw and L. Susskind, *Ann. Phys. (NY)* 93 (1975) 267; S. Coleman, *Ann. Phys. (NY)* 101 (1976) 239.

The correlation functions of the massless Thirring model are given in

R. Klaiber, *Lectures in Theoretical Physics* 10A (1967) 141, A. Barut and W. Brittin eds. (Gordon and Breach, London 1968).

The correspondence between the massive Thirring and the sine-Gordon models has been established in

S. Coleman, *Phys. Rev.* D11 (1975) 208 8.

For the bosonization of the $SU(N)$ Thirring model, see

T. Banks, D. Horn and H. Neuberger, *Nucl. Phys.* B108 (1976) 119.

For further references on bosonization techniques, see

E. Witten, *Commun. Math. Phys.* 92 (1984) 455; I. Affleck, *String, Fields and Critical Phenomena*, Les Houches 1988, E. Brézin and J. Zinn-Justin eds. (Elsevier, Amsterdam 1989).

For a discussion of the solitons in the sine-Gordon model see, for example,

R. Rajaraman, *Solitons and Instantons* (North-Holland, Amsterdam 1982),

and references therein. The spectrum and the S-matrix of the $SU(N)$ Nambu–Jona-Lasinio model are discussed in

J.H. Lowenstein, *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, J.-B. Zuber and R. Stora eds. (Elsevier 1984).

APPENDIX A32
A FEW ADDITIONAL RESULTS

A32.1 The Schwinger Model

In this section, we examine the role of the anomaly in the Schwinger model from two other point of views.

A32.1.1 The general anomaly

The anomaly of the axial current $J_\mu^S = i\bar{\psi}\gamma_S\gamma_\mu\psi$ is (equation (20.101))

$$\partial_\mu J_\mu^S = -\frac{ie}{2\pi}\epsilon_{\mu\nu}F_{\mu\nu}\,.$$

On the other hand, the vector current $J_\mu = \bar{\psi}\gamma_\mu\psi$ is exactly conserved:

$$\partial_\mu J_\mu = 0\,.$$

We have observed that in dimension 2, the two currents are related (equation (32.25)),

$$J_\mu^S = -\epsilon_{\mu\nu}J_\nu\,,$$

and thus

$$\partial_\mu J_\nu - \partial_\nu J_\mu = \frac{ie}{\pi}F_{\mu\nu}\,.$$

The gauge field equation of motion yields

$$ieJ_\mu + \partial_\nu F_{\nu\mu} = 0\,.$$

Combining these equations, we get

$$\left(-\partial^2 + (e^2/\pi)\right)J_\mu = 0\,.$$

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 32.4.1.

A32.1.2 The Schwinger model: one-loop calculation

We know from the general analysis that we can regularize the fermion determinant while preserving the QED gauge invariance. However, then we cannot preserve in general the chiral gauge invariance. For example, in dimensional regularization, the anticommutation properties of γ_S together with the relation (32.24) cannot be maintained (see Section 9.6.2). We here explicitly work out the one-loop contribution to the A_μ field two-point function (figure 32.1) in the Feynman gauge. Using dimensional regularization which preserves non-chiral gauge invariance and giving temporarily a mass M to the fermion field to avoid IR problems, we find

$$\Gamma^{(2)}_{\mu\nu}(k) = k^2\delta_{\mu\nu} - e^2\int\frac{\mathrm{d}^d q}{(2\pi)^d}\,\mathrm{tr}\left[\gamma_\mu\frac{M - i\not{q}}{M^2+q^2}\gamma_\nu\frac{M - i(\not{q}+\not{k})}{M^2+(q+k)^2}\right]. \qquad (A32.1)$$

APPENDIX A32
A FEW ADDITIONAL RESULTS

A32.1 The Schwinger Model

In this section, we examine the role of the anomaly in the Schwinger model from two other point of views.

A32.1.1 The general anomaly

The anomaly of the axial current $J_\mu^S = i\bar{\psi}\gamma_S\gamma_\mu\psi$ is (equation (20.101))

$$\partial_\mu J_\mu^S = -\frac{ie}{2\pi}\epsilon_{\mu\nu}F_{\mu\nu}\,.$$

On the other hand, the vector current $J_\mu = \bar{\psi}\gamma_\mu\psi$ is exactly conserved:

$$\partial_\mu J_\mu = 0\,.$$

We have observed that in dimension 2, the two currents are related (equation (32.25)),

$$J_\mu^S = -\epsilon_{\mu\nu}J_\nu\,,$$

and thus

$$\partial_\mu J_\nu - \partial_\nu J_\mu = \frac{ie}{\pi}F_{\mu\nu}\,.$$

The gauge field equation of motion yields

$$ieJ_\mu + \partial_\nu F_{\nu\mu} = 0\,.$$

Combining these equations, we get

$$\left(-\partial^2 + (e^2/\pi)\right)J_\mu = 0\,.$$

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 32.4.1.

A32.1.2 The Schwinger model: one-loop calculation

We know from the general analysis that we can regularize the fermion determinant while preserving the QED gauge invariance. However, then we cannot preserve in general the chiral gauge invariance. For example, in dimensional regularization, the anticommutation properties of γ_S together with the relation (32.24) cannot be maintained (see Section 9.6.2). We here explicitly work out the one-loop contribution to the A_μ field two-point function (figure 32.1) in the Feynman gauge. Using dimensional regularization which preserves non-chiral gauge invariance and giving temporarily a mass M to the fermion field to avoid IR problems, we find

$$\Gamma^{(2)}_{\mu\nu}(k) = k^2\delta_{\mu\nu} - e^2\int\frac{\mathrm{d}^d q}{(2\pi)^d}\,\mathrm{tr}\left[\gamma_\mu\frac{M - i\not{q}}{M^2+q^2}\gamma_\nu\frac{M - i(\not{q}+\not{k})}{M^2+(q+k)^2}\right]. \qquad (A32.1)$$

APPENDIX A32
A FEW ADDITIONAL RESULTS

A32.1 The Schwinger Model

In this section, we examine the role of the anomaly in the Schwinger model from two other point of views.

A32.1.1 The general anomaly

The anomaly of the axial current $J^S_\mu = i\bar{\psi}\gamma_S\gamma_\mu\psi$ is (equation (20.101))

$$\partial_\mu J^S_\mu = -\frac{ie}{2\pi}\epsilon_{\mu\nu}F_{\mu\nu}\,.$$

On the other hand, the vector current $J_\mu = \bar{\psi}\gamma_\mu\psi$ is exactly conserved:

$$\partial_\mu J_\mu = 0\,.$$

We have observed that in dimension 2, the two currents are related (equation (32.25)),

$$J^S_\mu = -\epsilon_{\mu\nu}J_\nu\,,$$

and thus

$$\partial_\mu J_\nu - \partial_\nu J_\mu = \frac{ie}{\pi}F_{\mu\nu}\,.$$

The gauge field equation of motion yields

$$ieJ_\mu + \partial_\nu F_{\nu\mu} = 0\,.$$

Combining these equations, we get

$$\left(-\partial^2 + (e^2/\pi)\right)J_\mu = 0\,.$$

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 32.4.1.

A32.1.2 The Schwinger model: one-loop calculation

We know from the general analysis that we can regularize the fermion determinant while preserving the QED gauge invariance. However, then we cannot preserve in general the chiral gauge invariance. For example, in dimensional regularization, the anticommutation properties of γ_S together with the relation (32.24) cannot be maintained (see Section 9.6.2). We here explicitly work out the one-loop contribution to the A_μ field two-point function (figure 32.1) in the Feynman gauge. Using dimensional regularization which preserves non-chiral gauge invariance and giving temporarily a mass M to the fermion field to avoid IR problems, we find

$$\Gamma^{(2)}_{\mu\nu}(k) = k^2\delta_{\mu\nu} - e^2\int\frac{\mathrm{d}^d q}{(2\pi)^d}\,\mathrm{tr}\left[\gamma_\mu\frac{M - i\not{q}}{M^2+q^2}\gamma_\nu\frac{M - i(\not{q}+\not{k})}{M^2+(q+k)^2}\right]. \qquad (A32.1)$$

It is easy to verify that the dimensionally regularized one-loop contribution is transverse as expected. Setting

$$\int \frac{\mathrm{d}^d q}{(2\pi)^d} \operatorname{tr}\left[\gamma_\mu \frac{M - i\not{q}}{M^2+q^2}\gamma_\nu \frac{M - i(\not{q}+\not{k})}{M^2+(q+k)^2}\right] = D\left(\delta_{\mu\nu} - \frac{k_\mu k_\mu}{k^2}\right), \tag{A32.2}$$

and taking the trace of both members we obtain

$$D(k)(d-1) = \operatorname{tr}\mathbf{1} \int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{dM^2 + (d-2)\left(q^2 + q\cdot k\right)}{\left(q^2+M^2\right)\left[(q+k)^2+M^2\right]}. \tag{A32.3}$$

The r.h.s. contains a term proportional to M^2 which is UV convergent and which goes to zero 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:

$$\int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{q^2}{\left(q^2+M^2\right)^2} \underset{d\to 2}{\sim} \frac{1}{2\pi(2-d)} \quad \Rightarrow \quad D = -1/\pi\,. \tag{A32.4}$$

The gauge field inverse two-point function at one-loop order follows

$$\Gamma^{(2)}_{\mu\nu}(k) = \left(k^2\delta_{\mu\nu} - k_\mu k_\nu\right)\left(1 + \frac{e^2}{\pi k^2}\right) + k_\mu k_\nu\,, \tag{A32.5}$$

a result consistent with the exact expression (32.77).

A32.2 The $SU(N)$ Thirring Model

In Section 32.6, 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 of the property that more general models can be bosonized. We here consider in particular the $SU(N)$ Thirring model, a model described in terms of N Dirac fermions, with the action

$$\begin{aligned} \mathcal{S}(\bar\psi,\psi) &= -\int \mathrm{d}^2 x \left[\bar\psi\cdot\not{\partial}\psi - \tfrac{1}{2}f\left(\bar\psi\cdot\gamma_\mu\psi\right)^2 - \tfrac{1}{2}g\bar\psi^i\gamma_\mu\psi^j\bar\psi^j\gamma_\mu\psi^i\right] & (A32.6)\\ &= -\int \mathrm{d}^2 x \left[\bar\psi\cdot\not{\partial}\psi - \frac{1}{2N}\left((g+Nf)J^0_\mu J^0_\mu + gJ^\alpha_\mu J^\alpha_\mu\right)\right], & (A32.7)\end{aligned}$$

where J^0_μ and J^α_μ are the $U(1)$ and $SU(N)$ currents:

$$J^0_\mu = \bar\psi\cdot\gamma_\mu\psi\,,\quad J^\alpha_\mu = \bar\psi^i\gamma_\mu t^\alpha_{ij}\psi^j\,.$$

This model has a $SU(N)\times U(1)$ vector together with a chiral $U(1)$ symmetry. After a Fierz transformation, the second interaction term becomes

$$\bar\psi^i\gamma_\mu\psi^j\bar\psi^j\gamma_\mu\psi^i = -\bar\psi^i\psi^i\bar\psi^j\psi^j + \bar\psi^i\gamma_S\psi^i\bar\psi^j\gamma_S\psi^j\,.$$

We then recognize the action of the Nambu–Jona-Lasinio model. Strictly speaking, this model corresponds to the special case $f=0$ but such a model is not multiplicatively renormalizable. Another current interaction has to be added for renormalization purpose. Due to the global chiral invariance of the model the coupling constant $g+Nf$ associated with the current J^0_μ is not renormalized. The remaining RG β_g-function in $d = 2 + \varepsilon$ at one-loop order then is

$$\beta_g = \varepsilon g - \frac{N}{\pi} g^2 , \tag{A32.8}$$

showing that the model is asymptotically free and has a non-perturbative spectrum for g positive and small. For $g < 0$ instead the model is IR free: the fermions remain massless.

The model can be bosonized like the Thirring model to yield a generalized SG model of the form

$$\mathcal{S}(\boldsymbol{\theta}) = \int \mathrm{d}^2 x \left\{ \frac{1}{4\pi^2} \left[(\pi + g) \textstyle\sum_i \left(\partial_\mu \theta^i\right)^2 + f \left(\sum_i \partial_\mu \theta^i\right)^2 \right] - G \textstyle\sum_{i,j} \cos(\theta^i - \theta^j) \right\} \tag{A32.9}$$

with

$$G = g \frac{\Lambda^2}{\pi^2} . \tag{A32.10}$$

The physics of the $SU(N)$ model is again the physics of spontaneous mass generation. In perturbation theory, we see only massless fermions but for $g > 0$ (where the forces between fermions are attractive) the spectrum contains massive particles. Chiral symmetry, however, 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 32.1 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 $\theta_i - \theta_j$ involve only massive particles. This another example of the Kosterlitz–Thouless mechanism (see Chapter 33).

Note that the spectrum has been obtained exactly

$$m_n \propto \frac{N}{\pi} \sin(n\pi/N), \quad n = 1, 2 \ldots < N ,$$

n odd corresponding to fermions and n even to bosons. This result confirms that fermions become massive.

A32.2.1 Derivation

The method is a simple extension of the method explained in Section 32.6 in the $SU(2)$ case. Introducing the currents

$$j^i_\mu = \bar{\psi}^i \gamma_\mu \psi^i , \tag{A32.11}$$

(no sum over i) we can write the action ($A32.6$) as

$$\mathcal{S}(\bar{\psi}, \psi) = - \int \mathrm{d}^2 x \left[\bar{\psi} \cdot \not{\partial} \psi - \frac{1}{2} g_{ij} j^i_\mu j^j_\mu - \frac{1}{2} g \sum_{i \neq j} \bar{\psi}^i \gamma_\mu \psi^j \bar{\psi}^j \gamma_\mu \psi^i \right] \tag{A32.12}$$

with

$$g_{ij} = g \delta_{ij} + f . \tag{A32.13}$$

Introducing vector fields A^i_μ, we transform the action into

$$\mathcal{S}(\bar{\psi}, \psi, \mathbf{A}_\mu) = - \int \mathrm{d}^2 x \left[\sum_i \bar{\psi}^i \left(\not{\partial} + i \not{A}^i \right) \psi^i - \frac{1}{2} \tilde{g}_{ij} A^i_\mu A^j_\mu - \frac{1}{2} g \sum_{i \neq j} \bar{\psi}^i \gamma_\mu \psi^j \bar{\psi}^j \gamma_\mu \psi^i \right] \tag{A32.14}$$

with

$$\tilde{g}_{ij} \equiv [g^{-1}]_{ij} = \frac{1}{g}\left(\delta_{ij} - \frac{f}{g+Nf}\right). \qquad (A32.15)$$

Combining the usual $U(1)$ gauge transformations on the fermions with a parametrization of the vector fields,

$$\begin{aligned}\psi^i(x) &\mapsto \mathrm{e}^{i\chi^i(x)+i\gamma_S\varphi^i(x)}\,\psi^i(x),\\ \bar{\psi}^i(x) &\mapsto \bar{\psi}^i\,\mathrm{e}^{-i\chi^i(x)+i\gamma_S\varphi^i(x)},\\ A^i_\mu &= -\partial_\mu\chi^i - i\epsilon_{\mu\nu}\partial_\nu\varphi^i\,,\end{aligned}$$

we find

$$\begin{aligned}\mathcal{S}(\bar{\psi},\psi,\varphi,\chi) = \int \mathrm{d}^2x\Big[&-\bar{\psi}^i\not{\partial}\psi^i + \frac{1}{2}\tilde{g}_{ij}\partial_\mu\chi^i\partial_\mu\chi^j - \frac{1}{2}\left(\tilde{g}_{ij}+\frac{\delta_{ij}}{\pi}\right)\partial_\mu\varphi^i\partial_\mu\varphi^j\\ &-\frac{1}{2}g\sum_{i\neq j}\bar{\psi}^i\,\mathrm{e}^{i\gamma_S(\varphi^i-\varphi^j)}\,\gamma_\mu\psi^j\bar{\psi}^j\,\mathrm{e}^{-i\gamma_S(\varphi^i-\varphi^j)}\,\gamma_\mu\psi^i\Big]. \qquad (A32.16)\end{aligned}$$

In terms of the chiral components of the mass operator $\bar{\psi}^i\psi^i$:

$$\varsigma^i_\pm = \bar{\psi}^i_\pm\psi^i_\pm\,,$$

and using the identity (32.111), we can write

$$\bar{\psi}^i\,\mathrm{e}^{i\gamma_S(\varphi^i-\varphi^j)}\,\gamma_\mu\psi^j\bar{\psi}^j\,\mathrm{e}^{-i\gamma_S(\varphi^i-\varphi^j)}\,\gamma_\mu\psi^i \underset{i\neq j}{=} -2\left(\mathrm{e}^{2i(\varphi^i-\varphi^j)}\,\varsigma^i_+\varsigma^j_- + \mathrm{e}^{2i(\varphi^j-\varphi^i)}\,\varsigma^j_+\varsigma^i_-\right).$$

We then expand the four-fermion interaction in perturbation theory. As in the $SU(2)$ model, we see that the average over the fermions $\bar{\psi}^i,\psi^i$ can be replaced by an average over bosons ϑ^i. After summation, this leads to a SG interaction of the form

$$-g\frac{\Lambda^2}{\pi^2}\sum_{i\neq j}\cos\left[\sqrt{4\pi}\left(\vartheta^i-\vartheta^j\right)+2\left(\varphi^i-\varphi^j\right)\right].$$

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 find the action ($A32.9$).

A32.2.2 Four-fermion current interactions: RG β-function

We leave as an exercise to calculate the RG β-function in the case of a four-fermion, current–current, interaction:

$$\mathcal{S}(\bar{\psi}^a,\psi^a) = -\int \mathrm{d}^2x\left(\bar{\psi}^a\not{\partial}\psi^a - \tfrac{1}{2}g_{abcd}\bar{\psi}^a\gamma_\mu\psi^b\bar{\psi}^c\gamma_\mu\psi^d\right), \qquad (A32.17)$$

where g_{abcd} is chosen to be symmetric in the exchange $(ab)\leftrightarrow(cd)$. The result is

$$\pi\beta_{abcd} = g_{aicj}\,g_{ibjd} - g_{aijd}\,g_{ibcj} + O\left(g^3\right). \qquad (A32.18)$$

A32.3 Solitons in the Sine-Gordon 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 also time-dependent solutions can be obtained. 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 (32.64) is a differential equation in the space coordinate x:

$$\theta'' = \alpha \sin\theta \,.$$

Finite energy solutions necessarily connect minima of the potential. Thus,

$$\theta = 4 \arctan \mathrm{e}^{\sqrt{\alpha} x} \,.$$

The corresponding energy then is the space integral of the lagrangian density (after subtraction of the vacuum energy):

$$M_{\text{sol.}} = \frac{1}{t} \int \mathrm{d}x \left[\tfrac{1}{2} \left(\partial_x \theta\right)^2 + \alpha(1-\cos\theta)\right] = 8\sqrt{\alpha}/t + O(1) \,,$$

which coincides with the Thirring fermion mass (32.99) for t small.

33 THE $O(2)$ CLASSICAL SPIN MODEL IN TWO DIMENSIONS

Having established in Chapter 32 a few properties of two-dimensional models, we now have the necessary technical background to discuss the abelian $O(2)$ spin model near and in two dimensions. As we have seen in Chapter 31, at low temperature its long distance properties can be described in terms of the $O(2)$ non-linear σ-model. We recall that the $O(2)$ case is special because the RG β-function reduces in the low temperature expansion to the dimensional term $(d-2)t$ and, therefore, the properties, from the RG point of view, are quite different. In particular, in two dimensions, 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 is locally a flat manifold, that is, which cannot be distinguished from a straight line. Therefore, if we parametrize the spin $\mathbf{S}(x)$ as

$$\mathbf{S}(x)=\begin{cases}\cos\theta(x)\\ \sin\theta(x)\end{cases}, \tag{33.1}$$

the euclidean action (31.12) in zero field becomes

$$\mathcal{S}(\theta)=\frac{\Lambda^{d-2}}{2t}\int \mathrm{d}^d x\,[\partial_\mu\theta(x)]^2\,, \tag{33.2}$$

that is, a free massless field action. This explains the form of the RG β-function. Nevertheless, because the physical fields are $\sin\theta$ or $\cos\theta$ or equivalently $\mathrm{e}^{\pm i\theta}$ rather than θ itself, a field renormalization is necessary and thus temperature-dependent anomalous dimensions are generated (equations (31.27*b*) and (32.12)).

The simplicity of the action (33.2), however, leads to a mystery: since the field θ is massless the correlation length remains infinite for all t. By contrast, a simple high temperature analysis of the corresponding spin model on the lattice shows that the correlation length is finite for t large enough. A phase transition at finite temperature is necessary to explain this phenomenon. The action (33.2), therefore, cannot represent by itself the long distance properties of the lattice model for all temperatures.

As the one-dimensional example of Section 3.3 indicates, it is necessary to take in some way into account the property that θ is a cyclic variable. This condition, which is not obviously incorporated in the action (33.2), is irrelevant at low temperature, 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 Kosterlitz–Thouless (KT) phase transition, which separates an infinite correlation length phase *without order* (the low temperature phase of the $O(2)$ model) from a finite correlation length phase. Additional information can be obtained from the equivalence between the SG model and several fermion models, established in Sections 32.5,32.6.

33.1 The Spin Correlation Functions

We first calculate the spin correlation functions in d dimensions with the free action (33.2):

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\epsilon_i\theta(x_i)} \right\rangle = \int [\mathrm{d}\theta] \exp\left[-\frac{\Lambda^{d-2}}{2t}\int \mathrm{d}^d x \, (\partial_\mu\theta)^2 + i\sum_i \epsilon_i\theta(x_i)\right] \tag{33.3}$$

with $\epsilon_i = \pm 1$. The method has been explained in Section 32.1. The result is

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\epsilon_i\theta(x_i)} \right\rangle = \exp\left[-\tfrac{1}{2}t\Lambda^{2-d}\sum_{i,j}\epsilon_i\epsilon_j\Delta(x_i - x_j)\right]. \tag{33.4}$$

in which $\Delta(x)$ is the massless propagator:

$$\Delta(x) = \frac{1}{(2\pi)^d}\int \frac{\mathrm{d}^d p}{p^2}\,\mathrm{e}^{ipx} = \Gamma(d/2-1)\frac{x^{2-d}}{4\pi^{d/2}} \qquad \text{for } d > 2\,. \tag{33.5}$$

$\Delta(0)$ diverges with the cut-off Λ:

$$\int \frac{\mathrm{d}^d p}{p^2} \propto \Lambda^{d-2}.$$

The limit $d = 2$ is singular, as we have discussed in Section 32.1, because the field is massless (IR divergence).

Remark. To define renormalized correlation functions, we have to introduce a renormalization scale μ and the corresponding coupling constant t_{r}:

$$t\Lambda^{2-d} = t_{\mathrm{r}}\mu^{2-d},$$

and cancel the divergent terms in the exponential in the r.h.s. of equation (33.4):

$$\sum_{i,j}\epsilon_i\epsilon_j\Delta(x_i - x_j) = n\Delta(0) + 2\sum_{i<j}\epsilon_i\epsilon_j\Delta(x_i - x_j). \tag{33.6}$$

As noted in Section 32.1, the fields $\mathrm{e}^{\pm i\theta(x)}$, which are composite fields in terms of $\theta(x)$, require a field renormalization Z:

$$Z = \mathrm{e}^{-t_{\mathrm{r}}\mu^{2-d}\Delta(0)}\,. \tag{33.7}$$

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 $d = 2+\varepsilon$, we find

$$\Delta(x) = \frac{1}{2\pi\varepsilon} - \frac{1}{4\pi}\left(\ln x^2 + \ln\pi + \gamma\right) + O(\varepsilon), \tag{33.8}$$

$$\Delta(0) = \frac{1}{2\pi\varepsilon} + \frac{1}{2\pi}\ln\Lambda + \text{ terms finite when } \varepsilon \to 0 \text{ or } \Lambda \to \infty\,, \tag{33.9}$$

33.1 The Spin Correlation Functions

We first calculate the spin correlation functions in d dimensions with the free action (33.2):

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\epsilon_i\theta(x_i)} \right\rangle = \int [\mathrm{d}\theta] \exp\left[-\frac{\Lambda^{d-2}}{2t} \int \mathrm{d}^d x\, (\partial_\mu \theta)^2 + i \sum_i \epsilon_i \theta(x_i) \right] \tag{33.3}$$

with $\epsilon_i = \pm 1$. The method has been explained in Section 32.1. The result is

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\epsilon_i\theta(x_i)} \right\rangle = \exp\left[-\tfrac{1}{2} t \Lambda^{2-d} \sum_{i,j} \epsilon_i \epsilon_j \Delta(x_i - x_j) \right]. \tag{33.4}$$

in which $\Delta(x)$ is the massless propagator:

$$\Delta(x) = \frac{1}{(2\pi)^d} \int \frac{\mathrm{d}^d p}{p^2}\, \mathrm{e}^{ipx} = \Gamma(d/2 - 1) \frac{x^{2-d}}{4\pi^{d/2}} \qquad \text{for } d > 2\,. \tag{33.5}$$

$\Delta(0)$ diverges with the cut-off Λ:

$$\int \frac{\mathrm{d}^d p}{p^2} \propto \Lambda^{d-2}.$$

The limit $d = 2$ is singular, as we have discussed in Section 32.1, because the field is massless (IR divergence).

Remark. To define renormalized correlation functions, we have to introduce a renormalization scale μ and the corresponding coupling constant t_{r}:

$$t \Lambda^{2-d} = t_{\mathrm{r}} \mu^{2-d},$$

and cancel the divergent terms in the exponential in the r.h.s. of equation (33.4):

$$\sum_{i,j} \epsilon_i \epsilon_j \Delta(x_i - x_j) = n\Delta(0) + 2 \sum_{i<j} \epsilon_i \epsilon_j \Delta(x_i - x_j). \tag{33.6}$$

As noted in Section 32.1, the fields $\mathrm{e}^{\pm i\theta(x)}$, which are composite fields in terms of $\theta(x)$, require a field renormalization Z:

$$Z = \mathrm{e}^{-t_{\mathrm{r}} \mu^{2-d} \Delta(0)}\,. \tag{33.7}$$

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 $d = 2 + \varepsilon$, we find

$$\Delta(x) = \frac{1}{2\pi\varepsilon} - \frac{1}{4\pi} \left(\ln x^2 + \ln \pi + \gamma \right) + O(\varepsilon), \tag{33.8}$$

$$\Delta(0) = \frac{1}{2\pi\varepsilon} + \frac{1}{2\pi} \ln \Lambda + \text{ terms finite when } \varepsilon \to 0 \text{ or } \Lambda \to \infty\,, \tag{33.9}$$

where $\gamma = -\psi(1)$ is Euler's constant. It follows that the sum of the divergent contributions in the exponential of equation (33.4) takes the form

$$-t\frac{\Lambda^{2-d}}{4\pi\varepsilon}\sum_{i,j}\epsilon_i\epsilon_j = -t\frac{\Lambda^{2-d}}{4\pi\varepsilon}\left(\sum_i \epsilon_i\right)^2 . \tag{33.10}$$

Therefore, all correlation functions vanish in the $d = 2$ limit, except 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. Note, however, that this property does not depend on the compact nature of the $O(2)$ group. All correlation functions invariant under a translation of $\theta(x)$ by a constant have a finite limit. In particular, as shown in Section 32.1, we can replace the signs ϵ_i by any set of numbers with a vanishing sum.

The non-vanishing limit, more explicitly, is then

$$\left\langle \prod_{i=1}^{n} \mathrm{e}^{i\epsilon_i\theta(x_i)}\right\rangle \propto \prod_{i<j}\left(\Lambda\,|x_i - x_j|\right)^{\epsilon_i\epsilon_j t/2\pi} . \tag{33.11}$$

This result is rather surprising: although the $O(2)$ symmetry is not broken the correlation functions decay algebraically at large distance and the correlation length is thus infinite. In addition, the power behaviour depends continuously on the temperature: in the RG sense, we have a line of fixed points and this is consistent with the property that the RG β-function vanishes identically. Specializing equation (33.11) to the two-point function,

$$\left\langle \mathrm{e}^{i\theta(x)}\,\mathrm{e}^{-i\theta(0)}\right\rangle \sim x^{-t/2\pi}, \tag{33.12}$$

we obtain the value of the exponent

$$\eta = t/2\pi . \tag{33.13}$$

It remains, however, to understand the absence of phase transition in this model. In particular, no invariant relevant operator can be constructed which would modify the action (33.2). Since we do not have yet all the necessary ingredients to examine this question we postpone it until Section 33.4. Let us finally note that $t > 4\pi$ implies $\eta > 2$. This is an unphysical 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 the *a priori* largest possible value of t for which the action (33.2) can represent the $O(2)$ model below T_c.

33.2 Correlation Functions in a Field

If we add a magnetic field term, the action (33.2) becomes

$$\mathcal{S}(\theta) = \frac{\Lambda^{d-2}}{t} \int \mathrm{d}^d x \left\{ \frac{1}{2} \left[\partial_\mu \theta(x)\right]^2 - h \cos\theta(x) \right\}. \tag{33.14}$$

We recognize the SG action studied in Section 32.3. Since we know the two RG functions

$$\beta(t) = (d-2)t\,, \qquad \zeta(t) = t/2\pi\,, \tag{33.15}$$

we can use the expressions of Chapter 31 to find the scaling form of correlation functions. For $d > 2$, the two functions $M_0(t)$ and $\xi(t)$ defined by equations (31.33,31.34) are

$$M_0(t) = \exp\left[-\frac{1}{2}\int_0^t \frac{\zeta(t')}{\beta(t')}\mathrm{d}t'\right] = \exp\left(-\frac{t}{4\pi\varepsilon}\right), \tag{33.16}$$

$$\xi(t) = \Lambda^{-1} t^{1/\varepsilon} \exp\left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{\varepsilon t'}\right)\mathrm{d}t'\right] = \Lambda^{-1} t^{1/\varepsilon}, \tag{33.17}$$

and, therefore, (equation (31.44))

$$W^{(n)}(p_i, t, h) = t^{(n-1)d/(d-2)}\, \mathrm{e}^{-nt/4\pi\varepsilon}\, F^{(n)}\left(p_i t^{1/\varepsilon}, h\,\mathrm{e}^{-t/4\pi\varepsilon}\, t^{-2/\varepsilon}\right). \tag{33.18}$$

Dimension 2. In dimension 2, the situation again is different since the β-function vanishes. In perturbation theory, the action (33.14) is super-renormalizable, which corresponds to the relevance of the magnetic field for long distance properties. The RG equations have been written in Chapter 31 and Section 32.3. For any temperature t, the correlation functions have a scaling behaviour. Let us rewrite RG equation (31.23) in this particular case:

$$\left(\Lambda\frac{\partial}{\partial\Lambda} + \frac{n}{2}\zeta(t) + \frac{1}{2}\zeta(t)h\frac{\partial}{\partial h}\right) W^{(n)} = 0\,. \tag{33.19}$$

Using the dimensional relation

$$\Lambda\frac{\partial}{\partial\Lambda} + 2h\frac{\partial}{\partial h} + p_i\frac{\partial}{\partial p_i} = 2(1-n)\,, \tag{33.20}$$

we obtain the scaling form of correlation functions:

$$W^{(n)}(p_i) = h^{t/(8\pi - t)} h^{1-n} F^{(n)}\left(p_i h^{-4\pi/(8\pi - t)}\right). \tag{33.21}$$

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 of the exponent η (equation (33.13)).

In particular, calling M the magnetization, we have

$$M = W^{(1)} \sim h^{t/(8\pi - t)}. \tag{33.22}$$

We have already noticed in Section 32.3, 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 above when examining the behaviour of the spin two-point function, for $t > 4\pi$ the model no longer represents the long distance physics of the lattice $O(2)$ ferromagnet.

33.3 The Coulomb Gas in Two Dimensions

We now postpone further discussion of the $O(2)$ model and study a model which is not obviously related, the two-dimensional Coulomb gas. We first show that the Coulomb gas is, in any dimension, related to the SG field theory.

33.3.1 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 (33.5). Therefore, the r.h.s. of equation (33.4) is, up to a multiplicative factor, the Boltzmann weight for a gas of particles with charges $\epsilon_i = \pm 1$, and temperature $2\Lambda^{d-2}/t$. Comparing then with the identities of Section 32.2.2, we guess that there is a relation between the SG field theory and the partition function of the Coulomb gas. To establish the correspondence, we generalize the identities of Section 28.2.1 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 an external background charge density $\rho_{\text{ext.}}$. In terms of the *charge density* $\rho(x)$ of the particles,

$$\rho(x) = \sum_i \epsilon_i \delta(x - x_i), \quad \epsilon_i = \pm 1 , \tag{33.23}$$

the potential energy of the gas can then be written as

$$\mathcal{V}(\rho) = \frac{1}{2} \int \mathrm{d}^d x \mathrm{d}^d y \, \big(\rho(x) + \rho_{\text{ext.}}(x)\big) V(x-y) \big(\rho(y) + \rho_{\text{ext.}}(y)\big), \tag{33.24}$$

provided the neutrality condition is satisfied:

$$\int \mathrm{d}^d x \big(\rho(x) + \rho_{\text{ext.}}(x)\big) = 0 . \tag{33.25}$$

The neutrality condition implies that the total background charge takes integer values.

In the partition function, we impose the definition (33.23) by the functional integral

$$\int [\mathrm{d}\phi] \exp \left[i \int \mathrm{d}^d x \, \phi(x) \rho(x) - i \sum_i \epsilon_i \phi(x_i)) \right] ,$$

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

$$\mathcal{Z} = \int [\mathrm{d}\phi][\mathrm{d}\rho] \exp\left[-\mathcal{S}(\rho,\phi)\right]$$

with

$$\mathcal{S}(\rho,\phi) = \mathcal{V}(\rho)/T - \int \mathrm{d}^d x \left[i\phi(x)\rho(x) + 2z \cos\phi(x) \right] , \tag{33.26}$$

because in the expansion of $\exp[2z \int \mathrm{d}^d \cos\phi(x)]$ in powers of z the terms which 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

$$\int \mathrm{d}^d y\, V(x-y)\big(\rho(y)+\rho_{\text{ext.}}(y)\big) - iT\phi(x) = 0\,,$$

which can be inverted to yield

$$\rho(x)+\rho_{\text{ext.}}(x) = -iT\nabla^2\phi(x). \tag{33.27}$$

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

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}T\big(\partial_\mu\phi(x)\big)^2 - 2z\cos\phi(x) + i\rho_{\text{ext.}}(x)\phi(x)\right]. \tag{33.28}$$

From this action (33.28) 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 $\mathrm{e}^{\pm iq\phi(x)}$. 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 32.3 is

$$T = 1/t\,, \quad z = \alpha/2t\,.$$

Remark. We have seen that for $t = 4\pi$, the free energy becomes infinite in the absence of a short distance cut-off. This implies that for $t \geq 4\pi$, that is, $T \leq 1/4\pi$ the Coulomb gas is only stable if the charged particles have a hard core.

The phase transition. We have shown in Section 32.3 that the quantum SG model must undergo a phase transition at $t = 8\pi$ for α small, that is, low fugacity and thus low particle density, between a phase with a finite correlation length at low t, that is, at high temperature T in the Coulomb gas, and a phase with infinite correlation length at high t, that is, at low T in the Coulomb gas. In the Coulomb gas language, the nature of these 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 pairs. In the Coulomb gas language, 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. In the phase of molecular bound states instead 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 is non-zero only at coinciding points.

33.3.2 Renormalization and renormalization group

We now examine more precisely the properties of this 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. Fortunately, we have derived in Chapter 32 a remarkable relation between the SG model and a two fermion model which is a free field theory just at the transition point. This allows the use of perturbative arguments in the fermion model to confirm the RG equations obtained more directly and to gain some further insight into the problem.

At $T = 1/8\pi$ (and $z \to 0$), the SG model is just renormalizable. 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:

$$u = 1 - 1/8\pi T\,, \qquad v = 2z/T\Lambda^2. \tag{33.29}$$

We assume that the set $\{u, v\}$ is multiplicatively renormalizable. To these coupling constants correspond two RG β-functions. The property that the free average of the product of an odd number of operators $\mathrm{e}^{\pm i\phi}$ vanishes (chiral symmetry in the equivalent fermion formulation) implies a parity symmetry in v. The RG function β_v is given at leading order in v in terms of the function $\zeta(t)$ of equations (33.15). Thus,

$$\beta_v = v\left[\tfrac{1}{2}\zeta\left(T^{-1}\right) - 2\right] + O(v^3) = -2uv + O(v^3)\,. \tag{33.30}$$

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

$$\beta_u = -2v^2, \qquad \beta_v = -2uv\,. \tag{33.31}$$

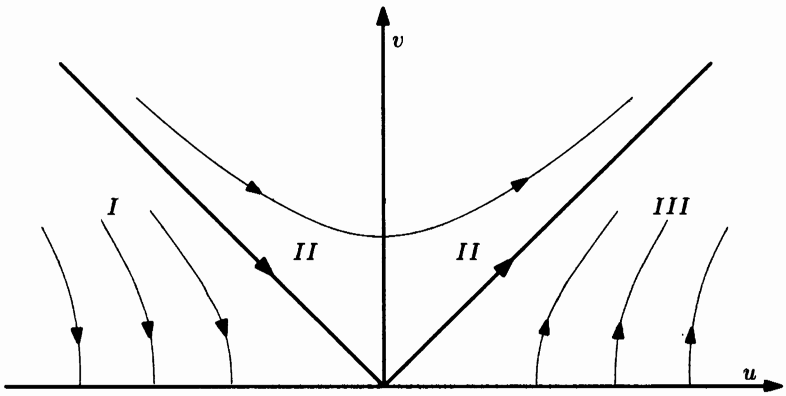


Fig. 33.1 The RG flow of the 2D Coulomb gas.

The RG flow. At this order, it is easy to find curves that are RG invariant:

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(u^2(\lambda) - v^2(\lambda)\right) = 2u\beta_u - 2v\beta_v = 0\,. \tag{33.32}$$

For $|u| > |v|$, we parametrize these hyperbolas by

$$u = a\frac{1+s^2}{1-s^2} \quad |s| < 1\,, \tag{33.33}$$

$$v = a\frac{2s}{1-s^2}. \tag{33.34}$$

The RG equation for the parameter s is then

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}s(\lambda) = -2as(\lambda)\,.$$

The solution is

$$s(\lambda) = s_0\lambda^{-2a}. \tag{33.35}$$

In the long distance limit λ goes to zero. 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 one leaves the perturbative regime. This is consistent with the relevance of v for large T. If instead we start with $T < 1/8\pi$, which implies $u < 0$ and $a > 0$, then $s(\lambda)$ goes to zero for small λ. $v(\lambda)$ then goes to zero 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.

For $|u| < |v|$, we parametrize the hyperbolas instead by

$$u = a\frac{2s}{1-s^2}, \quad |s| < 1, \tag{33.36}$$

$$v = a\frac{1+s^2}{1-s^2}. \tag{33.37}$$

We then find

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}s(\lambda) = -a(1+s^2)\,.$$

The solution can be written as

$$\arctan s(\lambda) = -a\ln\lambda + \arctan s_0\,. \tag{33.38}$$

Therefore, irrespective of the sign of a, that is, the initial sign of u, $u(\lambda)$ and $v(\lambda)$ always increase until one leaves the perturbative regime. This again is a region in which v is relevant. The conclusions are that the half-plane $v > 0$ is divided in three regions (see figure 33.1) 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 $u + v = 0$ is thus 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.

33.3.3 The correlation length near the phase transition

Let us now characterize the behaviour of the correlation length in the high temperature phase when one approaches the phase transition. We set

$$v + u = \tau\,, \tag{33.39}$$

$$v - u = 2w\,. \tag{33.40}$$

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. It, therefore, satisfies

$$\xi\big(\tau(\lambda), w(\lambda)\big) = \lambda\xi\big(\tau(1), w(1)\big). \tag{33.41}$$

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 a RG invariant:

$$\tau(\lambda)w(\lambda) = \tau w .$$

The RG equation for $\tau(\lambda)$ then reads

$$\lambda\frac{\mathrm{d}\tau}{\mathrm{d}\lambda} = -2\tau w - \tau^2(\lambda). \tag{33.42}$$

A short calculation gives the function $\tau(\lambda)$:

$$\arctan\left(\tau(\lambda)/\sqrt{2w\tau}\right) - \arctan\left(\tau/\sqrt{2w\tau}\right) = -\sqrt{2w\tau}\ln\lambda . \tag{33.43}$$

We evaluate the l.h.s. 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

$$\xi \sim \exp\left(\frac{\pi}{2\sqrt{2w\tau}}\right), \tag{33.44}$$

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

$$\beta_u = -2v^2 - 4v^2u\,, \tag{33.45}$$
$$\beta_v = -2uv + 5v^3. \tag{33.46}$$

None of the conclusions drawn from the analysis at leading order is qualitatively affected.

33.4 $O(2)$ Spin Model and Coulomb Gas

We have seen that the action (33.2) cannot represent the $O(2)$ spin model for all temperatures and we feel that somehow we have to introduce the condition that the $\theta(x)$ is a cyclic variable. We also know from the analysis of Section 31.1 that at higher temperature, the non-linear σ-model approximation fails because there may be 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, in a turn around these points, the direction of the field changes by an angle multiple of 2π. We, therefore, consider a configuration of the field $\theta(x)$ which 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:

$$\theta(x) = \theta_1(x) + \theta_V(x, x_i) \tag{33.47}$$

with

$$\theta_V(x, x_i) = \sum_i n_i \arctan \frac{(x - x_i)_2}{(x - x_i)_1}, \quad n_i \in \mathbb{Z}. \tag{33.48}$$

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 $\Lambda = 1$)

$$\mathcal{S}(\theta) = \frac{1}{2t} \int \mathrm{d}^2 x \, [\partial_\mu \theta_1(x) + \partial_\mu \theta_V(x)]^2 . \tag{33.49}$$

We now use the identities

$$\partial_\mu \theta_V(x) = \sum_i n_i \epsilon_{\mu\nu} \frac{(x - x_i)_\nu}{(x - x_i)^2} \tag{33.50}$$

$$= \sum_i n_i \epsilon_{\mu\nu} \partial_\nu \ln |x - x_i| . \tag{33.51}$$

We observe from expression (33.49) that the energy can be finite only if

$$\sum_i n_i = 0 .$$

An integration by parts shows that the cross-term in the r.h.s. of equation (33.49) vanishes and yields

$$\int \mathrm{d}^2 x \, [\partial_\mu \theta_V(x)]^2 = \int \mathrm{d}^2 x \sum_{ij} n_i n_j \partial_\mu \ln |x - x_i| \, \partial_\mu \ln |x - x_j|$$

$$= -2\pi \sum_{ij} n_i n_j \ln |x_i - x_j| . \tag{33.52}$$

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 32.3.2 (equation (32.71)) that the most relevant terms correspond to $n_i = \pm 1$ which 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

$$t' t = 4\pi^2 . \tag{33.53}$$

In particular, by introducing an auxiliary field $\theta_2(x)$ and using the identity (33.4), we can write an effective action:

$$\mathcal{S}(\theta_1, \theta_2) = \int \mathrm{d}^2 x \left[\frac{1}{2t} (\partial_\mu \theta_1)^2 + \frac{t}{8\pi^2} (\partial_\mu \theta_2)^2 - \frac{At}{4\pi^2} \cos \theta_2 \right] . \tag{33.54}$$

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.

33.5 The Critical Two-Point Function in the $O(2)$ Model

We now want to compute the two-point correlation function of the $O(2)$ model near $t = \pi/2$. We, therefore, have to calculate the expectation value:

$$\left\langle e^{i(\theta_1(y)+\theta_V(y))} e^{-i(\theta_1(x)+\theta_V(x))} \right\rangle = |x-y|^{-t/2\pi} \left\langle e^{i(\theta_V(y)-\theta_V(x))} \right\rangle_{\mathrm{CG}} . \tag{33.55}$$

To find an interpretation to this expectation value in terms of a modification of the energy of the Coulomb gas, we make a few transformations:

$$\begin{aligned} \theta_V(y) - \theta_V(x) &= \int_x^y \mathrm{d}s_\mu \partial_\mu \theta_V(s) \\ &= \frac{2\pi}{t} \sum_i n_i \frac{t}{2\pi} \int_x^y \mathrm{d}s_\mu \epsilon_{\mu\nu} \partial_\nu \ln|s - x_i| . \end{aligned} \tag{33.56}$$

To calculate the Coulomb gas expectation value in the r.h.s. of equation (33.55), we want to write the sum of the term (33.56) and the Coulomb gas energy (33.55) as arising from the free field average in a source. The source $J_2(s)$ for the free field $\theta_2(s)$:

$$J_2(s) = J(s) + K(s), \qquad \begin{cases} J(s) = i \sum_i n_i \delta(s - x_i) , \\ K(s) = \dfrac{t}{2\pi} \displaystyle\int_x^y \mathrm{d}s'_\mu \epsilon_{\mu\nu} \partial_\nu \delta(s - s') , \end{cases} \tag{33.57}$$

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 (33.56), but in addition a KK contribution $D(x, y)$ arises of the form

$$\begin{aligned} D(x,y) &\equiv \frac{2\pi^2}{t} \int \mathrm{d}s\, \mathrm{d}s' K(s) K(s') \Delta(s - s') \\ &= -\frac{t}{4\pi} \int_x^y \mathrm{d}s_\mu\, \mathrm{d}s'_\nu \epsilon_{\mu\rho} \epsilon_{\nu\sigma} \partial^s_\rho \partial^{s'}_\sigma \ln|s - s'| . \end{aligned} \tag{33.58}$$

We can thus write the Coulomb expectation value (33.55) as a free field expectation value in the source (33.57) 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

$$D(x,y) = -\frac{t}{2} \int_x^y \mathrm{d}s \cdot \mathrm{d}s' \delta(s - s') + \frac{t}{4\pi} \int_x^y \mathrm{d}s_\mu\, \mathrm{d}s'_\nu \partial^s_\mu \partial^{s'}_\nu \ln|s - s'| . \tag{33.59}$$

The integration over s and s' can be performed. Only one term coming from the second integral, gives a contribution which depends on $x - y$, the others are cut-off dependent. The final result is

$$D(x,y) = -\frac{t}{2\pi} \ln|x - y| + \text{ const.} \tag{33.60}$$

If we replace the Coulomb gas average in equation (33.55) by the free field average with the source (33.57), 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 field appearing in expression (33.57) (see also equation (32.89*b*)):

$$\exp[i\theta(x)] \longmapsto \exp\left[-\frac{t}{2\pi}\int^x \mathrm{d}s_\mu \epsilon_{\mu\nu}\partial_\nu\theta_2(s)\right]. \tag{33.61}$$

The critical two-point function. To calculate the two-point function at the critical temperature, we can now use the RG considerations of Section 33.3.1. Criticality corresponds to the special line $u = -v$ in the notation of Section 33.3.1. Note that the relation between t and u on the critical line is

$$t = \pi/2(1-u). \tag{33.62}$$

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 (33.54), the two-point function is still given by equation (33.12).

RG equations. The critical spin two-point correlation function $W^{(2)}$ satisfies the RG equation:

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(t)\frac{\partial}{\partial t} + \zeta(t)\right] W^{(2)}(x,t,\Lambda) = 0. \tag{33.63}$$

The functions $\zeta(t)$ and $\beta(t)$ are (equations (33.15,33.31,33.62))

$$\begin{aligned} \zeta(t) &= t/2\pi + O\left(t-\frac{\pi}{2}\right)^2, \\ \beta(t) &= -\frac{4}{\pi}\left(t-\frac{\pi}{2}\right)^2 + O\left(t-\frac{\pi}{2}\right)^3. \end{aligned} \tag{33.64}$$

For $t < \pi/2$, the theory is IR free. The effective coupling constant at scale λ behaves like

$$t(\lambda) = \frac{\pi}{2}\left(1+\frac{1}{2\ln\lambda}\right) + o\left(\frac{1}{\ln\lambda}\right) \text{ for } \lambda \to 0. \tag{33.65}$$

$W^{(2)}(x,t)$ being dimensionless, the solution of the RG equation can be written as

$$W^{(2)}(x/\lambda,t) = Z^2(\lambda)W^{(2)}(x,t(\lambda)) \tag{33.66}$$

with

$$Z(\lambda) = \exp\left[\frac{1}{2}\int_1^\lambda \frac{\mathrm{d}\sigma}{\sigma}\zeta(t(\sigma))\right] \sim \lambda^{1/8}\left|\ln\lambda\right|^{1/16}. \tag{33.67}$$

Therefore, at large distance, the critical two-point correlation function $W^{(2)}$ behaves like

$$W^{(2)}(x,t) \underset{x\to\infty}{\sim} x^{-1/4}(\ln x)^{1/8}. \tag{33.68}$$

This is the celebrated KT result for the phase transition of the classical 2D XY model.

33.6 The Generalized Thirring Model

We now use the results of Section 32.6 to justify the RG functions (33.31) and thus the RG flow near the phase transition. The correspondence with the SG model (see equation (32.113)), in the notation of the effective $O(2)$ spin action (33.54) is

$$A \sim f\,, \qquad t = \tfrac{1}{2}\pi(1 + g/\pi). \tag{33.69}$$

From these expressions, we see 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 (A32.18))

$$\begin{aligned} \beta_g &= -2f^2/\pi\,, \\ \beta_f &= -2fg/\pi\,. \end{aligned} \tag{33.70}$$

We recognize the RG functions (33.31) 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 that decouples. Note that, as discussed in Section A32.2, this does not mean that chiral symmetry in the fermion model is spontaneously broken; the expectation 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. For $g < 0$ instead, the force between fermion 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 Appendix A32.2).

For the two-parameter model we have the three regions of figure 33.1, a region I in which particles are massless, and regions II and III where particles are massive.

Note finally that in the language of the simple massive Thirring model of Section 32.5 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.

Bibliographical Notes

The low temperature phase of the model was considered by

V.L. Berezinskii, *Zh. Eksp. Teor. Fiz.* 59 (1970) 907 (*Sov. Phys. JETP* 32 (1970) 493).

Relevant is also the article

J. Villain, *J. Phys. C: Solid State Phys.* 6 (1973) 581.

The properties of the Kosterlitz–Thouless transition have been studied in

J.M. Kosterlitz and D.J. Thouless, *J. Phys. C: Solid State Phys.* 6 (1973) L97; J.M. Kosterlitz, *J. Phys. C: Solid State Phys.* 7 (1974) 1046; J.V. José, L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, *Phys. Rev.* B16 (1977) 1217.

For the relation between the Coulomb gas and the $O(2)$ XY model, see

B. Nienhuis in *Phase Transitions and Critical Phenomena* vol. 11, C. Domb and J.L. Lebowitz eds. (Academic Press, New York 1987).

Our discussion of the Kosterlitz–Thouless transition follows the lines of

P.B. Wiegmann, *J. Phys. C: Solid State Phys.* 11 (1978) 1583; D.J. Amit, Y.Y. Goldschmidt and G. Grinstein, *J. Phys. A: Math. Gen.* 13 (1980) 585.

34 CRITICAL PROPERTIES OF GAUGE THEORIES

In this chapter, we consider lattice models with gauge symmetry and discuss their properties from the point of view of phase transitions and spectrum structure. We concentrate on pure lattice gauge theories (without fermions) and study them mainly with lattice methods. Physically, this means that we cannot investigate many properties of a realistic theory like QCD where fermions are coupled through a gauged $SU(3)$ colour group, but we can still investigate one important question:

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?

Other problems which we do not consider here, can also be discussed in this framework: for example, the appearance of massive group singlet bound states in the spectrum (gluonium), the question of a deconfinement transition at finite physical temperature in QCD.

We first construct lattice models with gauge symmetry. We show that, as anticipated, they provide a lattice regularization of the continuum gauge theories studied in Chapters 18–21: the low temperature or small coupling expansion of the lattice model is a regularized continuum perturbation theory. We then discuss pure gauge theories (without matter fields) on the lattice. We discover that gauge theories have properties quite different from the ferromagnetic systems we have studied so far. In particular, the absence of a local order parameter will force us to examine the behaviour of a non-local quantity, a functional of loops called hereafter Wilson's loop to distinguish between the confined and deconfined phases. Results will be obtained in the high temperature or strong coupling limit and in the mean field approximation.

34.1 Gauge Invariance on the Lattice

The construction of lattice gauge theories is based on an idea that has already been introduced in Chapter 19: *parallel transport.*

We start from a model possessing a global (rigid) symmetry group G, and we want to make it gauge invariant.

To each site i of a lattice, we associate a set of dynamic variables φ_i, representing matter fields, on which acts an orthogonal representation $\mathcal{D}(G)$ of the group G:

$$\varphi_{\mathbf{g}} = \mathbf{g}\varphi\,, \qquad \mathbf{g} \in \mathcal{D}(G)\,. \tag{34.1}$$

A model is gauge invariant (local invariance) if it is invariant under independent group transformation on each lattice site i. For the φ-measure of integration as well as for all the terms in the lattice action which depend only on one site, global invariance implies local invariance as in the continuum. Problems arise only with terms which connect different lattice sites.

34 CRITICAL PROPERTIES OF GAUGE THEORIES

In this chapter, we consider lattice models with gauge symmetry and discuss their properties from the point of view of phase transitions and spectrum structure. We concentrate on pure lattice gauge theories (without fermions) and study them mainly with lattice methods. Physically, this means that we cannot investigate many properties of a realistic theory like QCD where fermions are coupled through a gauged $SU(3)$ colour group, but we can still investigate one important question:

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?

Other problems which we do not consider here, can also be discussed in this framework: for example, the appearance of massive group singlet bound states in the spectrum (gluonium), the question of a deconfinement transition at finite physical temperature in QCD.

We first construct lattice models with gauge symmetry. We show that, as anticipated, they provide a lattice regularization of the continuum gauge theories studied in Chapters 18–21: the low temperature or small coupling expansion of the lattice model is a regularized continuum perturbation theory. We then discuss pure gauge theories (without matter fields) on the lattice. We discover that gauge theories have properties quite different from the ferromagnetic systems we have studied so far. In particular, the absence of a local order parameter will force us to examine the behaviour of a non-local quantity, a functional of loops called hereafter Wilson's loop to distinguish between the confined and deconfined phases. Results will be obtained in the high temperature or strong coupling limit and in the mean field approximation.

34.1 Gauge Invariance on the Lattice

The construction of lattice gauge theories is based on an idea that has already been introduced in Chapter 19: *parallel transport.*

We start from a model possessing a global (rigid) symmetry group G, and we want to make it gauge invariant.

To each site i of a lattice, we associate a set of dynamic variables φ_i, representing matter fields, on which acts an orthogonal representation $\mathcal{D}(G)$ of the group G:

$$\varphi_{\mathbf{g}} = \mathbf{g}\varphi\,, \qquad \mathbf{g} \in \mathcal{D}(G)\,. \tag{34.1}$$

A model is gauge invariant (local invariance) if it is invariant under independent group transformation on each lattice site i. For the φ-measure of integration as well as for all the terms in the lattice action which depend only on one site, global invariance implies local invariance as in the continuum. Problems arise only with terms which connect different lattice sites.

Let us consider, for example, a term in the action of the form $\varphi_i \cdot \varphi_j$, i and j being different sites on the lattice. Such a term is invariant under global but not local transformations:

$$\varphi_i \cdot \varphi_j \mapsto \varphi_i{}^T\mathbf{g}_i\,\mathbf{g}_j\varphi_j\,, \tag{34.2}$$

(where ${}^T\mathbf{g}$ means $\mathbf{g}$ transposed). To render it invariant it is necessary to introduce a new dynamic variable, a matrix $\mathbf{U}_{ij}$ belonging to the representation $\mathcal{D}(G)$, which depends on the two sites i, j and transforms like

$$\mathbf{U}_{ij} \mapsto \mathbf{g}_i\mathbf{U}_{ij}\,{}^T\mathbf{g}_j\,. \tag{34.3}$$

Then, the quantity

$$\varphi_i\mathbf{U}_{ij}\varphi_j \tag{34.4}$$

is gauge invariant. Moreover, if $\mathbf{U}_{ij}$ and $\mathbf{U}_{jk}$ are two matrices transforming with the rule (34.4) then the product of matrices $\mathbf{U}_{ij}\mathbf{U}_{jk}$ transforms like

$$\mathbf{U}_{ij}\mathbf{U}_{jk} \mapsto \mathbf{g}_i\mathbf{U}_{ij}\mathbf{U}_{jk}\,{}^T\mathbf{g}_k\,. \tag{34.5}$$

In the transformation (34.3), we recognize the transformation of a parallel transporter. In the continuum, a parallel transporter depends not only on the end-points i, j but also on the curve joining them. Moreover, in a local field theory one needs only transport along infinitesimal curves which can be expressed in terms of a gauge field or connection, element of the representation of the Lie algebra.

On the lattice curves follow links, the segments which connect adjacent sites. The minimum displacement is a link and two arbitrary lattice sites can be joined by a path formed of links of the lattice. As a consequence of the composition rule (34.5), one can thus take as dynamic variables elements $\mathbf{U}_\ell$ of the group representation associated with parallel transport along oriented links of the lattice, which transform like

$$\mathbf{U}_\ell \equiv \mathbf{U}_{ab} \mapsto \mathbf{g}_a\mathbf{U}_\ell\,{}^T\mathbf{g}_b\,,$$

where the link ℓ goes from site b to a. It is consistent with the transformation law to choose

$$\mathbf{U}_{ba} = {}^T\mathbf{U}_{ab} = \mathbf{U}_{ab}^{-1}. \tag{34.6}$$

Then, we can choose for matrix U_{ij} any parallel transporter product of link variables along a path C joining j to i:

$$\mathbf{U}[C(i,j)] = \prod_{\text{links }\ell\in C(i,j)} \mathbf{U}_\ell\,,$$

where the product is ordered along the path.

Relation with the continuum formulation: the abelian example. In continuum field theory, in the abelian $U(1)$ example, we have already explicitly constructed the parallel transporter (equation (18.30)) which is an element of the $U(1)$ group. In terms of the gauge field A_μ, it reads

$$U(x,y) = \exp\left[-ie\int_x^y A_\mu(s)\mathrm{d}s_\mu\right], \tag{34.7}$$

in which e is the gauge coupling constant, and the gauge field is integrated over some differentiable curve going from x to y. Indeed in a gauge transformation a charged field φ and the gauge field

$$\begin{aligned}\varphi(x) &\mapsto \mathrm{e}^{i\Lambda(x)}\,\varphi(x),\\ A_\mu(x) &\mapsto A_\mu(x) - \frac{1}{e}\partial_\mu\Lambda(x),\end{aligned} \tag{34.8}$$

the transformation of $U(x,y)$ is

$$e\int_x^y A_\mu(s)\mathrm{d}s_\mu \mapsto e\int_x^y A_\mu(s)\mathrm{d}s_\mu - \Lambda(y) + \Lambda(x) \;\Rightarrow\; U(x,y) \mapsto \mathrm{e}^{i\Lambda(y)-i\Lambda(x)}\,U(x,y). \tag{34.9}$$

The non-abelian case. In the non-abelian case, the explicit relation is more complicated because the gauge field $\mathbf{A}^\alpha_\mu(x)t_\alpha$ is an element of the Lie algebra of G and the matrices representing the field at different points do not commute. It can be formally written as

$$\mathbf{U}(x,y) = \mathrm{P}\left\{\exp\left[\int_x^y \mathbf{A}^\alpha_\mu(s)t^\alpha \mathrm{d}s_\mu\right]\right\}, \tag{34.10}$$

in which the symbol P means path-ordered integral.

34.2 The Pure Gauge Theory

We now discuss the pure gauge theory and its formal continuum limit as obtained from a low temperature, strong coupling expansion.

34.2.1 Action and partition function

We now have to construct a gauge invariant interaction for the gauge elements $\mathbf{U}$. It follows from the transformation (34.3) 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

$$\mathcal{S}(\mathbf{U}) = -\sum_{\text{plaquettes}} \operatorname{tr}\mathbf{U}_{ij}\mathbf{U}_{jk}\mathbf{U}_{kl}\mathbf{U}_{li}\,, \tag{34.11}$$

in which β_p is the plaquette coupling. The appearance of products of parallel transporters along closed loops is not surprising 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 along 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 a partition function corresponding to the action (34.11):

$$\mathcal{Z} = \int \prod_{\text{links}\{ij\}} \mathrm{d}\mathbf{U}_{ij}\;\mathrm{e}^{-\beta_p \mathcal{S}(\mathbf{U})}\,. \tag{34.12}$$

We integrate over $\mathbf{U}_{ij}$ with the group invariant (de Haar) measure associated with the group G. In contrast to continuum gauge theories, the expression (34.12) is well-defined on the lattice (at least as long as the volume is finite) because the group is compact and thus the volume of the group is finite. Therefore, gauge fixing is not required and a *completely gauge invariant formulation* of the theory is possible.

34.2.2 Low temperature analysis

We first want to understand the precise connection between the lattice theory (34.12) and the continuum field theory. For this purpose, we investigate the lattice theory at low temperature, that is, at 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. We start from a first plaquette 1234. Without loss of generality, we can set

$$\mathbf{U}_{12} = \mathbf{g}_1^{-1}\mathbf{g}_2 \,, \qquad \mathbf{g}_1, \mathbf{g}_2 \in \mathcal{D}(G). \tag{34.13}$$

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

$$\mathbf{U}_{23} = \mathbf{g}_2^{-1}\mathbf{g}_3 \,, \qquad \mathbf{U}_{34} = \mathbf{g}_3^{-1}\mathbf{g}_4 \,. \tag{34.14}$$

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 an orthogonal matrix $\mathbf{U}$ is maximum when all its eigenvalues are 1.) In particular,

$$\mathbf{U}_{12}\mathbf{U}_{23}\mathbf{U}_{34}\mathbf{U}_{41} = \mathbf{1} \,, \tag{34.15}$$

which yields

$$\mathbf{U}_{41} = \mathbf{g}_4^{-1}\mathbf{g}_1 \,. \tag{34.16}$$

If we now take an adjacent plaquette the argument can be repeated for all links but one, which has already been fixed. In this way, we can show that the minimum of the action is a pure gauge. Thus, when the coupling constant β_p becomes very large, 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 this 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. 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:

$$\mathbf{U}(x, x + an_\mu) = 1 - a\mathbf{A}_\mu(x) + O\left(a^2\right), \tag{34.17}$$

in which a is the lattice spacing, x the point on the lattice, and n_μ the unit vector in the direction μ. We know from the discussion of Section 19.1 that the matrix $\mathbf{A}_\mu(x)$ is the connection or gauge field. We have already shown that the transformation (34.3) 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 orthogonal matrix $\mathbf{U}$ associated with the link $(x, x + an_\mu)$ in terms of the antisymmetric matrix $\mathbf{A}_\mu(x)$ as

$$\ln \mathbf{U}(x + an_\mu, x) = -a\mathbf{A}_\mu(x + \tfrac{1}{2}an_\mu) + O\left(a^3\right). \tag{34.18}$$

We verify below that we need $\mathbf{U}$ up to order a^2. With the parametrization (34.18), equation (34.6) implies that the term of order a^2 vanishes. We now define the antisymmetric matrix $\mathbf{F}_{\mu\nu}(x)$ by

$$\begin{aligned} e^{-a^2\mathbf{F}_{\mu\nu}(x)} &= \mathbf{U}(x, x+an_\nu)\mathbf{U}(x+an_\nu, x+a(n_\mu+n_\nu)) \\ &\quad \times \mathbf{U}(x+a(n_\mu+n_\nu), x+an_\mu)\mathbf{U}(x+an_\mu, x). \end{aligned} \tag{34.19}$$

To calculate $\mathbf{F}_{\mu\nu}(x)$, we introduce the expansion (34.18) and use repeatedly the Baker–Hausdorf formula:

$$\ln\left(e^A\, e^B\right) = A + B + \frac{1}{2}[A, B] + \cdots . \tag{34.20}$$

Applied to the product of several factors, it takes the form

$$\ln\left(e^{A_1}\, e^{A_2} \ldots e^{A_n}\right) = \sum_i A_i + \frac{1}{2}\sum_{i<j}[A_i, A_j] + \cdots , \tag{34.21}$$

and, therefore,

$$\begin{aligned} a^2\mathbf{F}_{\mu\nu}(x) &= a\left[\mathbf{A}_\mu(x+\tfrac{1}{2}an_\mu) + \mathbf{A}_\nu(x+an_\mu+\tfrac{1}{2}an_\nu) - \mathbf{A}_\mu(x+an_\nu+\tfrac{1}{2}an_\mu)\right. \\ &\quad \left. -\mathbf{A}_\nu(x+\tfrac{1}{2}an_\nu)\right] + a^2\left[\mathbf{A}_\mu(x), \mathbf{A}_\nu(x)\right] + O\left(a^3\right). \end{aligned} \tag{34.22}$$

At leading order, we recover the curvature tensor

$$\mathbf{F}_{\mu\nu}(x) = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu] + O(a). \tag{34.23}$$

We obtain one term in the plaquette action by taking the trace of expression (34.19). Since $\mathbf{F}_{\mu\nu}$ is an antisymmetric matrix $\operatorname{tr}\mathbf{F}_{\mu\nu}$ vanishes. Thus,

$$\operatorname{tr} e^{-a^2\mathbf{F}_{\mu\nu}(x)} = \operatorname{tr}\mathbf{1} + a^4\operatorname{tr}\mathbf{F}^2_{\mu\nu}(x) + O\left(a^6\right). \tag{34.24}$$

This result shows that the leading term of the small field expansion of the plaquette action (34.11) is the standard gauge action studied in Chapter 19. The relation between β_p and the bare coupling constant e_0 of continuum gauge theories is thus

$$a^4\beta_p \sim e_0^{-2}. \tag{34.25}$$

As anticipated in Chapters 18,19, we conclude that the low temperature expansion, in a fixed gauge, of lattice gauge theories provides indeed a lattice regularization of continuum gauge theories. We have here discussed only the pure gauge action, but the generalization to 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.

34.3 Wilson's Loop and Confinement

In Section 20.2.2, we have calculated the RG β-functions for non-abelian gauge theories and shown that pure gauge theories are asymptotically free in four dimensions, which means that the origin in the coupling constant space is an UV fixed point and also implies that the effective interaction increases at large distance. Therefore, as in the case of the 2D 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 that the spectrum of the symmetric phase consists only in neutral states, that is, states which are singlets for the group transformations.

Clearly, it would be convenient to identify a local order parameter, that is, a local observable whose expectation value would distinguish between the QED phase of abelian gauge theories, in which charge states can be produced, from the so-called *confined* phase. However, in gauge theories such a local order parameter does not exist (see Elitzur's theorem). 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 18,19) 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, it has been suggested by Wilson to study, in pure gauge theories, a gauge invariant non-local quantity, the energy of the vacuum in presence of largely separated static charges. We thus first study this quantity in pure abelian gauge theories, in which, in the continuum, all calculations can be done explicitly.

34.3.1 Wilson's loop in continuum abelian gauge theories

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 18.4 (equation (18.48)):

$$\psi(A) = \exp\left[-ie \oint_{C_0} A_i(s)\mathrm{d}s_i\right], \tag{34.26}$$

in which the charges are located at both ends of the curve C_0.

By evaluating the behaviour for large time T of the matrix element

$$W(C_0) = \left\langle \psi \left| \mathrm{e}^{-HT} \right| \psi \right\rangle, \tag{34.27}$$

in which H is the gauge hamiltonian in the temporal gauge, we obtain the energy $E(C_0)$ of the vacuum in presence of static charges:

$$W(C_0) \underset{T\to\infty}{\sim} \mathrm{e}^{-TE(C_0)}. \tag{34.28}$$

If the charges are separated by a distance R, we expect E to depend only on R and not on C_0.

The loop functional $W(C_0)$ can be calculated from a functional integral:

$$W(C_0) = \left\langle \exp\left[-ie \oint_{C_0'} A_\mu(s)\mathrm{d}s_\mu\right]\right\rangle,$$

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 $-C_0$ at time T, respectively. The expectation value here means average over gauge field configurations.

Since in the temporal gauge the time component of A_μ vanishes, we can add to C_0' two straight lines in the time direction 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 figure 34.1):

$$W(C_0) \equiv W(C) = \left\langle \exp\left[-ie \oint_C A_\mu(s)\mathrm{d}s_\mu\right]\right\rangle . \tag{34.29}$$

The advantage of the representation (34.29) is that it is explicitly gauge invariant since it is the expectation value of the parallel transporter corresponding to a closed loop in space–time.

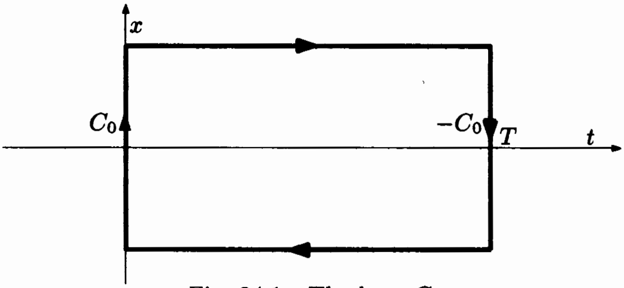


Fig. 34.1 The loop C.

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 (34.29) can be evaluated explicitly. To simplify calculations we take for C_0 also a straight line and use Feynman's gauge. The quantity $W(C)$ then is given by

$$W(C) = \int [\mathrm{d}A_\mu] \exp\left[-\mathcal{S}(A_\mu) + \int \mathrm{d}^d x\, J_\mu(x) A_\mu(x)\right] \tag{34.30}$$

with

$$\mathcal{S}(A_\mu) = \frac{1}{2}\int \mathrm{d}^d x\, [\partial_\mu A_\nu(x)]^2 \tag{34.31}$$

and

$$J_\mu(x) = -ie \oint_C \delta(x-s)\mathrm{d}s_\mu . \tag{34.32}$$

The result is

$$\ln W(C) = -\frac{\Gamma(d/2-1)}{8\pi^{d/2}} \mathrm{e}^2 \oint_{C\times C} \mathrm{d}\mathbf{s}_1 \cdot \mathrm{d}\mathbf{s}_2 \left|\mathbf{s}_1 - \mathbf{s}_2\right|^{2-d} . \tag{34.33}$$

The integral in the r.h.s. exhibits a short distance singularity, and a short distance cut-off thus is required. Moreover, to normalize the r.h.s. of equation (34.33), we divide it by

$W(C)$ taken for $R = a$, a being a fixed distance. We now write more explicitly the integrals:

$$\oint_{C\times C} \frac{\mathrm{d}\mathbf{s}_1 \cdot \mathrm{d}\mathbf{s}_2}{2\left|\mathbf{s}_1 - \mathbf{s}_2\right|^{d-2}} = \int_0^T |u-t|^{2-d}\,\mathrm{d}u\,\mathrm{d}t + \int_0^R |x-y|^{2-y}\,\mathrm{d}x\,\mathrm{d}y$$
$$-\int_0^R \left[(x-y)^2+T^2\right]^{1-d/2}\mathrm{d}x\,\mathrm{d}y - \int_0^T \left[(t-u)^2+R^2\right]^{1-d/2}\mathrm{d}t\,\mathrm{d}u\,. \tag{34.34}$$

The first term in the r.h.s. 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. The third term decreases with T for $d > 2$ which we now assume. Only the last term increases with T:

$$\int_0^T \left\{\left[(t-u)^2+R^2\right]^{1-d/2} - \left[(t-u)^2+a^2\right]^{1-d/2}\right\}\mathrm{d}t\,\mathrm{d}u$$
$$\sim \sqrt{\pi}\,\frac{\Gamma((d-3)/2)}{\Gamma(d/2-1)}\left(R^{3-d}-a^{3-d}\right)T\,. \tag{34.35}$$

Therefore, the vacuum energy $E(R)$ in presence of the static charges has the form

$$E(R)-E(a) = \frac{e^2}{4\pi^{(d-1)/2}}\Gamma((d-3)/2)\left(a^{3-d}-R^{3-d}\right). \tag{34.36}$$

We recognize the Coulomb potential between two charges.

For $d \le 3$, 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 above 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 $W(C)$. We here obtain

$$\ln W\left[C(R)\right] - \ln W\left[C(a)\right] = \frac{1}{2\pi^{d/2}}\Gamma(d/2-1)e^2\left\{\int_0^R \left[(u-t)^2+R^2\right]^{1-d/2}\mathrm{d}u\,\mathrm{d}t\right.$$
$$-\int_0^a \left[(u-t)^2+a^2\right]^{1-d/2}\mathrm{d}u\,\mathrm{d}t$$
$$\left.-\int_a^R |u-t|^{2-d}\mathrm{d}u\,\mathrm{d}t\right\}. \tag{34.37}$$

For $d > 3$, dimensions in which the Coulomb potential decreases, the r.h.s. is dominated by terms which correspond to the region $|s_1 - s_2| \ll R$ in equation (34.33):

$$\ln W\left[C(R)\right] - \ln W\left[C(a)\right] \sim \text{const.} \times R\,. \tag{34.38}$$

This is called the perimeter law since $\ln W(C)$ is proportional to the perimeter of C and is, therefore, relevant to the $d = 4$ Coulomb phase.

Instead for $d \le 3$, $\ln W(C)$ increases as R^{4-d}. The reason is that two charges separated on C by a distance of order R, feel 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 confinement situations.

34.3.2 Non-abelian gauge theories

It follows from the discussion of Section 19.3 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. The same arguments as in the abelian case, show that the expectation value of the operator e^{-TH} in the corresponding state is given by the average, in the sense of the functional integral, of the parallel transporter along a closed loop:

$$W(C) = \left\langle \mathrm{P}\exp\left[-i\oint_C \mathbf{A}_\mu(s)\mathrm{d}s_\mu\right]\right\rangle, \tag{34.39}$$

in which we recall that 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, we find of course at leading order the same results as in the abelian case. However, we know from renormalization group, that we cannot trust perturbation theory at large distances. Therefore, to get a qualitative idea about the phase structure we first use the lattice model to calculate $W(C)$ in the large coupling or high temperature limit $\beta_p \to 0$.

Strong coupling expansion for Wilson's loop. We here assume that the group we consider has a *non-trivial centre.* We shall take the explicit example of gauge elements on the lattice belonging to the fundamental representation of $SU(N)$ (whose centre is $\mathbb{Z}_N$, with elements the identity multiplied by roots z of unity, $z^N = 1$).

We calculate $W(C)$ by expanding the integrand in expression (34.12) in powers of β_p. We choose for simplicity for the loop C a rectangle although the generalization to other contours is easy.

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 link 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 get a non-vanishing result, is at least equal to the area of the rectangle, the minimal area surface having the loop as boundary. We can then perform the integrations which are just factorized group integrations. In this way, we get 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. It is indeed 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 just get

$$W(C) \sim \mathrm{e}^{RT \ln \beta_p}. \tag{34.40}$$

This results indicates that the potential between the static charges is linearly rising at large distance. Static charges creating the loop cannot simply be screened by the gauge field, in which case we would again get 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. If, for example, the group is $SO(3)$, 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 (34.40) 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, but this question has not been definitively settled. 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). 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 33. The separation of vortices in the Kosterlitz–Thouless (KT) phase transition is here replaced by the separation of magnetic monopoles.

The string tension. The coefficient in front of the area is called the string tension σ,

$$\sigma(\beta_p) \underset{\beta_p\to 0}{\sim} -\ln\beta_p\,. \tag{34.41}$$

If no phase transition occurs when β_p varies from zero to infinity, the gauge theory leads to confinement. In this case, the behaviour of the string tension for β_p small is predicted by the renormalization group. Since σ has the dimension of a mass squared one finds

$$\sigma(e_0) \sim (e_0^2)^{-\beta_2/\beta_3^2} \exp\left(-1\,/\beta_2 e_0^2\right). \tag{34.42}$$

in which e_0^2 is related to β_p by equation (34.25) and β_2, β_3 are two first coefficients of the RG β-function which are given in equation (35.43). A physical quantity relevant to the continuum limit can then be obtained by dividing $\sqrt{\sigma}$ by its asymptotic behaviour. Let us define Λ_L as

$$\Lambda_L = a^{-1}(\beta_2 e_0^2)^{-\beta_3/2\beta_2^2} \exp\left(-1/2\beta_2 e_0^2\right), \tag{34.43}$$

then $\Lambda_L\,/\sqrt{\sigma}$ has a continuum limit. When one calculates σ by non-perturbative lattice methods, the verification of the scaling behaviour (34.42) indicates that the result is relevant to the continuum field theory and not only a lattice artifact.

It is possible to systematically expand σ in powers of β_p. The possibility of verifying that 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 to believe that, independently of the group, the string tension is affected by a singularity associated with the roughening transition, transition which, however, is not related to bulk properties. At strong coupling, the contributions to the large distance. Static charges creating the loop cannot simply be screened by the gauge field, in which case we would again get 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. If, for example, the group is $SO(3)$, 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 (34.40) 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, but this question has not been definitively settled. 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). 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 33. The separation of vortices in the Kosterlitz–Thouless (KT) phase transition is here replaced by the separation of magnetic monopoles.

The string tension. The coefficient in front of the area is called the string tension σ,

$$\sigma(\beta_p) \underset{\beta_p \to 0}{\sim} -\ln \beta_p \,. \tag{34.41}$$

If no phase transition occurs when β_p varies from zero to infinity, the gauge theory leads to confinement. In this case, the behaviour of the string tension for β_p small is predicted by the renormalization group. Since σ has the dimension of a mass squared one finds

$$\sigma(e_0) \sim (e_0^2)^{-\beta_2/\beta_3^2} \exp\left(-1\,/\beta_2 e_0^2\right) . \tag{34.42}$$

in which e_0^2 is related to β_p by equation (34.25) and β_2, β_3 are two first coefficients of the RG β-function which are given in equation (35.43). A physical quantity relevant to the continuum limit can then be obtained by dividing $\sqrt{\sigma}$ by its asymptotic behaviour. Let us define Λ_L as

$$\Lambda_L = a^{-1}(\beta_2 e_0^2)^{-\beta_3/2\beta_2^2} \exp\left(-1/2\beta_2 e_0^2\right) , \tag{34.43}$$

then $\Lambda_L\,/\sqrt{\sigma}$ has a continuum limit. When one calculates σ by non-perturbative lattice methods, the verification of the scaling behaviour (34.42) indicates that the result is relevant to the continuum field theory and not only a lattice artifact.

It is possible to systematically expand σ in powers of β_p. The possibility of verifying that 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 to believe that, independently of the group, the string tension is affected by a singularity associated with the roughening transition, 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. From the numerical point of view, it seems that the plaquette–plaquette correlation function is the most promising case for strong coupling expansion.

Remark. 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 instead a flux tube between static charges in such a way that the force remains the same as in one space dimension.

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 obtained by generalizing the arguments developed for ordinary symmetries in Sections 23.2 and 23.5. 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 different from zero only 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 which do not vanish at large distances. Therefore, it does not forbid a spontaneous breaking of the global symmetry associated with the gauge group.

34.4 Mean Field Approximation

We have shown that the pure gauge lattice model yields at low temperature or coupling the continuum gauge theory. The continuum model allows, in perturbation theory, the separation of charges at large distances. On the contrary, at high temperature, charges are confined in the lattice model.

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. From the numerical point of view, it seems that the plaquette–plaquette correlation function is the most promising case for strong coupling expansion.

Remark. 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 instead a flux tube between static charges in such a way that the force remains the same as in one space dimension.

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 obtained by generalizing the arguments developed for ordinary symmetries in Sections 23.2 and 23.5. 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 different from zero only 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 which do not vanish at large distances. Therefore, it does not forbid a spontaneous breaking of the global symmetry associated with the gauge group.

34.4 Mean Field Approximation

We have shown that the pure gauge lattice model yields at low temperature or coupling the continuum gauge theory. The continuum model allows, in perturbation theory, the separation of charges at large distances. On the contrary, at high temperature, charges are confined in the lattice model.

It is, therefore, 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 $d > 2$. It is, therefore, natural to also study gauge theories in the mean field approximation.

The general mean field formalism has been described in Appendix A24. We introduce two sets of real matrices ϕ_ℓ and H_ℓ, in which the index ℓ stands for link. Then we write the partition function

$$\mathcal{Z} = \int \prod_\ell \mathrm{d}U_\ell \exp\left[-\beta_p \mathcal{S}(U)\right], \tag{34.44}$$

in which $\mathcal{S}(U)$ is the lattice action (34.11), as

$$\mathcal{Z} = \int \prod_{\text{links } \ell} \mathrm{d}\phi_\ell \, \mathrm{d}H_\ell \, \mathrm{d}U_\ell \exp\left[-\beta_p \mathcal{S}(\phi) + \sum_{\text{links}} \operatorname{tr} H_\ell \left(\phi_\ell - U_\ell\right)\right]. \tag{34.45}$$

The introduction of the variables ϕ allows to express the action in terms of an average link variable. Since the average of an orthogonal (unitary) matrix is not orthogonal (resp. not unitary), we have defined ϕ as an arbitrary real (resp. complex) matrix. The variables H_ℓ represent directly at leading order the mean field which approximates the effect of the plaquette interaction.

The integral over the matrices U now factorizes into a product of integrals over each link variable:

$$\int \mathrm{d}U \, \mathrm{e}^{-\operatorname{tr} HU} = \mathrm{e}^{-\rho(H)}, \tag{34.46}$$

in which $\rho(H)$ is thus a $G \times G$ group invariant function of H (H transforming under independent right and left multiplication).

The partition function becomes

$$\mathcal{Z} = \int \prod_\ell \mathrm{d}H_\ell \, \mathrm{d}\phi_\ell \exp\left\{-\left[\beta_p \mathcal{S}(\phi) + \sum_\ell \left(\rho(H_\ell) - \operatorname{tr} H_\ell \phi_\ell\right)\right]\right\}. \tag{34.47}$$

We then 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 in which H_ℓ and ϕ_ℓ are constant on the lattice and multiple of the identity (up to a gauge transformation):

$$\phi_\ell = \varphi \mathrm{I}, \qquad H_\ell = h \mathrm{I}, \tag{34.48}$$

in which I is the identity matrix. Calling $\mathcal{S}(\varphi, h)$ the lattice action per link we then find

$$\mathcal{S}(h, \varphi) = \operatorname{tr} \mathrm{I} \left[-\tfrac{1}{2}(d-1)\beta_p \varphi^4 + V(h) - h\varphi\right], \tag{34.49}$$

in which we have defined $V(h)$ by

$$V(h) = \frac{\rho(h\mathrm{I})}{\operatorname{tr} \mathrm{I}} \tag{34.50}$$

with $V(h) = -\frac{1}{4}h^2 + O(h^4)$ for $SU(2)$. The saddle point equations are

$$\varphi = V'(h), \qquad h = -2(d-1)\beta_p \varphi^3. \tag{34.51}$$

We can eliminate φ and obtain

$$h = -2(d-1)\beta_p \left[V'(h)\right]^3 . \tag{34.52}$$

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 r.h.s. of equation (34.52) is at least cubic in h instead of being linear. Thus, the equation has never a non-trivial solution arbitrarily close to zero. For β_p small there exists only the trivial solution $h = 0$, which, 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 β_c, h jumps from zero 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 the non-linear σ-model. Above β_c, the expectation value of Wilson's loop is given by

$$W(C) = \left\langle \operatorname{tr} \prod_{\text{all links } \ell \in C} \phi_\ell \right\rangle \sim \phi^{P(C)}, \tag{34.53}$$

in which $P(C)$ is the perimeter of the loop. Therefore, Wilson's loop follows a perimeter law and the phase is deconfined. Above β_c we are in the low temperature phase which can be described by a continuum field theory and perturbation theory.

Discussion. As we have shown in Appendix A24, mean field theory is valid in high dimensions. Continuum field theory tells us that the zero temperature ($\beta_p = \infty$) is an IR stable fixed point for $d > 4$. Thus, the mean field result can only apply 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. The dimension $d = 4$ for gauge theories 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 37). 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 will not describe the numerical methods which have been used in lattice gauge theories. They are generalizations of the methods which we have briefly described in the case of simpler systems in Appendix A4.3. 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 has been determined.

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 already shown in Section 9.6 that non-trivial problems arise when one tries to construct a chiral invariant lattice action. One has the choice only between writing an action which is not explicitly chiral symmetric and in which one tries to restore chiral symmetry by adjusting the fermion mass term (Wilson's fermions), writing a chiral symmetric action with too many fermions (staggered or Kogut–Susskind fermions), or, as it has been more recently discovered various Dirac operators satisfying the Ginsparg–Wilson relation (Section 9.6.2). In the latter solution several implementations can be interpreted as adding for the fermions an extra space dimension, which increases the already difficult computer problem. Indeed, 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 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 up to now 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, the numerical study of the effect dynamical fermions is still in a preliminary stage.

Bibliographical Notes

Lattice gauge theories have been introduced in

K.G. Wilson, *Phys. Rev.* D10 (1974) 2445; F.J. Wegner, *J. Math. Phys.* 12 (1971) 2259.

Compact QED is studied in

A.M. Polyakov, *Phys. Lett.* 59B (1975) 82.

High temperature expansion is discussed in

R. Balian, J.M. Drouffe and C. Itzykson, *Phys. Rev.* D11 (1975) 2098; J.M. Drouffe and J.-B. Zuber, *Phys. Rep.* 102 (1983) 1.

Elitzur's theorem is derived in

S. Elitzur, *Phys. Rev.* D12 (1975) 3978.

For a review of the early development of lattice gauge theories, see

J.B. Kogut in *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, J.-B. Zuber and R. Stora eds. (North-Holland, Amsterdam 1984); *Rev. Mod. Phys.* 55 (1983) 775.

Calculations of the spectrum in non-abelian gauge theories can be found in

M. Lüscher, *Nucl. Phys.* B219 (1983) 233; M. Lüscher and G. Münster, *Nucl. Phys.* B232 (1984) 445; J. Koller and P. van Baal, *Nucl. Phys.* 302 (1988) 1; P. van Baal, *Nucl. Phys.* B307 (1988) 274.

Monte Carlo simulations were initiated by

M. Creutz, *Phys. Rev. Lett.* 45 (1980) 313.

Early articles on lattice gauge theories and Monte Carlo simulations are reprinted in

Lattice Gauge Theories and Monte Carlo Simulations, C. Rebbi ed. (World Scientific, Singapore 1983).

See also

M. Creutz, L. Jacobs and C. Rebbi, *Phys. Rep.* 95 (1983) 201; I. Montvay, *Rev. Mod. Phys.* 59 (1987) 263.

For Monte Carlo methods in statistical physics see, for example,

Monte Carlo Methods in Statistical Physics, K. Binder ed. (Springer-Verlag, New York 1979).

Relevant are also the lectures

M. Lüscher in *Critical Phenomena, Random Systems, Gauge Theories*, Les Houches 1984, K. Osterwalder and R. Stora eds. (Elsevier, Amsterdam 1986) and *Fields, Strings and Critical Phenomena*, Les Houches 1988, E. Brézin and J. Zinn-Justin eds. (Elsevier, Amsterdam 1989).

Finite temperature transitions are discussed in

B. Svetitsky, *Phys. Rep.* 132 (1986) 1; E. Marinari, *Phys. Rep.* 184 (1989) 131.

For pure gauge theories in two dimensions see, for example,

D.J. Gross and E. Witten, *Phys. Rev.* D21 (1980) 446.

APPENDIX A34
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 has no real dynamical degrees of freedom. Still the gauge field generates a force between particles which, as we have seen in Section 34.3, leads to confinement, even in the abelian case. An example of such a situation has been encountered in the Schwinger model in Section 32.4. This is a question that we here examine in more detail on the lattice for pure gauge theories.

Lattice gauge theories in two dimensions. Let us consider a pure gauge lattice action (34.11) in two dimensions. The partition function is given by

$$\mathcal{Z} = \int \prod_{\text{links}\{ij\}} \mathrm{d}\mathbf{U}_{ij} \exp\left(\beta \sum_{\text{plaquettes}} \operatorname{tr} \mathbf{U}_{ij}\mathbf{U}_{jk}\mathbf{U}_{kl}\mathbf{U}_{li}\right). \tag{A34.1}$$

We assume free boundary conditions in the time direction (to avoid closed loop variables surviving due to the topologic properties of the two-dimensional space manifold). We can then use the gauge invariance to set to 1 all link variables in the time direction. This is the equivalent of the temporal gauge of the continuum theory. However, in two dimensions this has the remarkable effect of decoupling all links in the orthogonal direction. The partition function factorizes. If we call L the size in this direction, we find

$$\mathcal{Z} = \mathcal{Z}_t^L \tag{A34.2}$$

with

$$\mathcal{Z}_t = \int \prod_i \mathrm{d}U_i \exp\left(\beta \sum_i \operatorname{tr} U_{i+1}{}^T U_i\right), \tag{A34.3}$$

in which i now is the coordinate in the euclidean time direction. We recognize the partition function of a simple one-dimensional model with two-body 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

$$V_{i+1} = U_{i+1}{}^T U_i\,. \tag{A34.4}$$

The integrals over the matrices V_i factorize and we find

$$\mathcal{Z}(\beta) = z^{\Omega}(\beta), \tag{A34.5}$$

in which Ω is the area of the lattice, and $z(\beta)$ is given by

$$z(\beta) = \int \mathrm{d}U \exp\left(\beta \operatorname{tr} U\right). \tag{A34.6}$$

We hereafter 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. It is also easy to calculate the expectation value of Wilson's loop. Calling R and T the sizes of the loop in space and time, one finds

$$W(C) = \mathrm{e}^{-RT\sigma(\beta)} \tag{A34.7}$$

APPENDIX A34
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 has no real dynamical degrees of freedom. Still the gauge field generates a force between particles which, as we have seen in Section 34.3, leads to confinement, even in the abelian case. An example of such a situation has been encountered in the Schwinger model in Section 32.4. This is a question that we here examine in more detail on the lattice for pure gauge theories.

Lattice gauge theories in two dimensions. Let us consider a pure gauge lattice action (34.11) in two dimensions. The partition function is given by

$$\mathcal{Z}=\int\prod_{\text{links}\{ij\}}\mathrm{d}\mathbf{U}_{ij}\exp\left(\beta\sum_{\text{plaquettes}}\operatorname{tr}\mathbf{U}_{ij}\mathbf{U}_{jk}\mathbf{U}_{kl}\mathbf{U}_{li}\right). \qquad (A34.1)$$

We assume free boundary conditions in the time direction (to avoid closed loop variables surviving due to the topologic properties of the two-dimensional space manifold). We can then use the gauge invariance to set to 1 all link variables in the time direction. This is the equivalent of the temporal gauge of the continuum theory. However, in two dimensions this has the remarkable effect of decoupling all links in the orthogonal direction. The partition function factorizes. If we call L the size in this direction, we find

$$\mathcal{Z}=\mathcal{Z}_t^L \qquad (A34.2)$$

with

$$\mathcal{Z}_t=\int\prod_i\mathrm{d}U_i\exp\left(\beta\sum_i\operatorname{tr}U_{i+1}{}^TU_i\right), \qquad (A34.3)$$

in which i now is the coordinate in the euclidean time direction. We recognize the partition function of a simple one-dimensional model with two-body 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

$$V_{i+1}=U_{i+1}{}^TU_i\,. \qquad (A34.4)$$

The integrals over the matrices V_i factorize and we find

$$\mathcal{Z}(\beta)=z^{\Omega}(\beta), \qquad (A34.5)$$

in which Ω is the area of the lattice, and $z(\beta)$ is given by

$$z(\beta)=\int\mathrm{d}U\exp\left(\beta\operatorname{tr}U\right). \qquad (A34.6)$$

We hereafter 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. It is also easy to calculate the expectation value of Wilson's loop. Calling R and T the sizes of the loop in space and time, one finds

$$W(C)=\mathrm{e}^{-RT\sigma(\beta)} \qquad (A34.7)$$

with

$$\sigma(\beta) = -\ln\left(\frac{z'(\beta)}{Nz(\beta)}\right). \tag{A34.8}$$

We have again assumed that the centre of the group is non-trivial and called N the trace of the unit matrix. As expected, Wilson's loop has an area law for all groups and all couplings in two dimensions.

For β small, $z(\beta)$ has the expansion

$$z(\beta) = 1 + z_2\beta^2 + O\left(\beta^4\right). \tag{A34.9}$$

This yields for the string tension

$$\sigma(\beta) \sim -\ln\beta\,, \tag{A34.10}$$

in agreement with expression (34.40). For β large, $z(\beta)$ can be calculated by steepest descent which is also perturbation theory, and one finds

$$\ln z(\beta) = N(\beta - K\ln\beta) + O\left(\beta^{-1}\right), \tag{A34.11}$$

in which K is a constant. Therefore,

$$\sigma(\beta) \sim K/\beta\,, \tag{A34.12}$$

in agreement with perturbation theory in the continuum.

Finally, since in two dimensions the gauge field has no physical degrees of freedom, no particle propagates, and no gluonium state can be found.

35 UV FIXED POINTS IN QUANTUM FIELD THEORY

Up to now, we have mainly discussed the IR behaviour of field theories. In this chapter, we use RG equations to characterize instead the large momentum behaviour of renormalized field theories. This assumes implicitly that a universal large momentum physics, that is, a property of the continuum, can be defined. This implies also the existence of a crossover scale between low and large momentum physics.

It has been observed experimentally that the quarks, fundamental particles of the theory of Strong Interactions, behave like free particles at the shortest distances presently accessible. Therefore, the discussion of the large momentum behaviour and the identification of field theories which behave approximately as free field theories at short distance is directly relevant to Particle Physics. Examining the large momentum behaviour of all field theories renormalizable in four dimensions, we show that only theories having a non-abelian gauge symmetry can be asymptotically free, that is, behave as a free field theory at large euclidean momenta (Coleman–Gross theorem). We begin our investigation with scalar ϕ^4-like field theories and then add fermions and gauge fields. As an application, we calculate the total cross section of electron–positron annihilation into hadrons at large momentum.

Another more theoretical reason for discussing the large momentum behaviour is the apparent connection between the existence of non-trivial renormalized quantum field theories and the presence of UV fixed points. The absence of identified UV fixed points in theories like the ϕ^4 field theory or QED leads to the so-called triviality problem which we examine in Section 35.1.1. If we then consider the scalar self-interaction of ϕ^4 type as the main source in the Standard Model of the Higgs mass, we obtain a semi-quantitative bound on the Higgs mass (see Section 35.1.2).

Before discussing these issues, we want to again emphasize the existence of two ways of understanding quantum field theory.

As we have explained several times in Chapters 25 and 26 in the discussion of critical phenomena, we consider the bare theory with cut-off as real in the sense that it is an effective representation of some unknown underlying microscopic physics.

In another presentation of quantum field theory, as applied to Particle Physics, the bare theory is a formal intermediate step to construct the renormalized theory, only the latter one being physical.

In what follows we take the former point of view but some arguments will involve only the renormalized theory, and, therefore, will be valid from all point of views.

35.1 The $(\phi^2)^2$ Field Theory: Large Momentum Behaviour and Triviality

Large momentum behaviour can entirely be discussed in the framework of the massless theory. We thus consider the massless ϕ^4 field theory and the corresponding RG equations. Therefore, the extension of the analysis to dimensions $d < 4$ requires the framework of the $\varepsilon = 4 - d$ expansion, even though the large momentum behaviour can also be discussed at fixed dimension in the massive theory.

We consider the bare action

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} r_c \phi^2 + \frac{1}{4!} g_0 \Lambda^{4-d} \left(\phi^2\right)^2\right], \tag{35.1}$$

where g_0 is dimensionless and a cut-off Λ is implied. In Section 25.4, we have derived the bare RG equation (equation (25.51)):

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0} - \frac{n}{2} \eta(g_0)\right] \Gamma^{(n)}(p_i; g_0, \Lambda) = 0\,, \tag{35.2}$$

up to corrections decreasing like power of Λ, with

$$\beta(g_0) = (d-4) g_0 + \frac{N+8}{48\pi^2} g_0^2 + O(g_0^3), \quad \eta(g_0) = \frac{N+2}{72(8\pi^2)^2} g_0^2 + O(g_0^3).$$

Dimensions $d < 4$. In Section 25.5, we have solved the RG equations by the 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 smaller than the IR fixed point $< g_0^*$, the effective coupling at scale λ, $g_0(\lambda)$, decreases, and $g_0 > g_0^*$ it increases, and of course if initially $g_0 = g_0^*$ it is invariant.

Since increasing momenta amounts to decrease the cut-off, a question arises: does one find some universal (continuum) non-IR physics before the momenta become of order Λ. Additional information is provided by solving the RG equation differently, setting

$$\Gamma^{(n)}(p_i; g_0, \Lambda) = \zeta^{n/2}(g_0) [M(g_0)]^{d-n(d-2)/2} G^{(n)}(p_i/M(g_0)),$$

where $M(g_0)$ is a RG invariant mass scale:

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0}\right) M(g_0) = 0\,.$$

For $g_0 < g_0^*$

$$M(g_0) = \Lambda g_0^{1/(4-d)} \exp\left[\int_{g_0}^{0} \mathrm{d}g' \left(\frac{1}{\beta(g')} - \frac{1}{(d-4)g'}\right)\right] \tag{35.3}$$

Moreover,

$$\ln \zeta(g_0) = \int_0^{g_0} \mathrm{d}g' \frac{\eta(g')}{\beta(g')}.$$

Since functions depend only the mass scale M_0, a continuum large momentum behaviour can be found only if one can find momenta of order $M(g_0)$ or larger but still much smaller than Λ. This yields for $g_0 < g_0^*$ the condition

$$M(g_0) \ll \Lambda \;\Rightarrow\; g_0 \ll 1\,. \tag{35.4}$$

This condition can be satisfied only if the initial coupling constant g_0 is close to a UV fixed point value. In the $(\phi^2)^2$ theory, in dimensions $d < 4$, this is realized if g_0 is very small, close to the gaussian fixed point $g_0 = 0$. Then, the mass scale $M(g_0)$ is a crossover scale between IR and UV behaviours.

The condition (35.4) is equivalent to demand that the effective coupling constant $g_0(\mu/\Lambda)$ at scale $\mu \ll \Lambda$, which is also a possible definition of the renormalized coupling g defined at renormalization scale μ,

$$g \sim g_0(\mu/\Lambda), \quad \int_{g_0}^{g_0(\mu/\Lambda)} \frac{\mathrm{d}g'}{\beta(g')} = \ln(\mu/\Lambda), \tag{35.5}$$

should be significantly different from its IR fixed point value.

35.1.1 The renormalized ϕ^4_4 field theory

The gaussian fixed point. A first conclusion, important 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 represent the physics of Strong Interactions at experimentally accessible short distances.

Triviality. We now show that the same arguments allow us to discuss the existence of a non-trivial renormalized ϕ^4 field theory in four dimensions.

The question we investigate here is whether it is possible to find, for arbitrarily large values of the cut-off Λ, a bare coupling constant g_0 which yields a given renormalized coupling constant g. This problem of course has always a formal perturbative solution but we want to discuss this question beyond perturbation theory.

Again, we can solve the massless RG equations in the same way, and introduce the RG invariant scale

$$M(g_0) = \Lambda \exp\left[\int_{g_0}^{a} \frac{\mathrm{d}g'}{\beta(g')}\right],$$

where a is a (small) fixed value. The condition

$$M(g_0) \ll \Lambda\,,$$

now plays an even more important role. Because the IR fixed point corresponds to $g_0 = 0$, the existence of a non-trivial continuum UV behaviour is also the condition for the existence of a non-free ϕ^4 field theory in the infinite cut-off limit. The same condition is recovered from equation (35.5):

$$\int_{g_0}^{g} \frac{\mathrm{d}g'}{\beta(g')} = \ln(\mu/\Lambda).$$

For $\mu/\Lambda \to 0$ either g goes to the IR fixed point $g = 0$, 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. Various considerations, however, 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 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 regularized by the lattice by taking the large g_0 limit.

In the case of a one component system (Ising-like), we have presented conclusive numerical evidence in Chapter 29 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 has found no evidence of additional fixed points and are consistent with triviality.

35.1.1 The renormalized ϕ_4^4 field theory

The gaussian fixed point. A first conclusion, important 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 represent the physics of Strong Interactions at experimentally accessible short distances.

Triviality. We now show that the same arguments allow us to discuss the existence of a non-trivial renormalized ϕ^4 field theory in four dimensions.

The question we investigate here is whether it is possible to find, for arbitrarily large values of the cut-off Λ, a bare coupling constant g_0 which yields a given renormalized coupling constant g. This problem of course has always a formal perturbative solution but we want to discuss this question beyond perturbation theory.

Again, we can solve the massless RG equations in the same way, and introduce the RG invariant scale

$$M(g_0) = \Lambda \exp\left[\int_{g_0}^{a} \frac{\mathrm{d}g'}{\beta(g')}\right],$$

where a is a (small) fixed value. The condition

$$M(g_0) \ll \Lambda\,,$$

now plays an even more important role. Because the IR fixed point corresponds to $g_0 = 0$, the existence of a non-trivial continuum UV behaviour is also the condition for the existence of a non-free ϕ^4 field theory in the infinite cut-off limit. The same condition is recovered from equation (35.5):

$$\int_{g_0}^{g} \frac{\mathrm{d}g'}{\beta(g')} = \ln(\mu/\Lambda).$$

For $\mu/\Lambda \to 0$ either g goes to the IR fixed point $g = 0$, 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. Various considerations, however, 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 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 regularized by the lattice by taking the large g_0 limit.

In the case of a one component system (Ising-like), we have presented conclusive numerical evidence in Chapter 29 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 has 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 30 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 30.6. Again, we have found that it becomes a free field theory in four dimensions. Therefore, we have strong evidence that no non-trivial continuum 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.

The Higgs model. 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 Standard Model of weak electromagnetic interactions which contains a $(\phi^2)^2$ interaction. Of course then other coupling constants contribute to the ϕ^4 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, we realize that the model has only been tested in some limited range of energies. Therefore, demanding that the quantum field theory should be consistent on all scales has no real physical 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 will give reasonable answers. Correspondingly the renormalized coupling will be allowed to vary in a limited range, which goes to zero (logarithmically) for large cut-off. In this sense, the Standard Model 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, even though the theory is perturbatively renormalizable: the theory is not consistent on all scales.

This conclusion is specially unavoidable if the ϕ^4 Higgs coupling is not very small. Indeed since the gauge and Yukawa couplings are quite small (even in the case of the top quark), the Higgs coupling then probably dominates the Higgs sector and a purely ϕ^4 analysis is relevant.

One may thus wonder about the meaning of the correlation functions as defined by renormalized perturbation theory. We argue in Chapter 42 that there are intrinsic difficulties in the reconstruction of correlation functions from the knowledge of their perturbative expansion. However, irrespective of this problem, it follows from RG arguments that to a finite renormalized coupling constant can only correspond a complex bare coupling constant, the imaginary part vanishing at all orders in perturbation theory. As a consequence, the correlation functions, depending on the summation procedure, will either be complex, or will not satisfy field equations beyond perturbation theory.

35.1.2 An upper-bound on the Higgs mass

The Higgs field through its various couplings gives masses to all other fields. The observed masses determine the corresponding couplings. Only the Higgs mass and thus the Higgs self-coupling are unknown parameters. Note, however, that it is likely that the renormalized ϕ^4 coupling g is such that perturbation theory remains at least semi-quantitatively applicable. Otherwise, the successes of the Standard Model would be 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. As we have argued above for g large enough probably the Higgs mass is mostly determined by the Higgs self-coupling. We, therefore, examine below the pure ϕ^4 field theory. Since we remain in the perturbative regime RG arguments are still applicable. A bound on m_H, the Higgs mass, can then be derived.

We solve equation (35.5), using the expansion of the β-function,

$$\beta(g_0) = \beta_2 g_0^2 + \beta_3 g_0^3 + O\left(g_0^4\right),$$

for g small. We find

$$\ln(\Lambda/\mu) = \frac{1}{\beta_2 g} + \frac{\beta_3}{\beta_2^2} \ln g + K(g_0) + O(g), \tag{35.6}$$

where the function $K(g_0)$, according to the previous discussion, is bounded but can only be determined by non-perturbative methods. For g small, we can use perturbation theory to relate the ϕ (=Higgs) expectation value, which is known from the Z mass ($\langle\phi\rangle \sim 250\,\mathrm{GeV}$), and the Higgs mass (see Section 27.2). At leading order, we find

$$m_H^2 = \tfrac{1}{3} g \left\langle\phi\right\rangle^2 + O\left(g^2\right). \tag{35.7}$$

To minimize higher order corrections, we choose for g the renormalized coupling constant at scale $\langle\phi\rangle$. We can then eliminate g between equations (35.7) and (35.6), and find

$$\ln\left(\frac{\Lambda}{\langle\phi\rangle}\right) = \frac{1}{3\beta_2}\frac{\langle\phi\rangle^2}{m_H^2} + \frac{2\beta_3}{\beta_2^2}\ln\left(\frac{m_H}{\langle\phi\rangle}\right) + \tilde{K}(g_0) + O(g).$$

If we assume that we can neglect in the r.h.s. all terms but the two first ones, we obtain a relation between the two ratios $\Lambda/\langle\phi\rangle$ and $m_H/\langle\phi\rangle$. Moreover, if the Higgs is really associated to a physical particle its mass must be smaller than the cut-off (which at this point only represents the onset of new physics beyond the Standard Model). Taking for the two coefficients of the β-function, the values from equation (29.3) for $N = 4$, $8\pi^2\beta_2 = 2$, $\beta_3/\beta_2^2 = -13/24$, we obtain an upper-bound for m_H:

$$m_H < 2.6\,\langle\phi\rangle \;\Rightarrow m_H < 640\ \mathrm{GeV},$$

($\langle\phi\rangle \approx 250$ GeV) value which can be compared with computer simulation values which vary in the range 670–700 GeV. Note that the corresponding value of g is such that perturbation theory at leading order should still be semi-quantitatively correct.

Conversely, if we know the physical coupling constant at scale μ, we can infer from equation (35.6) an upper-bound on the cut-off or scale of new physics. Clearly, this bound is very sensitive to small corrections since the equation determines $\ln(\Lambda/\mu)$.

35.1.3 The ϕ_4^4 field theory for negative renormalized coupling

In the preceding analysis, we have assumed that the renormalized coupling constant is always positive. It has been argued that the renormalized coupling constant could also be negative. This would mean that the identification between renormalized coupling and effective bare coupling is impossible, and the bare renormalization group somewhat pathological.

We have, therefore, to discuss this issue entirely within the framework of the renormalized theory.

However, a question arises immediately: does this theory correspond to a hamiltonian bounded from below? We have shown in Section 7.10 that the 1PI functional is related theory. Since we remain in the perturbative regime RG arguments are still applicable. A bound on m_H, the Higgs mass, can then be derived.

We solve equation (35.5), using the expansion of the β-function,

$$\beta(g_0) = \beta_2 g_0^2 + \beta_3 g_0^3 + O\left(g_0^4\right),$$

for g small. We find

$$\ln(\Lambda/\mu) = \frac{1}{\beta_2 g} + \frac{\beta_3}{\beta_2^2} \ln g + K(g_0) + O(g), \tag{35.6}$$

where the function $K(g_0)$, according to the previous discussion, is bounded but can only be determined by non-perturbative methods. For g small, we can use perturbation theory to relate the ϕ (=Higgs) expectation value, which is known from the Z mass ($\langle\phi\rangle \sim 250\,\text{GeV}$), and the Higgs mass (see Section 27.2). At leading order, we find

$$m_H^2 = \tfrac{1}{3} g \langle\phi\rangle^2 + O\left(g^2\right). \tag{35.7}$$

To minimize higher order corrections, we choose for g the renormalized coupling constant at scale $\langle\phi\rangle$. We can then eliminate g between equations (35.7) and (35.6), and find

$$\ln\left(\frac{\Lambda}{\langle\phi\rangle}\right) = \frac{1}{3\beta_2} \frac{\langle\phi\rangle^2}{m_H^2} + \frac{2\beta_3}{\beta_2^2} \ln\left(\frac{m_H}{\langle\phi\rangle}\right) + \tilde{K}(g_0) + O(g).$$

If we assume that we can neglect in the r.h.s. all terms but the two first ones, we obtain a relation between the two ratios $\Lambda/\langle\phi\rangle$ and $m_H/\langle\phi\rangle$. Moreover, if the Higgs is really associated to a physical particle its mass must be smaller than the cut-off (which at this point only represents the onset of new physics beyond the Standard Model). Taking for the two coefficients of the β-function, the values from equation (29.3) for $N = 4$, $8\pi^2\beta_2 = 2$, $\beta_3/\beta_2^2 = -13/24$, we obtain an upper-bound for m_H:

$$m_H < 2.6\,\langle\phi\rangle \ \Rightarrow m_H < 640\ \text{GeV},$$

($\langle\phi\rangle \approx 250$ GeV) value which can be compared with computer simulation values which vary in the range 670–700 GeV. Note that the corresponding value of g is such that perturbation theory at leading order should still be semi-quantitatively correct.

Conversely, if we know the physical coupling constant at scale μ, we can infer from equation (35.6) an upper-bound on the cut-off or scale of new physics. Clearly, this bound is very sensitive to small corrections since the equation determines $\ln(\Lambda/\mu)$.

35.1.3 The ϕ_4^4 field theory for negative renormalized coupling

In the preceding analysis, we have assumed that the renormalized coupling constant is always positive. It has been argued that the renormalized coupling constant could also be negative. This would mean that the identification between renormalized coupling and effective bare coupling is impossible, and the bare renormalization group somewhat pathological.

We have, therefore, to discuss this issue entirely within the framework of the renormalized theory.

However, a question arises immediately: does this theory correspond to a hamiltonian bounded from below? We have shown in Section 7.10 that the 1PI functional is related to the average of the hamiltonian in states with fixed field expectation value. We have seen in Section 26.5 that RG equations relate the thermodynamic potential for small magnetization to IR fixed points. Not surprisingly the large constant field behaviour of the 1PI functional is governed by UV fixed points. Following the arguments of Section 26.5, we can derive the RG equation satisfied by $\mathcal{V}(\varphi)$, the thermodynamic potential density at constant field. We consider a massless theory renormalized at scale μ, and subtract $\mathcal{V}(\varphi)$ at $\varphi = 0$. We obtain the equation

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} - \frac{1}{2}\eta(g)\varphi\frac{\partial}{\partial\varphi}\right]\mathcal{V}(\varphi, g, \mu) = 0\,. \tag{35.8}$$

Solving the equation by the method of characteristics and using dimensional analysis, we find

$$\mathcal{V}(\lambda\varphi, g, \mu) = \lambda^4\mathcal{V}\bigl(Z^{1/2}(\lambda)\varphi, g(\lambda), \mu\bigr), \tag{35.9}$$

in which $g(\lambda)$ and $Z(\lambda)$ are defined by equations in terms of the corresponding RG functions by

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}g(\lambda) = \beta\bigl(g(\lambda)\bigr), \quad \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln Z(\lambda) = -\eta\bigl(g(\lambda)\bigr). \tag{35.10}$$

The solutions are

$$\ln\lambda = \int_g^{g(\lambda)}\frac{\mathrm{d}g'}{\beta(g')}, \quad Z(\lambda) = \exp\left[-\int_1^\lambda\frac{\mathrm{d}\sigma}{\sigma}\eta\bigl(g(\sigma)\bigr)\right]. \tag{35.11}$$

We see that to study the large field behaviour of $\mathcal{V}(\varphi)$, we have to increase λ and, therefore, study the UV limit of the theory. If we start from $g < 0$ and small, $g(\lambda)$ approaches the origin for λ large. The field renormalization $Z(\lambda)$ then tends towards 1, 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 $\boldsymbol{\varphi}$, we obtain

$$\mathcal{V}(\lambda\varphi, g, \mu) \sim \frac{1}{4!}\lambda^4 g(\lambda)\varphi^4, \qquad g(\lambda) \sim -\frac{48\pi^2}{(N+8)\ln\lambda}\,. \tag{35.12}$$

The consequences of this result are obvious: by increasing λ, $\mathcal{V}$ can be made arbitrarily large and negative. It follows that the corresponding hamiltonian is not bounded from below.

35.2 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 a massless field $\phi_i(x)$ which in the tree approximation has the form

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x\left[\frac{1}{2}\sum_i\left(\partial_\mu\phi_i\right)^2 + \frac{\mu^{4-d}}{4!}\sum_{ijkl} g_{ijkl}\phi_i\phi_j\phi_k\phi_l\right]. \tag{35.13}$$

The corresponding RG equations have been established in Section 11.6. The renormalized n-point correlation functions satisfy equation (11.89):

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_{i'j'k'l'}\frac{\partial}{\partial g_{i'j'k'l'}}\right)\Gamma^{(n)}_{i_1 i_2\ldots i_n} - \frac{1}{2}\sum_{m=1}^{n}\eta_{i_m j_m}\Gamma^{(n)}_{i_1 i_2\ldots j_m\ldots i_n} = 0\,. \tag{35.14}$$

We also need the explicit leading order of β_{ijkl} (equation (11.94)) in four dimensions:

$$\beta_{ijkl} = \frac{1}{16\pi^2}\left(g_{ijmn}g_{mnkl} + g_{ikmn}g_{mnjl} + g_{ilmn}g_{mnjk}\right) + O\left(|g|^3\right). \tag{35.15}$$

Moreover, we know that η_{ij} is of order g^2. Following the arguments given in Chapter 26, it is easy to verify that the thermodynamic potential density in a constant field φ, and subtracted at $\varphi = 0$, satisfies

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_{i'j'k'l'}\frac{\partial}{\partial g_{i'j'k'l'}} - \frac{1}{2}\eta_{ij}\varphi_i\frac{\partial}{\partial\varphi_j}\right)\mathcal{V}(\varphi) = 0\,. \tag{35.16}$$

We again solve this equation by the method of characteristics (see Chapters 25 and 26). Introducing a scale parameter λ, scale dependent coupling constants $g_{ijkl}(\lambda)$ and field renormalization matrix $Z_{ij}(\lambda)$ defined by

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}g_{ijkl}(\lambda) = \beta_{ijkl}\left(g(\lambda)\right), \qquad g_{ijkl}(1) = g_{ijkl}\,, \tag{35.17}$$

$$\left(\lambda\frac{\mathrm{d}Z^{1/2}}{\mathrm{d}\lambda}Z^{-1/2}\right)_{ij} = -\frac{1}{2}\eta_{ij}\left(g(\lambda)\right), \qquad Z_{ij}(1) = \delta_{ij}\,, \tag{35.18}$$

we obtain

$$\mathcal{V}(\lambda\varphi_i, g_{ijkl}, \mu) = \lambda^4\mathcal{V}\left(Z_{ij}^{1/2}(\lambda)\varphi_j, g_{ijkl}(\lambda), \mu\right). \tag{35.19}$$

We know from the analysis of Section 35.1.3 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 we 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:

$$\mathcal{V}(\lambda\varphi) \sim \lambda^4\frac{g_{ijkl}(\lambda)}{4!}\varphi_i\varphi_j\varphi_k\varphi_l\,. \tag{35.20}$$

We have immediately taken into account that for g_{ijkl} small the renormalization matrix Z_{ij} goes to 1. The boundness of the hamiltonian implies, therefore, that $g_{ijkl}(\lambda)\varphi_i\varphi_j\varphi_k\varphi_l$ must be a non-negative quartic form. Let us examine the motion of the quantity

$$G(\lambda) = g_{ijkl}(\lambda)\varphi_i\varphi_j\varphi_k\varphi_l\,, \tag{35.21}$$

using the explicit form of the β function given by equation (35.15):

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}G(\lambda) = \frac{3}{16\pi^2}\varphi_i\varphi_j g_{ijmn}(\lambda)\varphi_k\varphi_l g_{klmn}(\lambda). \tag{35.22}$$

The r.h.s. 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 r.h.s. vanish:

$$\varphi_i\varphi_j g_{ijmn}(\lambda) = 0\,, \qquad \forall m, n\,. \tag{35.23}$$

However, $G(\lambda)$ then vanishes identically. The argument is valid for all vectors φ. This implies that we started from a free field theory.

We thus 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 continuum limit.

35.3 Theories with Scalar Bosons and Fermions

We now consider theories renormalizable in four dimensions and involving only scalar bosons and fermions. We first discuss a theory with one scalar and one Dirac fermion field for which we have already calculated the RG β-functions in Section 11.7.

A simple example. In the case of one scalar boson field coupled to one spin 1/2 Dirac field through an interaction term of the form

$$\mathcal{S}_{\mathrm{Int}} = \int \mathrm{d}^4 x \left[-u\bar{\psi}(x)\psi(x)\phi(x) + \frac{g}{4!}\phi^4(x) \right], \tag{35.24}$$

the RG β-functions at one-loop order in four dimensions are given by equations (11.132):

$$\begin{aligned} \beta_{u^2} &= \frac{5}{8\pi^2} u^4 + O\left(u^6, u^2 g^2\right), \\ \beta_g &= \frac{1}{8\pi^2}\left(\frac{3}{2}g^2 + 4gu^2 - 24u^4\right) + O\left(g^3, g^2u^2, gu^4, u^6\right). \end{aligned} \tag{35.25}$$

The fermions generate negative contributions to the ϕ^4 coupling RG function, which is no longer obviously positive. Instead β_{u^2} now is strictly positive. Therefore, for u^2 small the running coupling constant $u^2(\lambda)$ increases for large λ. Since u^2 is positive, it grows in absolute value and the theory cannot be asymptotically free. There is, however, one case which must be examined separately: if u is of order g then the two-loop contribution of order $u^2 g^2$ 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 (11.69) and has to be added in equation (11.126). The function β_{u^2} then becomes

$$\beta_{u^2} = \frac{5}{8\pi^2}u^4 + \frac{1}{48(8\pi^2)^2}u^2 g^2 + O\left(u^6, u^2 g^3\right). \tag{35.26}$$

Therefore, this additional term is also positive and the conclusion is the same.

The general case. The most general interaction, renormalizable in four dimensions, has the form

$$\mathcal{S}_{\mathrm{Int}}(\phi, \bar{\psi}, \psi) = \int \mathrm{d}^4 x \left[\frac{1}{4!} g_{ijkl}\phi_i\phi_j\phi_k\phi_l - \bar{\psi}_a \left(u^i_{ab} + i\gamma_5 v^i_{ab}\right)\phi_i\psi_b \right], \tag{35.27}$$

in which $(u^i)_{ab}$ and $(v^i)_{ab}$ are hermitian matrices (see Appendix A8). Since the diagrams contributing to the RG functions have been calculated in the one component case in Section 11.7, we just have to take into account the additional geometric factors. It is convenient to set

$$z^i_{ab} = u^i_{ab} + iv^i_{ab}\,. \tag{35.28}$$

A remark simplifies the calculation. If one calculates the Feynman diagrams in the 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 $\mathbf{z}^{i\dagger}$. The renormalization constants then are

$$\mathbf{Z}_\psi = \mathbf{I} - \frac{1}{16\pi^2\varepsilon}\mathbf{z}^{i\dagger}\mathbf{z}^i + O(2 \text{ loops}), \tag{35.29}$$

$$\left(\mathbf{Z}_\phi\right)_{ij} = \delta_{ij} - \frac{1}{8\pi^2\varepsilon}\operatorname{tr}\left(\mathbf{z}^{i\dagger}\mathbf{z}^j + \mathbf{z}^{j\dagger}\mathbf{z}^i\right) + O(2 \text{ loops}), \tag{35.30}$$

($\mathbf{I}$ 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:

$$16\pi^2\beta_z^i = \tfrac{1}{2}\left(\mathbf{z}^j\mathbf{z}^{j\dagger}\mathbf{z}^i + \mathbf{z}^i\mathbf{z}^{j\dagger}\mathbf{z}^j\right) + \operatorname{tr}\left(\mathbf{z}^{i\dagger}\mathbf{z}^j\right)\mathbf{z}^j + \operatorname{tr}\left(\mathbf{z}^i\mathbf{z}^{j\dagger}\right)\mathbf{z}^j + 2\mathbf{z}^j\mathbf{z}^{i\dagger}\mathbf{z}^j . \tag{35.31}$$

Let us write the flow equation for the quantity $\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^i$:

$$\begin{aligned} 8\pi^2\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^i &= \tfrac{1}{2}\left(\operatorname{tr}\mathbf{z}^j\mathbf{z}^{j\dagger}\mathbf{z}^i\mathbf{z}^{i\dagger} + \operatorname{tr}\mathbf{z}^{j\dagger}\mathbf{z}^j\mathbf{z}^{i\dagger}\mathbf{z}^i\right) + \operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j \\ &\quad + \operatorname{tr}\mathbf{z}^i\mathbf{z}^{j\dagger}\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j + 2\operatorname{tr}\mathbf{z}^j\mathbf{z}^{i\dagger}\mathbf{z}^j\mathbf{z}^{i\dagger} . \end{aligned} \tag{35.32}$$

The matrices $\mathbf{z}^i\mathbf{z}^{i\dagger}$ and $\mathbf{z}^{i\dagger}\mathbf{z}^i$ are positive, therefore, the two first terms in the r.h.s. of equation (35.32), 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. Indeed, we have

$$\begin{aligned} \operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j &= \operatorname{Re}\left(\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\right) \\ &\le \left(\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\right)^*\left(\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\right) = \operatorname{tr}\mathbf{z}^i\mathbf{z}^{j\dagger}\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j . \end{aligned} \tag{35.33}$$

It follows that

$$8\pi^2\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^i \ge 2\left[\left(\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\right)\left(\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^j\right) + \operatorname{tr}\mathbf{z}^j\mathbf{z}^{i\dagger}\mathbf{z}^j\mathbf{z}^{i\dagger}\right] . \tag{35.34}$$

Let us introduce the four-point vertex

$$M_{abcd} = z^i_{ab}z^i_{cd} . \tag{35.35}$$

It is easy to verify that equation (35.34) can then be rewritten as

$$8\pi^2\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^i \ge \left(M_{abcd} + M_{adbc}\right)\left(M^*_{abcd} + M^*_{adbc}\right) , \tag{35.36}$$

which proves that the r.h.s. is positive. We conclude that $\operatorname{tr}\mathbf{z}^{i\dagger}\mathbf{z}^i(\lambda)$ 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 $g_{iklm}g_{klmj}z^j_{ab}$, with a positive coefficient as we have seen before. It contributes in equation (35.32) by a term proportional to $g_{iklm}g_{klmj}z^j_{ab}z^{i\dagger}_{ba}$ that is positive:

$$g_{iklm}g_{klmj}z^j_{ab}z^{i\dagger}_{ab} = \left(g_{iklm}z^{i*}_{ab}\right)\left(g_{jklm}z^j_{ab}\right) . \tag{35.37}$$

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.

35.4 Gauge Theories

We have calculated in Section 18.9, the RG β-function for QED. In four dimensions, in a theory with n_{F} fermions and n_{B} bosons of charge e the β-function reads

$$\beta(e^2) = (4n_{\mathrm{F}} + n_{\mathrm{B}}) \frac{1}{3} \frac{e^4}{8\pi^2} + O\left(e^6\right). \tag{35.38}$$

Therefore, QED is IR free in four dimensions and like the ϕ^4 theory 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.

In Section 20.2.2, we have calculated the β-function for purely non-abelian gauge theories and found

$$\beta(g^2) = -\frac{g^4}{8\pi^2} \frac{11}{6} C(G) + O\left(g^6\right), \tag{35.39}$$

in which $C(G)$ is the Casimir of the group G. As we have already emphasized, 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 language 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.

Non-abelian gauge theories could, therefore, be consistent on all scales, but for the existence of low energy physics one condition must 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 condition is not too severe.

Gauge theories and fermions. In Section 20.2.2, 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

$$\operatorname{tr} t^a t^b = -\delta_{ab} T(R), \tag{35.40}$$

the β-function reads (equation (20.72))

$$\beta(g^2) = -\tfrac{1}{3}\left(11C(G) - 4T(R)\right) \frac{g^4}{8\pi^2} + O\left(g^6\right). \tag{35.41}$$

This result is sometimes expressed in terms of the equivalent of the fine structure constant $\alpha_s = g^2/4\pi$.

Before any calculation, we knew that the contribution of fermions would be 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,

$$C(G) = N \quad \text{for } SU(N), \qquad T(R) = \tfrac{1}{2} N_F,$$

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* $SU(3)$, asymptotic freedom imposes

$$N_F < 33/2\,. \tag{35.42}$$

For completeness, we also give here the β-function at two-loop order:

$$\beta(g^2) = \beta_2 g^4 + \beta_3 g^6 + O\left(g^8\right) \tag{35.43}$$

with

$$8\pi^2\beta_2 = -\tfrac{1}{3}\left[11C(G) - 4T(R)\right], \tag{35.44}$$

$$\left(8\pi^2\right)^2\beta_3 = -\tfrac{1}{6}\left[17C^2(G) - 6C(R)T(R) - 10C(G)T(R)\right]. \tag{35.45}$$

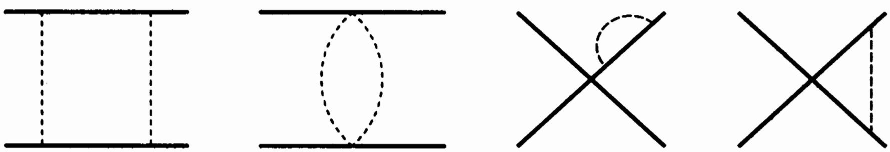
We recall that, as shown in Section 10.11, these two first coefficients are independent of the renormalization scheme. For the $SU(N)$ group, with N_F flavours in the fundamental representation, one finds

$$\left(8\pi^2\right)^2\beta_3 = -\tfrac{1}{6}\left[17N^2 - 5NN_F - 3N_F(N^2-1)/(2N)\right]. \tag{35.46}$$

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 (18.115), 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 35.2, by itself does not allow 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

$$\beta_u = au^2 + 2bug^2 + cg^4. \tag{35.47}$$

We know that a is positive. It is easy to verify that the contributions to c of the two diagrams (a) of figure 35.1 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 figure 35.1).



(a) (b)

Fig. 35.1 Dashed lines correspond to gauge fields.

From the form (35.46) 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

$$v = u/g^2. \tag{35.48}$$

Setting

$$\beta_{g^2} = dg^4 + \cdots, \tag{35.49}$$

For completeness, we also give here the β-function at two-loop order:

$$\beta(g^2) = \beta_2 g^4 + \beta_3 g^6 + O\left(g^8\right) \tag{35.43}$$

with

$$8\pi^2 \beta_2 = -\tfrac{1}{3}\left[11C(G) - 4T(R)\right], \tag{35.44}$$

$$\left(8\pi^2\right)^2 \beta_3 = -\tfrac{1}{6}\left[17C^2(G) - 6C(R)T(R) - 10C(G)T(R)\right]. \tag{35.45}$$

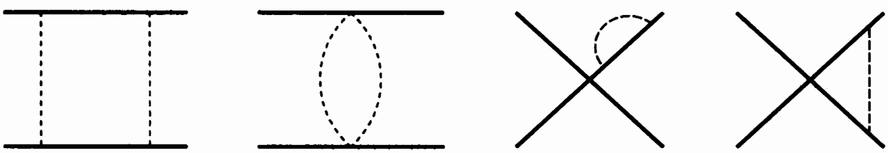
We recall that, as shown in Section 10.11, these two first coefficients are independent of the renormalization scheme. For the $SU(N)$ group, with N_F flavours in the fundamental representation, one finds

$$\left(8\pi^2\right)^2 \beta_3 = -\tfrac{1}{6}\left[17N^2 - 5NN_F - 3N_F(N^2-1)/(2N)\right]. \tag{35.46}$$

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 (18.115), 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 35.2, by itself does not allow 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

$$\beta_u = au^2 + 2bug^2 + cg^4. \tag{35.47}$$

We know that a is positive. It is easy to verify that the contributions to c of the two diagrams (a) of figure 35.1 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 figure 35.1).



(a) (b)

Fig. 35.1 Dashed lines correspond to gauge fields.

From the form (35.46) 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

$$v = u/g^2. \tag{35.48}$$

Setting

$$\beta_{g^2} = dg^4 + \cdots, \tag{35.49}$$

and using equation (10.88), which relates the β-functions in different parametrizations:

$$\beta_j(g)\frac{\partial \tilde{g}_i}{\partial g_j} = \tilde{\beta}_i(\tilde{g}), \tag{35.50}$$

we can calculate

$$\beta_v = \beta_{g^2}\frac{\partial v}{\partial g^2} + \beta_u \frac{\partial v}{\partial u}. \tag{35.51}$$

We obtain

$$\beta_v = g^2\left[av^2 + (2b-d)\,v + c\right]. \tag{35.52}$$

The theory can only be asymptotically free if the second degree polynomial in v has positive zeros. This requires

$$(b-d/2)^2 > ac\,, \qquad d-2b > 0\,. \tag{35.53}$$

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 (35.53) 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 results have been obtained:

(i) It is necessary to render the coefficient d of equation (35.49) 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. There exist, however, theories in which one can find a manifold of measure zero 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.

35.5 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. Presently, the existence of six quarks has been established, corresponding to three generations and six *flavours*. The sixth quark (the top) has been confirmed only recently, but its existence had been generally postulated because flavours are paired in successive generations. Moreover, experiments exclude an additional generation of quarks associated with light leptons (neutrinos of mass below 45 GeV). With six quarks, the theory is still asymptotically free and asymptotic freedom alone would leave 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 12 that information about the behaviour near the light cone can be obtained from RG arguments. It is clear from the discussion of Section 27.2, 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. We do not give here a detailed discussion of the theoretical 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. 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. Since the electromagnetic current is conserved, we can write in momentum space

$$\langle J_\mu(q) J_\nu(-q)\rangle = \left(\delta_{\mu\nu}q^2 - q_\mu q_\nu\right) F(q^2). \tag{35.54}$$

The cross section is proportional to $\operatorname{Im} F(q^2)$ 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 renormalization group and perturbation theory. Strictly speaking, we can only obtain the behaviour of $F(q^2)$ at short distance, that is, in the euclidean region q^2 large and positive. This behaviour is valid in any direction 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 $\exp\left(-\sqrt{q^2}\right)$. We hereafter ignore this difficulty but remember that the behaviour we obtain might only apply to some local average of $F(q^2)$ which smoothes such oscillations.

The electromagnetic current $J_\mu(q)$ requires no renormalization because it is exactly conserved, and has thus no anomalous dimension (see Appendix A13.2). However, the product of two currents which have dimension 3 has dimension 2. Since, due to current conservation, we have been able to extract a factor of dimension 2, $F(q^2)$ is logarithmically divergent. This situation is similar to the ϕ^2 two-point function in the ϕ^4 field theory. The two-point function needs an additive renormalization which induces an inhomogeneous term in the RG equations (see Section 10.7). Calling μ the renormalization scale, and associating a parameter m to the masses of the quarks, we can write the RG equation:

$$\left(\mu\frac{\partial}{\partial\mu} + \beta\frac{\partial}{\partial g} + \eta_m m\frac{\partial}{\partial m}\right) F(q^2, g, m, \mu) = B(g)\,. \tag{35.55}$$

The solution of a similar equation has already been discussed in Section 27.2 (see equations (27.24–A26.10)). Let us call $C(g)$ a particular solution of the inhomogeneous equation:

$$\beta(g)\frac{\partial}{\partial g}C(g) = B(g)\,. \tag{35.56}$$

Then, $F(q^2) - C$ satisfies a homogeneous equation which can be solved in the usual way:

$$F\left((\lambda q)^2, g, m, \mu\right) - C(g) = F\left(q^2, g(\lambda), m(\lambda)/\lambda, \mu\right) - C\left(g(\lambda)\right). \tag{35.57}$$

Since the theory is asymptotically free, the effective coupling constant $g(\lambda)$ goes to zero for large λ. The l.h.s. term $B(g)$ does not vanish for $g = 0$ (even in a free theory $F(q^2)$ is divergent),

$$B(g) = B_0 + B_1 g^2 + O\left(g^4\right), \tag{35.58}$$

and thus $C(g)$ behaves for g small like

$$C(g) = -\frac{B_0}{2\beta_2 g^2} + \left(\frac{B_1}{\beta_2} - \frac{B_0\beta_3}{\beta_2^2}\right)\ln g + O(1). \tag{35.59}$$

Therefore, at large scale λ, the r.h.s. of equation (35.57) is dominated by the singular terms of $C\left(g(\lambda)\right)$:

$$F(\lambda^2 q^2, g, m, \mu) = -C(g(\lambda)) + O(1). \tag{35.60}$$

From the definition of $g(\lambda)$ and equation (35.56), it follows that

$$\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}C\left(g(\lambda)\right) = B\left(g(\lambda)\right), \tag{35.61}$$

and, therefore, using the expansion of $g(\lambda)$ for λ large,

$$C\left(g(\lambda)\right) = B_0 \ln\lambda - \frac{B_1}{2\beta_2}\ln\ln\lambda + O(1). \tag{35.62}$$

The final result is

$$F(q^2) = -\frac{B_0}{2}\ln\left(\frac{q^2}{\mu^2}\right) + \frac{B_1}{2\beta_2}\ln\ln\left(\frac{q^2}{\mu^2}\right) + O(1) \quad \text{for } q^2 \to \infty. \tag{35.63}$$

It follows that

$$\operatorname{Im}\left(F(q^2)\right) = -\frac{B_0}{2}\pi + \frac{B_1}{2\beta_2}\frac{\pi}{\ln(q^2/\mu^2)} + o\left(\frac{1}{\ln(q^2/\mu^2)}\right) \quad \text{for } q^2 \to \infty. \tag{35.64}$$

One usually expresses this result in terms of the ratio $R(q^2)$ of the cross section for $\mathrm{e}_+\,\mathrm{e}_-$ into hadrons to $\mathrm{e}_+\,\mathrm{e}_-$ 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 18.9 (equation (18.105)). 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 which involves the charges Q_i of the quarks. Therefore, for q^2 large and negative, this ratio is just 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:

$$R(q^2) = \sum_i Q_i^2\left[1 - \frac{B_1}{B_0\beta_2}\frac{1}{\ln(q^2/\mu^2)} + o\left(\frac{1}{\ln(q^2/\mu^2)}\right)\right]. \tag{35.65}$$

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:

$$R(q^2) = \sum_i Q_i^2 \left[1 + \frac{1}{4\pi^2} g^2(q/\mu) + O\left(g^4(q/\mu)\right)\right]. \tag{35.66}$$

For the $SU(3)$ colour group and for the six flavours already observed the coefficient of the leading term is

$$\sum_i Q_i^2 = 3 \times \left(3 \times \left(\frac{2}{3}\right)^2 + 3 \times \left(\frac{1}{3}\right)^2\right) = 5. \tag{35.67}$$

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, R to be slowly varying in intermediate regions and close to the value obtained by taking into account only the quarks of smaller masses. For instance, all measurements have been made below the top threshold and, therefore, the largest relevant value is $11/3$.

Bibliographical Notes

For a discussion of the triviality of ϕ^4, see also
D.J.E. Callaway, *Phys. Rep.* 167 (1988) 241.
The upper bound on the Higgs mass is discussed, for example, in
M. Lüscher and P. Weisz, *Nucl. Phys.* B318 (1989) 705.
A early discussion of the applications of renormalization group to Strong Interaction Physics can be found in
D.J. Gross in *Methods in Field Theory*, Les Houches 1975, already quoted in Chapter 11.
The RG β-function for gauge theories has been calculated at one-loop order by
H.D. Politzer, *Phys. Rev. Lett.* 30 (1973) 1346; D.J. Gross and F. Wilczek, *Phys. Rev. Lett.* 30 (1973) 1343;
The general analysis of RG properties of four-dimensional field theories is due to
S. Coleman and D.J. Gross, *Phys. Rev. Lett.* 31 (1973) 851.
See also
S. Coleman in *Properties of the Fundamental Interactions*, Erice 1973, A. Zichichi ed., reprinted in *Aspects of Symmetries* (Cambridge University Press, Cambridge 1985).
Applications of asymptotically free theories to the moments of the deep inelastic structure functions are considered by
H. Georgi and H.D. Politzer, *Phys. Rev.* D9 (1974) 416; D.J. Gross and F. Wilczek, *Phys. Rev.* D9 (1974) 920.
An elementary textbook is
M. Le Bellac, *Des Phénomènes Critiques aux Champs de Jauge* (Editions du CNRS, 1988), english version: *Quantum and Statistical Field Theory* (Oxford Univ. Press, Oxford 1992).
Calculation of the RG β-function to higher orders is due to
W.E. Caswell, *Phys. Rev. Lett.* 33 (1974) 244; D.R.T. Jones, *Nucl. Phys.* B75 (1974) 531; A.A. Belavin and A.A. Migdal, *JETP Lett.* 19 (1974) 181; E.Sh. Egoryan and O.V. Tarasov, *Theor. Math. Phys.* 41 (1979) 863; O.V. Tarasov, A.A. Vladimirov and A.Y. Zharkov, *Phys. Lett.* 93B (1980) 429.

36 CRITICAL DYNAMICS

Up to now, we have discussed only the static equilibrium properties of critical systems. We shall now study the time evolution of classical statistical systems in the critical domain (note that from the point of view of Particle Physics this time is completely unphysical 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.

In contrast to the situation in static critical phenomena, there is, however, 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....

One can argue, however, that the dynamics of all these quantities can, in the critical domain, be described by coupled Langevin equations of the type considered in Chapters 4 and 17. We have already shown in Section 4.4 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

$$\dot{\varphi}_i(t) = -\frac{1}{2}\beta\Omega_{ij}\frac{\delta\mathcal{H}}{\delta\varphi_j} + F_i\bigl(\varphi(t)\bigr) + \nu_i(t), \tag{36.1}$$

in which $\nu_i(t)$ is the usual gaussian white noise:

$$\langle\nu_i(t)\rangle = 0\,, \qquad \langle\nu_i(t)\nu_j(t')\rangle = \Omega_{ij}\delta(t-t') \tag{36.2}$$

leads to the equilibrium distribution $\mathrm{e}^{-\beta\mathcal{H}}$ if the "streaming" term $F_i(\varphi)$ satisfies the conservation equation:

$$\frac{\partial}{\partial\varphi_i}\left[F_i(\varphi)\,\mathrm{e}^{-\beta\mathcal{H}(\varphi)}\right] = 0\,, \tag{36.3}$$

as can be verified immediately on the corresponding Fokker–Planck equation (see, for example, equations (4.28)).

This equation 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 beautiful simplicity of the statics. In specific examples, one writes the simplest phenomenological equation which has the required equilibrium properties and which incorporates all the known physical conditions. A few typical examples are discussed in this chapter.

The many techniques we have developed in Chapter 17 are then useful to predict the RG properties and to perform explicit perturbative calculations. Note that some of these techniques also apply to general local dynamical equations without reference to a possible equilibrium state.

Correlation and response functions. As mentioned above, we can be interested in the critical behaviour of relaxation towards equilibrium, time-dependent correlation functions:

$$W^{(n)}_{i_1\ldots i_n}(t_1,\ldots,t_n) = \langle \varphi_{i_1}(t_1)\ldots\varphi_{i_n}(t_n)\rangle_\nu\,, \tag{36.4}$$

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:

$$\mathcal{H}\bigl(\varphi(t)\bigr) \mapsto \mathcal{H}\bigl(\varphi(t)\bigr) - \int \mathrm{d}t\, h(t)\mathcal{O}\bigl(\varphi(t)\bigr). \tag{36.5}$$

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

$$R^{(n)}_{i_1\ldots i_n}(t_0;t_1,\ldots,t_n) = \frac{\delta}{\delta h(t_0)}\left.\langle\varphi_{i_1}(t_1;h)\ldots\varphi_{i_n}(t_n;h)\rangle\right|_{h=0}. \tag{36.6}$$

The causality of the Langevin equation implies

$$R^{(n)} = 0 \quad \text{for } t_i < t_0\,,\quad \forall\, 1\le i\le n\,. \tag{36.7}$$

If we take h constant, then the modification (36.5) is just a modification of the hamiltonian allowing at equilibrium to generate $\mathcal{O}(\varphi)$ correlation functions. Therefore,

$$\lim_{T\to+\infty}\int \mathrm{d}t\, R^{(n)}_{i_1,\ldots,i_n}(t;T,T,\ldots,T) = \langle\mathcal{O}(\varphi)\varphi_{i_1}\ldots\varphi_{i_n}\rangle\,, \tag{36.8}$$

in which the r.h.s. now is an average taken with the equilibrium distribution $\mathrm{e}^{-\beta\mathcal{H}(\varphi)}$. We can of course calculate higher order derivatives which respect to $h(t)$ and generalize the properties (36.7,36.8).

36.1 Dissipative Case: RG Equations near Four Dimensions

We first discuss a model with a purely dissipative dynamics and without conservation laws. In the classification of the review article of Halperin and Hohenberg, which we follow in this chapter, we consider model A. We have developed in Chapter 17 most of the technology we need to derive RG equations for the dynamics. In Appendix A36, we calculate the dynamic RG functions, both for the ϕ^4 field theory and the non-linear σ-model, at two-loop order.

The N-vector model near four dimensions. We first consider a dissipative dynamics for the N-vector model:

$$\dot\varphi(t,x) = -\frac{\Omega_0}{2}\left[\left(-\nabla^2 + r_0\right)\varphi(t,x) + g_0\frac{\Lambda^\varepsilon}{3!}\varphi\varphi^2\right] + \nu(t,x). \tag{36.9}$$

The equilibrium distribution is characterized by the hamiltonian $\mathcal{H}$:

$$\mathcal{H}(\varphi) = \int \mathrm{d}^d x\left[\frac{1}{2}\left(\partial_\mu\varphi\right)^2 + \frac{1}{2}r_0\varphi^2 + g_0\frac{\Lambda^\varepsilon}{4!}\left(\varphi^2\right)^2\right]. \tag{36.10}$$

In terms of the superfield

$$\boldsymbol{\phi} = \boldsymbol{\varphi} + \theta \bar{\mathbf{c}} + \mathbf{c}\bar{\theta} + \theta\bar{\theta}\boldsymbol{\lambda}, \tag{36.11}$$

the dynamic action $\mathcal{S}(\boldsymbol{\phi})$ takes a supersymmetric form

$$\mathcal{S}(\boldsymbol{\phi}) = \int \mathrm{d}t\, \mathrm{d}\bar{\theta}\, \mathrm{d}\theta \left[\int \mathrm{d}^d x \frac{2}{\Omega_0} \bar{\mathrm{D}}\phi \mathrm{D}\phi + \mathcal{H}(\boldsymbol{\phi})\right] \tag{36.12}$$

with

$$\bar{\mathrm{D}} = \frac{\partial}{\partial \bar{\theta}}, \quad \mathrm{D} = \frac{\partial}{\partial \theta} - \bar{\theta}\frac{\partial}{\partial t}. \tag{36.13}$$

36.1.1 *Supersymmetry and the fluctuation–dissipation theorem*

The addition to the static action of a perturbation proportional to the order parameter itself:

$$\mathcal{H}\big(\boldsymbol{\varphi}(t,x)\big) \mapsto \mathcal{H}\big(\boldsymbol{\varphi}(t,x)\big) - \int \mathrm{d}t\, \mathrm{d}^d x\, \mathbf{h}(t,x) \cdot \boldsymbol{\varphi}(t,x) \tag{36.14}$$

is equivalent to the addition of a term to the noise in equation (36.9):

$$\boldsymbol{\nu}(t,x) \mapsto \boldsymbol{\nu}(t,x) + \tfrac{1}{2}\Omega\, \mathbf{h}(t,x). \tag{36.15}$$

Therefore, differentiation with respect to $\mathbf{h}(t,x)$ generates $\boldsymbol{\lambda}(t,x)$ field correlation functions.

In Section 17.1.2, from supersymmetry, we have derived WT identities satisfied by the connected time-dependent correlation functions (equation (17.9)):

$$\sum_{j=1}^{n} \left(\frac{\partial}{\partial \bar{\theta}_j} + \theta_j \frac{\partial}{\partial t_j}\right) W^{(n)}(k_i, t_i, \theta_i, \bar{\theta}_i) = 0\,. \tag{36.16}$$

We have then obtained the general form of a two-point function consistent with supersymmetry and causality (equation (17.14)):

$$W^{(2)}(k,t,\boldsymbol{\theta},\boldsymbol{\theta}') = \left\{1 + \frac{1}{2}(\theta - \theta')\left[\bar{\theta} + \bar{\theta}' - (\bar{\theta} - \bar{\theta}')\epsilon(t-t')\right]\frac{\partial}{\partial t}\right\} A(k, t-t'),$$

where $\epsilon(t)$ is the sign function and we denote by $\boldsymbol{\theta}$ the set $\{\bar{\theta}, \theta\}$. In our example, $A(t)$ is in addition an even real function of t. In terms of the Fourier transform over time,

$$B(k,\omega) = \int_0^{+\infty} \mathrm{e}^{i\omega t} A(k,t)\mathrm{d}t\,, \tag{36.17}$$

the Fourier transform of $W^{(2)}$ takes the form

$$\begin{aligned} W^{(2)}(k,\omega,\boldsymbol{\theta},\boldsymbol{\theta}') = {} & B(k,\omega) + B^*(k,\omega) - i\omega\,(\theta - \theta')\left[\bar{\theta}' B(k,\omega) + \bar{\theta} B^*(k,\omega)\right] \\ & + A(k,0)\delta^2(\boldsymbol{\theta} - \boldsymbol{\theta}'), \end{aligned} \tag{36.18}$$

in which B^* is the complex conjugate of B.

The sum $B + B^*$ which is the $\langle \varphi\varphi \rangle$ connected two-point function, is a real function. The response function $\langle \lambda\varphi \rangle$ is the coefficient of $\theta'\bar{\theta}'$. Taking its imaginary part, we find the relation

$$\begin{aligned} \mathrm{Im}\left[\frac{\delta}{\delta h(k,\omega)} \langle \varphi(-k,-\omega)\rangle\right]\bigg|_{h=0} & \equiv \langle \lambda(k,\omega)\varphi - k, -\omega)\rangle \\ & = \tfrac{1}{2}\omega \left\langle \varphi(k,\omega)\varphi(-k,-\omega)\right\rangle, \end{aligned} \tag{36.19}$$

known under the name of fluctuation–dissipation theorem.

36.1.2 RG equations at and above T_c

We have shown in Section 17.1.3 that static and supersymmetric dynamic theories have the same upper-critical dimension. Fluctuations are, therefore, only relevant for dimensions $d \le 4$. From the discussion of Chapter 17, we know that the renormalized action $\mathcal{S}_{\mathrm{r}}(\boldsymbol{\phi})$ then has the form

$$\mathcal{S}_{\mathrm{r}}(\boldsymbol{\phi}) = \int \mathrm{d}\bar\theta\,\mathrm{d}\theta\,\mathrm{d}t \left[\int \mathrm{d}^d x \frac{2}{\Omega} Z_\omega \bar{\mathrm{D}}\phi \mathrm{D}\phi + \mathcal{H}_{\mathrm{r}}(\boldsymbol{\phi})\right], \tag{36.20}$$

in which $\boldsymbol{\phi}$ is now the renormalized field and $\mathcal{H}_{\mathrm{r}}(\boldsymbol{\phi})$ is the static renormalized hamiltonian.

To renormalize the action (36.12), we have introduced, in addition to the static renormalization constants, a renormalization of the parameter Ω:

$$\Omega_0 = \Omega Z/Z_\omega\,, \tag{36.21}$$

where Z is the field renormalization constant. The RG differential operator then takes the form

$$\mathrm{D}_{\mathrm{RG}} = \mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} + \eta_\omega(g)\Omega\frac{\partial}{\partial\Omega} - \frac{n}{2}\eta(g), \tag{36.22}$$

where $\eta_\omega(g)$ is a new independent RG function, defined as

$$\eta_\omega(g) = \mu\frac{\mathrm{d}}{\mathrm{d}\mu}\bigg|_{g_0,\Omega_0} \ln\Omega\,. \tag{36.23}$$

In the case of dimensional regularization, equation (36.23) becomes

$$\eta_\omega(g) = \beta(g)\frac{\mathrm{d}}{\mathrm{d}g}\ln(Z_\omega/Z). \tag{36.24}$$

The RG equations for the critical theory then read

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} + \eta_\omega(g)\Omega\frac{\partial}{\partial\Omega} - \frac{n}{2}\eta(g)\right]\Gamma^{(n)}(p_i,\omega_i,\boldsymbol{\theta}_i,\mu,\Omega,g) = 0\,, \tag{36.25}$$

in which μ is the renormalization scale, and we have used the vector notation $\boldsymbol{\theta} \equiv (\bar\theta,\theta)$.

At the IR fixed point g^*, they reduce to

$$\left(\mu\frac{\partial}{\partial\mu} + \eta_\omega\Omega\frac{\partial}{\partial\Omega} - \frac{n}{2}\eta\right)\Gamma^{(n)}(p_i,\omega_i,\boldsymbol{\theta}_i,\mu,\Omega) = 0 \tag{36.26}$$

with

$$\eta_\omega = \eta_\omega(g^*).$$

We set

$$z = 2 + \eta_\omega\,. \tag{36.27}$$

From dimensional analysis, we obtain

$$\Gamma^{(n)}\left(\lambda p_i, \rho\omega_i, \frac{\boldsymbol{\theta}_i}{\sqrt{\rho}}, \lambda\mu, \frac{\rho\Omega}{\lambda^2}\right) = \lambda^{d-n(d-2)/2}\rho^{1-n}\Gamma^{(n)}(p_i,\omega_i,\boldsymbol{\theta}_i,\mu,\Omega), \tag{36.28}$$

In the dimensional equation (36.28), we choose

$$\rho = \Omega \lambda^z \mu^{-\eta_\omega} . \tag{36.29}$$

Then, combining the solution of the RG equation with equation (36.28), we find the dynamic scaling relations

$$\Gamma^{(n)}(\lambda p_i, \omega_i, \boldsymbol{\theta}_i, \mu = 1, \Omega = 1) = \lambda^{d-n(d-2+\eta)/2-z(n-1)} F^{(n)}(p_i, \lambda^{-z}\omega_i, \boldsymbol{\theta}_i \lambda^{z/2}). \tag{36.30}$$

A few algebraic manipulations yield the corresponding relation for connected correlation functions:

$$W^{(n)}(\lambda p_i, \omega_i, \boldsymbol{\theta}_i, \mu = 1, \Omega = 1) = \lambda^{(d+z)(1-n)+n(d-2+\eta)/2} G^{(n)}(p_i, \omega_i \lambda^{-z}, \boldsymbol{\theta}_i \lambda^{z/2}). \tag{36.31}$$

The φ-field two-point correlation function is obtained for $n = 2$ and $\boldsymbol{\theta} = 0$:

$$W^{(2)}(p, \omega, \boldsymbol{\theta} = 0) \sim p^{-2+\eta-z} G^{(2)}(\omega / p^z). \tag{36.32}$$

The equal-time correlation function is obtained by integrating over ω. One verifies the consistency with static scaling.

The dynamic critical two-point function thus depends on a frequency scale which vanishes at small momentum like p^z or a time scale which diverges like p^{-z}. The RG function η_ω for this model is calculated in Section A36.1 (see equation (A36.13)):

$$\eta_\omega(\tilde{g}) = \frac{N+2}{72} \left[6 \ln(4/3) - 1\right] \tilde{g}^2 + O\left(\tilde{g}^3\right) \tag{36.33}$$

with as usual

$$\tilde{g} = N_d \, g \,, \quad N_d = \frac{2}{\Gamma(d/2)(4\pi)^{d/2}} \,. \tag{36.34}$$

The dynamic critical exponent z follows:

$$z = 2 + \frac{N+2}{2(N+8)^2} \left[6 \ln(4/3) - 1\right] \varepsilon^2 + 0\left(\varepsilon^3\right). \tag{36.35}$$

At this order, it can also be written as

$$z = 2 + c\eta + O\left(\varepsilon^3\right) \tag{36.36}$$

with

$$c = 6 \ln(4/3) - 1 \,. \tag{36.37}$$

The scaling behaviour of the response function can be obtained by considering, for example, the coefficient of $\theta\bar{\theta}$ in $W^{(2)}$.

Correlation functions above T_c in the critical domain. Let us just write the RG equations at the IR fixed point:

$$\left(\mu \frac{\partial}{\partial \mu} + \eta_\omega \Omega \frac{\partial}{\partial \Omega} - \eta_2 \sigma \frac{\partial}{\partial \sigma} - \frac{n}{2}\eta\right) \Gamma^{(n)}(p_i, \omega_i, \boldsymbol{\theta}_i, \sigma, \mu, \Omega) = 0 \,, \tag{36.38}$$

in which σ is a measure of the deviation from the critical temperature:

$$\sigma \sim T - T_c\,. \tag{36.39}$$

The dimensional relation (36.28) becomes

$$\Gamma^{(n)}(p_i,\omega_i,\boldsymbol{\theta}_i,\sigma,\mu,\Omega) = \lambda^{d-n(d-2)/2}\rho^{1-n}\Gamma^{(n)}\left(\frac{p_i}{\lambda},\frac{\omega_i}{\rho},\boldsymbol{\theta}_i\sqrt{\rho},\frac{\sigma}{\lambda^2},\frac{\mu}{\lambda},\frac{\Omega\lambda^2}{\rho}\right). \tag{36.40}$$

Combining this equation with the RG equation (36.38) and choosing

$$\lambda = \sigma^{\nu}\mu^{\nu\eta_2} \sim \xi^{-1}, \qquad \rho = \Omega\mu^{-\eta\omega}\lambda^{z} \sim \xi^{-z}, \tag{36.41}$$

in which ξ is the correlation length, we finally obtain

$$\begin{aligned}\Gamma^{(n)}(p_i,\omega_i,\boldsymbol{\theta}_i,\sigma,\mu=1,\Omega=1) \sim\ & \xi^{-d+n(d-2+\eta)/2+z(n-1)} \\ & \times F^{(n)}(p_i\xi,\omega_i\xi^{z},\boldsymbol{\theta}_i\xi^{-z/2}).\end{aligned} \tag{36.42}$$

We now see that all times are measured in terms of a *correlation time* τ which diverges at the critical temperature as ξ^{z}:

$$\tau \propto \xi^{z}. \tag{36.43}$$

36.2 Dissipative Case: RG Equations Near Two Dimensions

Near two dimensions, the N-vector model is described by a Langevin equation of the form (see Section 17.3, Appendix A36.2)

$$\dot{\boldsymbol{\varphi}} = -\tfrac{1}{2}\Omega\beta\left[-\nabla^2\boldsymbol{\varphi} + \boldsymbol{\varphi}\left(\boldsymbol{\varphi}\cdot\nabla^2\boldsymbol{\varphi}\right)\right] + \boldsymbol{\nu} - \boldsymbol{\varphi}\left(\boldsymbol{\varphi}\cdot\boldsymbol{\nu}\right), \tag{36.44}$$

$\boldsymbol{\nu}(t,x)$ being the noise of Section 36.1 and $\boldsymbol{\varphi}$ satisfying

$$\boldsymbol{\varphi}^2 = 1\,. \tag{36.45}$$

Setting

$$\beta = g_0/\Lambda^{d-2}, \tag{36.46}$$

we can derive from equation (36.44) a dynamic action

$$\mathcal{S}(\boldsymbol{\phi}) = \frac{\Lambda^{d-2}}{g_0}\int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t\left[\int \mathrm{d}^d x\frac{2}{\Omega_0}\bar{\mathrm{D}}\boldsymbol{\phi}\mathrm{D}\boldsymbol{\phi} + \mathcal{H}(\boldsymbol{\phi})\right] \tag{36.47}$$

with

$$\mathcal{H}(\boldsymbol{\phi}) = \frac{1}{2}\int \mathrm{d}^d x\left[\partial_\mu\boldsymbol{\phi}(t,x)\right]^2. \tag{36.48}$$

The superfield $\boldsymbol{\phi}$ satisfies the constraint

$$\boldsymbol{\phi}^2 = 1\,. \tag{36.49}$$

In the superfield form (36.47), we see that the dynamical actions for $(\boldsymbol{\phi}^2)^2$ field theory and the non-linear σ-model are related in the same way as the corresponding static hamiltonians.

The renormalized action $\mathcal{S}_\mathrm{r}(\phi)$ then reads

$$\mathcal{S}_\mathrm{r}(\phi) = \frac{\mu^\varepsilon}{g} \int \mathrm{d}\bar\theta\,\mathrm{d}\theta\,\mathrm{d}t \left[\int \mathrm{d}^d x \frac{2}{\Omega} Z_\omega \bar{\mathrm{D}}\phi \mathrm{D}\phi + \mathcal{H}_\mathrm{r}(\phi)\right], \tag{36.50}$$

in which μ is the renormalization scale and $d = 2 + \varepsilon$.

In the minimal subtraction scheme, the new RG function $\eta_\omega(g)$ is given by

$$\eta_\omega(g) = \beta(g) \frac{\mathrm{d}}{\mathrm{d}g} \ln(Z_\omega Z_g / Z). \tag{36.51}$$

The RG equations in zero magnetic field then read (see equation (31.19))

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} + \eta_\omega(g)\Omega\frac{\partial}{\partial\Omega} - \frac{n}{2}\zeta(g)\right] \Gamma^{(n)}(p_i, \omega_i, \boldsymbol{\theta}_i, g, \mu, \Omega) = 0\,. \tag{36.52}$$

Dimensional analysis yields (equation (36.28))

$$\Gamma^{(n)}(p_i, \omega_i, \boldsymbol{\theta}_i; g, \mu, \Omega) = \lambda^d \rho^{1-n} \Gamma^{(n)}\left(\frac{p_i}{\lambda}, \frac{\omega_i}{\rho}, \boldsymbol{\theta}_i\sqrt{\rho}; \frac{\Omega\lambda^2}{\rho}, \frac{\mu}{\lambda}\right). \tag{36.53}$$

We introduce in addition to the functions $M_0(g)$ and $\xi(g)$:

$$M_0(g) = \exp\left[-\frac{1}{2}\int_0^g \frac{\zeta(g')}{\beta(g')}\mathrm{d}g'\right], \tag{36.54}$$

$$\xi(g) = \mu^{-1} g^{1/\varepsilon} \exp\left[\int_0^g \left(\frac{1}{\beta(g')} - \frac{1}{\varepsilon g'}\right)\mathrm{d}g'\right], \tag{36.55}$$

(see equations (31.33,31.34)) and the new RG function

$$\tau(g) = \Omega^{-1}\xi^2(g) \exp\left[\int_0^g \frac{\eta_\omega(g')}{\beta(g')}\mathrm{d}g'\right]. \tag{36.56}$$

Combining equations (36.52–36.56) and choosing

$$\lambda = \xi^{-1}(g), \qquad \rho = \tau^{-1}(g), \tag{36.57}$$

we obtain

$$\Gamma^{(n)}(p_i, \omega_i, \boldsymbol{\theta}_i; g, \mu = 1, \Omega = 1) = \tau^{n-1}\xi^{-d} M_0^{-n} F^{(n)}(p_i\xi, \omega\tau, \boldsymbol{\theta}\tau^{-1/2}).$$

Near the critical temperature g^*, this expression agrees with the scaling form (36.42). The exponent z then is given by

$$z = 2 + \eta_\omega(g^*). \tag{36.58}$$

Equation (A36.32) gives $\eta_\omega(g)$ at two-loop order. The exponent z is then

$$z = 2 + \left(1 - \ln(4/3)\right)\frac{\varepsilon^2}{N-2} + O\left(\varepsilon^3\right), \tag{36.59}$$

or again in terms of η:

$$z = 2 + \left(1 - \ln(4/3)\right)\varepsilon\eta + O\left(\varepsilon^3\right). \tag{36.60}$$

36.3 Conserved Order Parameter

A simple modification of equation (36.9) ensures that the order parameter is conserved, which is the statement

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^d x\, \varphi(t,x) = 0\,. \tag{36.61}$$

We consider the equation

$$\dot{\varphi}(t,x) = \frac{\Omega_0}{2}\nabla_x^2 \frac{\delta \mathcal{H}}{\delta \varphi(t,x)} + \nu(t,x), \tag{36.62}$$

where the gaussian noise distribution satisfies

$$\langle \nu_i(t,x)\nu_j(t',x')\rangle = -\delta_{ij}\Omega_0\, \delta(t-t')\nabla_x^2\delta(x-x'). \tag{36.63}$$

The dynamic action $\mathcal{S}(\boldsymbol{\phi})$ is still supersymmetric:

$$\mathcal{S}(\boldsymbol{\phi}) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t \left[-\int \mathrm{d}^d x \frac{2}{\Omega_0}\bar{\mathrm{D}}\boldsymbol{\phi}\cdot\nabla^{-2}\mathrm{D}\boldsymbol{\phi} + \mathcal{H}(\boldsymbol{\phi})\right], \tag{36.64}$$

and this form shows that this Langevin equation generates the same equilibrium distribution as model A.

The appearance of a non-local term in the action reminds us of the effective field theory for uniaxial systems with dipolar forces. Power counting now is different. Since the propagator reads

$$\Delta(k,\omega,\boldsymbol{\theta},\boldsymbol{\theta}') = \frac{\Omega k^2\left[1-\frac{1}{2}i\omega(\theta-\theta')(\bar{\theta}+\bar{\theta}') + \frac{1}{4}\Omega k^2(k^2+r)\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}')\right]}{\omega^2+\frac{1}{4}\Omega^2\left(k^2\right)^2\left(k^2+r\right)^2}, \tag{36.65}$$

ω has the dimension of k^4. Above four dimensions, the characteristic frequency diverges like k^4 which means that the dynamical exponent z is 4 instead of 2 as in model A.

Above two dimensions, Feynman diagrams calculated with the propagator (36.65) are not singular at zero momentum for $\omega \neq 0$. Therefore, no counter-term singular in $\mathbf{k}$ can be generated and thus Ω_0 in the dynamic action remains unrenormalized. The renormalized action $\mathcal{S}_\mathrm{r}(\boldsymbol{\phi})$ is

$$\mathcal{S}_\mathrm{r}(\boldsymbol{\phi}) = \int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t \left[-\int \mathrm{d}^d x \frac{2}{\Omega}\bar{\mathrm{D}}\boldsymbol{\phi}\cdot\nabla^{-2}\mathrm{D}\boldsymbol{\phi} + \mathcal{H}_\mathrm{r}(\boldsymbol{\phi})\right]. \tag{36.66}$$

The field amplitude renormalization implies the relation

$$\Omega_0 = Z\Omega\,, \tag{36.67}$$

and, therefore, the RG function $\eta_\omega(g)$ is

$$\eta_\omega(g) = -\beta(g)\frac{\mathrm{d}}{\mathrm{d}g}\ln Z = -\eta(g). \tag{36.68}$$

The analysis then becomes quite similar to the previous case. However, relation (36.68) implies that the exponent z is no longer an independent exponent but given instead by

$$z = 4-\eta\,. \tag{36.69}$$

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. We, therefore, couple a field $e(x)$ to $\varphi^2(x)$ and consider the equilibrium hamiltonian

$$\mathcal{H}(\varphi,e)=\int \mathrm{d}^d x\left[\frac{1}{2}\left(\partial_\mu\varphi\right)^2+\frac{1}{2}r\varphi^2+\frac{u}{4!}\left(\varphi^2\right)^2+\frac{1}{2}\Lambda^{\varepsilon/2}ve(x)\varphi^2(x)+\frac{1}{2}e^2(x)\right]. \tag{36.70}$$

The static properties are not affected by this modification because after integration over the e-field, we recover the usual φ^4 theory. Let us call $K(x)$ the source for the correlation functions of $e(x)$. If we integrate over $e(x)$, we obtain a new hamiltonian

$$\mathcal{H}(\varphi,K)=\int \mathrm{d}^d x\left[\frac{1}{2}\left(\partial_\mu\varphi\right)^2+\frac{1}{2}r\varphi^2+\frac{1}{4!}\left(u-3\Lambda^{\varepsilon}v^2\right)\left(\varphi^2\right)^2+\Lambda^{\varepsilon/2}\frac{v}{2}K\varphi^2-\frac{K^2}{2}\right]. \tag{36.71}$$

In the static theory, the e-field is directly related to the φ^2 field which is also the standard energy operator. The e-field amplitude and the v coupling constant renormalizations can be expressed in terms of φ^2 multiplicative and $\left\langle\varphi^2\varphi^2\right\rangle$ additive renormalization constants. More precisely, we have the relation

$$\langle e(k)e(-k)\rangle_{\mathrm{c}}=1+\Lambda^{\varepsilon}v^2\left\langle\left(\tfrac{1}{2}\varphi^2\right)(k)\left(\tfrac{1}{2}\varphi^2\right)(-k)\right\rangle_{\mathrm{c}}, \tag{36.72}$$

and 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 correlation functions, we can derive the RG equations for $e(x)$ insertions. At T_c they take the form

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)\frac{\partial}{\partial g}+\beta_v(g,v)\frac{\partial}{\partial v}-\frac{l}{2}\eta_e(g,v)-\frac{n}{2}\eta(g)\right]\Gamma_e^{(l,n)}(q,p;g,v,\Lambda)=0\,, \tag{36.73}$$

in which $\Gamma_e^{(l,n)}$ is the correlation function with l $e(x)$ insertions.

Applying this equation to the case $l+n>2$ and using equation (25.72), we obtain the relation

$$\eta_e(g,v)=\varepsilon+2\beta_v/v+2\eta_2(g). \tag{36.74}$$

Then applying equation (36.73) to relation (36.72) and again using equation (25.72), we find

$$\varepsilon+2\beta_v/v+2\eta_2(g)=v^2B(g), \tag{36.75}$$

and, therefore,

$$\eta_e=v^2B(g), \tag{36.76a}$$

$$\beta_v=-\tfrac{1}{2}v\left[\varepsilon+2\eta_2-v^2B(g)\right]. \tag{36.76b}$$

At the IR fixed point g^*, the function β_v becomes

$$\beta_v=-\tfrac{1}{2}v\left[\alpha/\nu-v^2B(g^*)\right]. \tag{36.77}$$

At leading order in the ε-expansion $B(g^*)$ is positive:

$$B(g^*)=\frac{N}{16\pi^2}+O(\varepsilon)\,.$$

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.

(ii) $\alpha > 0$:

The IR fixed point is non-trivial, it corresponds to

$$v = v^* \equiv \pm \left(\frac{\alpha}{\nu B(g^*)}\right)^{1/2} . \tag{36.78}$$

At leading order α vanishes for $N = 4$:

$$\alpha = \frac{4 - N}{2(N+8)}\varepsilon + O\left(\varepsilon^2\right). \tag{36.79}$$

In three dimensions, numerical calculations (see Chapter 29) show that α is already slightly negative for $N = 2$. For $\alpha > 0$, the dynamics differs from the simple dynamics of Section 36.1 as we shall see below.

Of course 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, therefore, examine the dynamics of the model in the case $\alpha > 0$. The Langevin equations consistent with energy conservation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^d x\, e(t,x) = 0$$

are

$$\dot{\varphi} = -\frac{\Omega}{2}\frac{\delta\mathcal{H}}{\delta\varphi(t,x)} + \nu(t,x), \tag{36.80}$$

$$\dot{e} = \frac{\Omega'}{2}\nabla^2\frac{\delta\mathcal{H}}{\delta e(t,x)} + \nu'(t,x). \tag{36.81}$$

The new gaussian noise $\nu'(t,x)$ is defined by

$$\langle\nu'\rangle = 0\,, \qquad \langle\nu'(t,x)\nu'(x',t')\rangle = -\Omega'\delta(t-t')\nabla_x^2\delta(x-x'). \tag{36.82}$$

The dynamic action now written in terms of two superfields $\phi(t,x)$ and $E(t,x)$ is still supersymmetric and reads

$$\mathcal{S}(\phi,E) = \int \mathrm{d}\bar\theta\,\mathrm{d}\theta\,\mathrm{d}t\left\{\int \mathrm{d}^d x\left[-\frac{2}{\Omega'}\bar{\mathrm{D}}E\,\nabla_x^{-2}\,\mathrm{D}E + \frac{2}{\Omega}\bar{\mathrm{D}}\phi\mathrm{D}\phi\right] + \mathcal{H}(\phi,E)\right\}. \tag{36.83}$$

As we have already explained in the example of the φ-field conservation, 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

$$\Omega = \Omega_\mathrm{r}/Z_\omega\,, \tag{36.84}$$

at one-loop order we get

$$Z_\omega = 1 - \frac{\Omega}{\Omega + \Omega'} N_d v^2 \ln \Lambda \,, \tag{36.85}$$

while for Ω' we have

$$\Omega' = Z_e \Omega'_{\mathrm{r}} \,, \tag{36.86}$$

in which Z_e is the e-field renormalization. The corresponding RG function η_e is given by equation (36.76*a*).

The RG equations then read

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \beta_v \frac{\partial}{\partial v} + \eta_\omega \Omega \frac{\partial}{\partial \Omega} + \eta'_\omega \Omega' \frac{\partial}{\partial \Omega'} - \frac{l}{2}\eta_e - \frac{n}{2}\eta\right] \times \Gamma_e^{(l,n)}(q,p;g,v,\Lambda,\Omega,\Omega') = 0 \,. \tag{36.87}$$

Equation (36.86) leads to the identity

$$\eta'_\omega = \eta_e \,, \tag{36.88}$$

while at leading order

$$\eta_\omega = \frac{1}{8\pi^2} \frac{\Omega}{\Omega + \Omega'} v^2 . \tag{36.89}$$

To separate variables, we set

$$s = \Omega'/\Omega \,, \tag{36.90}$$

and take, for example, s and Ω' as independent variables. We then have

$$\eta_\omega \Omega \frac{\partial}{\partial \Omega} + \eta_e \Omega' \frac{\partial}{\partial \Omega'} = \eta_e \Omega' \frac{\partial}{\partial \Omega'} + \beta_s \frac{\partial}{\partial s} \tag{36.91}$$

with

$$\beta_s = \eta_e - \eta_\omega \,. \tag{36.92}$$

At the IR fixed in the (g,v) plane, the function β_s takes the form

$$\beta_s = \frac{\alpha}{\nu}\left(1 - \frac{2}{N}\frac{1}{1+s}\right) s + O\left(\varepsilon^2\right). \tag{36.93}$$

The fixed points in the variables s are

$$s^* = 0 \,, \qquad s^* = 2/N - 1 \,, \qquad s = \infty \,. \tag{36.94}$$

Since α is positive, the fixed point $s = 0$ is stable for $N > 2 + O(\varepsilon)$. However, the $s = 0$ limit is a peculiar limit and it is not clear whether this result is consistent with the ε-expansion.

The fixed point $s = \infty$ is never stable for $\alpha > 0$. It also corresponds to decoupling of the E sector.

Finally, for $N < 2$, $s^* = 2/N - 1$ 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

$$s^* = 2/N - 1 + O(\varepsilon) \Rightarrow z = 2 + \alpha/\nu \,. \tag{36.95}$$

36.5 A Non-Relaxational Model

General remarks. Let us consider the Langevin equation (36.1) (we have set $\beta = 1$):

$$\dot{\varphi}(t) = -\frac{1}{2}\Omega_{ij}\frac{\delta\mathcal{H}}{\delta\varphi_j(t)} + F_i(t) + \nu_i(t) \tag{36.96}$$

with a gaussian noise (36.2):

$$\langle \nu_i(t)\nu_j(t')\rangle = \Omega_{ij}\,\delta(t-t').$$

If the function F_i satisfies equation (36.3), that is,

$$\frac{\partial}{\partial\varphi_i}\left[F_i(\varphi)\,\mathrm{e}^{-\mathcal{H}(\varphi)}\right] = 0\,,$$

the equilibrium distribution is $\mathrm{e}^{-\mathcal{H}}$. A particular solution of this equation is provided by

$$F_i(\varphi) = \frac{\partial}{\partial\varphi_j}R_{ij}(\varphi) - R_{ij}(\varphi)\frac{\partial\mathcal{H}}{\partial\varphi_j}\,, \tag{36.97}$$

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 an 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 previous section there is a conserved density $e(x)$. The Langevin equation, however, reads

$$\dot{\varphi} = -\frac{\Omega}{2}\frac{\delta\mathcal{H}}{\delta\varphi^*(t,x)} - is\varphi\frac{\delta\mathcal{H}}{\delta e(t,x)} + \nu(t,x), \tag{36.98}$$

$$\dot{e} = \frac{\Omega'}{2}\nabla^2\frac{\delta\mathcal{H}}{\delta e(t,x)} + is\left[\varphi^*\frac{\delta\mathcal{H}}{\delta\varphi^*(t,x)} - \varphi\frac{\delta\mathcal{H}}{\delta\varphi(t,x)}\right] + \nu'(t,x). \tag{36.99}$$

The hamiltonian is

$$\mathcal{H}(\varphi,e) = \int \mathrm{d}^d x\left[\left|\partial_\mu\varphi\right|^2 + r\left|\varphi\right|^2 + \frac{1}{3!}u\left|\varphi\right|^4 + \frac{1}{2}e^2\right]. \tag{36.100}$$

The noise two-point functions are the same as in Section 36.4 (for the $N=2$ case):

$$\begin{aligned}\langle\nu(t,x)\nu^*(t',x')\rangle &= \Omega\,\delta(t-t')\delta(x-x'),\\ \langle\nu\nu\rangle = \langle\nu^*\nu^*\rangle &= 0\,,\\ \langle\nu'(t,x)\nu'(t',x')\rangle &= -\Omega'\delta(t-t')\nabla_x^2\delta(x-x').\end{aligned} \tag{36.101}$$

It is easy to verify that this model provides one example of Langevin equation (36.96) with a streaming term of the form (36.97) 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:

$$\int \mathrm{d}^d x \left(\varphi^* \frac{\delta \mathcal{H}}{\delta \varphi^*(t,x)} - \varphi \frac{\delta \mathcal{H}}{\delta \varphi(t,x)}\right) = 0 ,$$

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 theory is renormalizable in four dimensions and that the canonical dimensions of the fields are

$$[\varphi] = 1 , \quad [e] = 2 .$$

From the general results of Section 16.9, based on the BRS 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 $\partial^2(\varphi^*\varphi)$ in equation (36.99) and do not imply the equality of the coupling constants s in equations (36.98) and (36.99). We here 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 (36.98) is unchanged while an additional field independent term is added to the r.h.s. of equation (36.99): $\ddot{\alpha}/s$. 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 easy 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, and

$$v = s^2/\Omega\Omega' , \quad w = \Omega/\Omega' . \tag{36.102}$$

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 $4-\varepsilon$ dimensions as

$$\beta_v = -v\,(\varepsilon + \eta_\omega + \eta_{\omega'}) , \tag{36.103}$$

$$\beta_w = w\,(\eta_\omega - \eta_{\omega'}) . \tag{36.104}$$

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 (36.33)) 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 (36.103) implies

$$\varepsilon + \eta_\omega + \eta_{\omega'} = 0\,. \tag{36.105}$$

Equation (36.104) 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:

$$\eta_\omega = -\frac{v}{8\pi^2(1+w)}\,,\qquad \eta_{\omega'} = -\frac{v}{16\pi^2}\,. \tag{36.106}$$

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 (36.105) holds. Therefore, to all orders

$$\eta_\omega = \eta_{\omega'} = -\varepsilon/2\,. \tag{36.107}$$

The exponent z is then exactly calculable:

$$z = d/2\,. \tag{36.108}$$

For a more detailed discussion, we refer the interested reader to the literature.

Bibliographical Notes

The classification of the various dynamic models is taken from the review article

P.C. Hohenberg and B.I. Halperin, *Rev. Mod. Phys.* 49 (1977) 435.

The dynamic action associated with the Langevin equation has been introduced in

P.C. Martin, E.D. Siggia and H.A. Rose, *Phys. Rev.* A8 (1978) 423.

For an early discussion of the renormalization of dynamic theories in a field theory language see, for example,

C. De Dominicis and L. Peliti, *Phys. Rev.* B18 (1978) 353,

and references therein.

The non-linear σ-model has been considered in

R. Bausch, H.K. Janssen and Y. Yamazaki, *Z. Phys.* B37 (1980) 163.

For a review on early work on the critical dynamics of Helium, see

P.C. Hohenberg, *Physica* 109 & 110B (1982) 1436.

A review concerning Helium with many references is

V. Dohm, *J. Low Temp. Phys.* 69 (1987) 51.

The exponent z for model A is given at order ε^3 in

N.V. Antonov and A.N. Vasil'ev, *Theor. Math. Phys.* 60 (1984) 671.

The dynamics of various other models has also been discussed, like the interface model of Wallace and Zia:

R. Bausch, V. Dohm, H.K. Janssen and R.K.P. Zia, *Phys. Rev. Lett.* 47 (1981) 1837; J.C. Niel, *Phys. Rev.* B37 (1988) 2156,

and the Yang–Lee edge singularity

N. Breuer and H.K. Janssen, *Z. Phys.* B41 (1981) 55.

Finally, the critical behaviour in driven diffusive systems has been studied in

H.K. Janssen and B. Schmittmann, *Z. Phys.* B64 (1986) 503; K. Leung and J.L. Cardy, *J. Stat. Phys.* 44 (1986) 567.

APPENDIX A36
RG FUNCTIONS AT TWO-LOOPS: SUPERSYMMETRIC PERTURBATION THEORY

For illustration purpose, we here consider the examples of the $O(N)$ symmetric $(\phi^2)^2$ field theory and non-linear σ-model of Sections 36.1 and 36.2 and calculate the corresponding dynamic RG functions at two-loop order. From the form of the renormalized dynamical action RG equations can be derived. In the case of the purely dissipative Langevin equation (16.98,16.104), only one new RG function appears, associated with the time scale renormalization.

Perturbative calculations. In perturbative calculations of dynamic quantities, the formalism of Section 16.8.1 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 on the same footing as the usual commuting coordinates. In this way, the 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 $-\infty$ so that the system is at equilibrium at any finite time, and, therefore, time translation invariance is secured.

WT identities. Proper vertices, after Fourier transformation, satisfy the WT identities (17.9,17.10) corresponding to supersymmetry transformations (17.7),

$$\sum_{i=1}^{n}\left(\frac{\partial}{\partial\bar{\theta}_i}-i\omega_i\theta_i\right)\Gamma^{(n)}=0\,. \tag{A36.1}$$

These identities provide checks in perturbative calculations.

A36.1 The $(\phi^2)^2$ Field Theory: Dynamic Exponent

To the action (36.12) corresponds a propagator

$$\Delta(\mathbf{k},\omega,\boldsymbol{\theta},\boldsymbol{\theta}')=\frac{\Omega\left[1-\frac{1}{2}i\omega(\theta-\theta')(\bar{\theta}+\bar{\theta}')+\frac{1}{4}\Omega\left(k^2+m^2\right)\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}')\right]}{\omega^2+\frac{1}{4}\Omega^2\left(k^2+m^2\right)^2}. \tag{A36.2}$$

We have omitted the factor δ_{ij} corresponding to group indices. For practical calculations, it is actually more convenient to use a mixed representation for the propagator, Fourier transformed on space but not in time:

$$\begin{aligned}\Delta(\mathbf{k},t,\boldsymbol{\theta},\boldsymbol{\theta}')=&\left\{\frac{1}{k^2+m^2}+\tfrac{1}{4}\Omega(\theta-\theta')\left[\bar{\theta}-\bar{\theta}'-(\bar{\theta}+\bar{\theta}')\epsilon(t)\right]\right\}\\&\times\exp\left[-\tfrac{1}{2}\Omega\left(k^2+m^2\right)|t|\right],\end{aligned} \tag{A36.3}$$

in which $\epsilon(t)$ is the sign of t. Note that, in agreement with the analysis of Section 17.1.2, Δ can also be written as

$$\begin{aligned}\Delta(\mathbf{k},t,\boldsymbol{\theta},\boldsymbol{\theta}')=&\left\{1+\tfrac{1}{2}(\theta-\theta')\left[\bar{\theta}+\bar{\theta}'-(\bar{\theta}-\bar{\theta}')\epsilon(t)\right]\frac{\partial}{\partial t}\right\}\\&\times\frac{1}{k^2+m^2}\exp\left[-\tfrac{1}{2}\Omega\left(k^2+m^2\right)|t|\right].\end{aligned}$$

This form follows directly from the supersymmetry and the causality of the Langevin equation (see also equations (36.16–36.19)).

From the supersymmetric form (36.12) of the dynamical action, we deduce 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

$$V^{(4)} = m^{\varepsilon}\frac{g}{4!}\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_3)\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_4). \tag{A36.4}$$

We have already calculated all static renormalization constants to two-loop order in Sections 11.5,11.6. To calculate the new renormalization constant at leading order (i.e. two-loop order) we need only the two-point function. Using the expressions of Section 11.6, we find

$$\begin{aligned}\Gamma^{(2)} &= -\frac{2}{\Omega}\left[2-i\omega(\theta-\theta')(\bar\theta+\bar\theta')\right]+\left(k^2+m^2\right)\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}')\\ &\quad+\frac{N+2}{6}m^{\varepsilon}gD_1-\frac{(N+2)^2}{36}m^{2\varepsilon}g^2D_2-\frac{N+2}{18}m^{2\varepsilon}g^2D_3\,,\end{aligned} \tag{A36.5}$$

in which the three diagrams D_1, D_2 ,D_3 are given by

$$D_1=\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}')\frac{1}{(2\pi)^d}\int\frac{\mathrm{d}^dp}{p^2+m^2}, \tag{A36.6}$$

$$\begin{aligned}D_2 &= D_1\int\frac{\mathrm{d}^dp}{(2\pi)^d}\mathrm{d}t\,\mathrm{d}^2\bar\theta''\,\mathrm{e}^{-\Omega(p^2+m^2)|t|}\,\frac{1}{p^2+m^2}\\ &\quad\times\left\{\frac{1}{p^2+m^2}+\frac{\Omega}{2}\delta(\theta''-\theta)\left[\bar\theta''-\bar\theta-\varepsilon(t)(\bar\theta''+\bar\theta)\right]\right\},\end{aligned} \tag{A36.7}$$

$$\begin{aligned}D_3 &= \frac{1}{(2\pi)^{2d}}\int\mathrm{d}t\,\mathrm{e}^{i\omega t}\int\mathrm{d}^dp_1\mathrm{d}^dp_2\,\mathrm{e}^{-\Omega s(p_i)|t|/2}\prod_{i=1}^{3}\frac{1}{p_i^2+m^2}\\ &\quad\times\left\{1+\frac{\Omega}{4}s\left\{p_i\right\}\delta(\theta-\theta')\left[\bar\theta-\bar\theta'-\varepsilon(t)(\bar\theta+\bar\theta')\right]\right\}\end{aligned} \tag{A36.8}$$

with the definitions

$$\mathbf{p}_3=-(\mathbf{k}+\mathbf{p}_1+\mathbf{p}_2), \tag{A36.9}$$

$$s(p_i)=\sum_{i=1}^{3}(p_i^2+m^2). \tag{A36.10}$$

The diagrams D_1 and D_2 (after integration over t) are identical to the diagrams of the 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

$$\begin{aligned}D_3^{(2)} &= \left[1-\tfrac{1}{2}i\omega(\theta-\theta')(\bar\theta+\bar\theta')\right]\int\mathrm{d}t\,\mathrm{e}^{i\omega t}\\ &\quad\times\int\frac{\mathrm{d}^dp_1}{(2\pi)^d}\frac{\mathrm{d}^dp_2}{(2\pi)^d}\,\mathrm{e}^{-\Omega s(p_i)|t|/2}\prod_{i=1}^{3}\frac{1}{p_i^2+m^2}\,.\end{aligned} \tag{A36.11}$$

The divergent part of the integral is a constant which can be calculated at $\omega = 0$ and $\mathbf{k} = 0$. Integrating over t, p_2 and p_1 successively, one finally obtains

$$Z_\omega = 1 - N_d^2 g^2 \ln(4/3) \frac{N+2}{24\varepsilon} + O\left(g^3\right), \tag{A36.12}$$

and, therefore, using the results of Section 11.6,

$$\eta_\omega(g) = N_d^2 \frac{N+2}{72} g^2 \left(6\ln(4/3) - 1\right) + O\left(g^3\right). \tag{A36.13}$$

A36.2 The Non-Linear σ-Model

We now consider the $O(N)$ non-linear σ-model (as discussed in Chapter 31), still with the supersymmetric dynamics, to illustrate the discussion of Sections 17.3, 17.4 and 36.2. The dynamical bare action in a magnetic field can be written as

$$\mathcal{S}(\boldsymbol{\phi}_0) = \frac{1}{g_0} \int \mathrm{d}\bar\theta \, \mathrm{d}\theta \, \mathrm{d}t \left[\int \mathrm{d}^d x \frac{2}{\Omega_0} \bar{\mathrm{D}} \boldsymbol{\phi}_0 \cdot \mathrm{D}\boldsymbol{\phi}_0 + \mathcal{A}(\boldsymbol{\phi}_0) \right] \tag{A36.14}$$

with

$$\mathcal{A}(\boldsymbol{\phi}_0) = \int \mathrm{d}^d x \, \frac{1}{2} \left(\partial_\mu \boldsymbol{\phi}_0\right)^2 - \int \mathrm{d}^d x \, \mathbf{h}_0 \cdot \boldsymbol{\phi}_0 . \tag{A36.15}$$

Note that for practical reasons, we have adopted normalizations which differ from those of Sections 17.3,17.4.

The renormalized theory is defined by the substitutions

$$\boldsymbol{\phi}_0 = \sqrt{Z} \{\sigma(t,x), \quad \boldsymbol{\pi}(t,x)\} \tag{A36.16}$$

with, then,

$$\sigma(t,x) = \sqrt{Z^{-1} - \boldsymbol{\pi}^2}, \tag{A36.17}$$

and

$$g_0 = \mu^{-\varepsilon} g Z_g, \quad \Omega_0 g_0 = \Omega g \mu^{-\varepsilon} Z/Z_\omega, \quad h_0 \sqrt{Z}/g_0 = h\mu^\varepsilon/g . \tag{A36.18}$$

We have called μ the renormalization scale and set $\varepsilon = d-2$.

In the case of dimensional regularization, the dynamic RG function $\eta_\omega(g)$ then is given by

$$\eta_\omega(g) = \mu \frac{\mathrm{d}}{\mathrm{d}\mu}\bigg|_{g_0,\Omega_0} \ln \Omega = \beta(g) \frac{\mathrm{d}}{\mathrm{d}g} \ln(Z_\omega Z_g/Z). \tag{A36.19}$$

Perturbation theory. The propagator is the same as in equation ($A36.2$) up to the normalization:

$$\Delta(\mathbf{k},\omega,\boldsymbol{\theta},\boldsymbol{\theta}') = g\mu^{-\varepsilon} \frac{\Omega\left[1 - \frac{1}{2} i\omega(\theta-\theta')(\bar\theta+\bar\theta') + \frac{1}{4}\Omega\left(k^2+h\right)\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}')\right]}{\omega^2 + \frac{1}{4}\Omega^2(k^2+h)^2}. \tag{A36.20}$$

To calculate the two-point function at two-loop order, we need the π^4 and the π^6 vertices:

$$\begin{aligned} V^{(4)} &= \frac{\mu^\varepsilon}{8g}\delta_{i_1 i_2}\delta_{i_3 i_4}\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)\delta^2(\boldsymbol{\theta}_3-\boldsymbol{\theta}_4) \\ &\times\left\{\frac{2}{\Omega}\left[-2+i(\omega_1+\omega_2)(\theta_1-\theta_3)(\bar\theta_1+\bar\theta_3)\right]+\left((p_1+p_2)^2+h\right)\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_3)\right\}, \quad (A36.21) \\ V^{(6)} &= \frac{\mu^\varepsilon}{16g}\delta_{i_1 i_2}\delta_{i_3 i_4}\delta_{i_5 i_6}\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)\delta^2(\boldsymbol{\theta}_3-\boldsymbol{\theta}_4)\delta^2(\boldsymbol{\theta}_3-\boldsymbol{\theta}_5)\delta^2(\boldsymbol{\theta}_3-\boldsymbol{\theta}_6) \\ &\times\left\{\frac{2}{\Omega}\left[-2+i\,(\omega_1+\omega_2)(\theta_1-\theta_3)(\bar\theta_1+\bar\theta_3)\right]+\delta^2(\boldsymbol{\theta}_1-\boldsymbol{\theta}_3)\left((p_1+p_2)^2+h\right)\right\}. \quad (A36.22) \end{aligned}$$

Two-point function at one-loop order. The values of the two diagrams of figure 31.1 now 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 (31.55,30.12))

$$I(h)=\Omega_d(\sqrt{h})=\frac{1}{(2\pi)^d}\int\frac{\mathrm{d}^d p}{p^2+h}=N_d\left(-\frac{1}{\varepsilon}-\frac{1}{2}\ln h\right)+O(\varepsilon). \quad (A36.25)$$

The two-point function at one-loop order follows

$$\begin{aligned} \frac{g}{\mu^\varepsilon}\Gamma^{(2)} &= -\frac{2Z_\omega}{\Omega}\left[2-i\omega(\theta-\theta')(\bar\theta+\bar\theta')\right]+\left(\frac{Z}{Z_g}k^2+h\sqrt{Z}\right)\delta^2(\boldsymbol{\theta}-\boldsymbol{\theta}') \\ &\quad+\tfrac{1}{2}(N-1)gD_1+gD_2+O\left(g^2\right). \quad (A36.26) \end{aligned}$$

We recover the one-loop static results:

$$Z=1+(N-1)g/(2\pi\varepsilon)+O\left(g^2\right),\quad Z_g=1+(N-2)g/(2\pi\varepsilon)+O\left(g^2\right),$$

and in addition we find

$$Z_\omega=Z/Z_g+O\left(g^2\right). \quad (A36.27)$$

This equation together with equation ($A36.19$) implies the absence of dynamics renormalization at this order:

$$\eta_\omega(g)=O\left(g^2\right). \quad (A36.28)$$

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 figure 31.2. Since the static renormalization constants have already been calculated, to calculate $\eta_\omega(g)$ at two-loop order we need only $\Gamma^{(2)}$ for vanishing arguments: $\mathbf{k}=0\,,\omega=0\,,\boldsymbol{\theta}=0$. Then, the corresponding contributions are

$$\begin{array}{ccccc} 2(-4I^2), & (N-1)(-4I^2), & 0, & 0\,, & -(-4I^2+I/\pi), \\ -\frac{1}{2}(N-1)(-I/\pi), & 0\,, & 0, & -(-8I^2+16J), & -\frac{1}{2}(N-1)(4I^2-32J). \end{array}$$

The integral J is given by

$$J = \frac{1}{(2\pi)^{2d}} \int \frac{\mathrm{d}^d p \, \mathrm{d}^d q}{(p^2 + h)\left(p^2 + q^2 + (p+q)^2 + 3h\right)}$$
$$= \frac{N_d^2}{4\varepsilon^2} \left(1 + \frac{\varepsilon}{2} \ln \frac{3}{4} + O\left(\varepsilon^2\right)\right). \qquad (A36.29)$$

We obtain, therefore,

$$\Gamma^{(2)} = -\frac{4Z_\omega}{\Omega} \Big[1 + gZ_g \left(hZ_g/Z\right)^{\varepsilon/2} \Omega_d + \tfrac{1}{2}(3N-5)\Omega_d^2 g^2 J g^2$$
$$- 4(N-2)Jg^2 - \tfrac{1}{4}(N-3)N_d \Omega_d g^2 + O\left(g^3\right)\Big]. \qquad (A36.30)$$

Expanding all terms, we find, finally, the expression of Z_ω at two-loop order:

$$Z_\omega = 1 + \frac{g}{2\pi\varepsilon} + \left(\frac{N-1}{2\varepsilon^2}\right) \frac{g^2}{(2\pi)^2} - \frac{N-2}{2\varepsilon} \ln \frac{4}{3} \frac{g^2}{(2\pi)^2} + O\left(g^3\right). \qquad (A36.31)$$

The expansion of $\eta_\omega(g)$ at order g^2 follows

$$\eta_\omega(g) = (N-2)\left[1 - \ln(4/3)\right] \frac{g^2}{(2\pi)^2} + O\left(g^3\right). \qquad (A36.32)$$

37 FIELD THEORY IN A FINITE GEOMETRY: FINITE SIZE SCALING

Many 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 have some idea about 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. We 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 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 modify 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 periodic boundary conditions are emphasized. Finally, both static and dynamical finite size effects are described.

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 QFT will be discussed mainly in Chapter 38, though the 2D non-linear σ-model with one infinite size is relevant to both situations.

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.

37.1 Renormalization Group in Finite Geometries

We assume in this chapter that the size of our system is 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 which *do not break translation symmetry* to avoid surface effects which are of a different nature. Periodic boundary conditions certainly satisfy such a criterion. Depending on the specific symmetries of a model, other boundary conditions are also available (like anti-periodic boundary conditions for Ising-like systems).

The crucial observation which explains finite size scaling is that renormalization theory which leads to RG equations is completely *insensitive to finite size effects* since renormalizations are entirely due to *short distance singularities*. As a consequence RG equations are not modified. However, their solutions are different because correlation functions now depend on one additional dimensional parameter L. We discuss below the solution of RG equations both in the examples of the ϕ^4 field theory and of the non-linear σ-model.

37.1.1 The $(\phi^2)^2$ field theory for $d < 4$

In the continuum limit, the N-vector model can be described by a field theory with the action (Chapter 26)

$$\mathcal{S}(\phi) = \int \left\{ \frac{1}{2} \left[\partial_\mu \phi(x)\right]^2 + \frac{1}{2}(r_c + t)\,\phi^2(x) + \frac{1}{4!}u\left(\phi^2(x)\right)^2 \right\} \mathrm{d}^d x\,, \tag{37.1}$$

where ϕ is a N-component field and t characterizes the deviation from the critical temperature. 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). In terms of the dimensionless coupling constant $g = u\Lambda^{d-4}$, the corresponding correlation functions satisfy for $d < 4$ the RG equations (Section 26.5):

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{1}{2}\eta(g)\left(n + M\frac{\partial}{\partial M}\right) - \eta_2(g)t\frac{\partial}{\partial t}\right]\Gamma^{(n)}(p_i;t,M,g,L,\Lambda) = 0\,. \tag{37.2}$$

They can be solved in the usual way by setting

$$\Gamma^{(n)}(p_i;t,M,g,L,\Lambda) = Z^{-n/2}(\lambda)\Gamma^{(n)}\left(p_i;t(\lambda),M/\sqrt{Z(\lambda)},g(\lambda),L,\lambda\Lambda\right), \tag{37.3}$$

the various functions of λ being defined by (equations (26.26–26.28))

$$\begin{aligned} \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}g(\lambda) &= \beta\big(g(\lambda)\big), & g(1) &= g\,, \\ \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}t(\lambda) &= -\eta_2\big(g(\lambda)\big), & t(1) &= t\,, \\ \lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}Z(\lambda) &= \eta\big(g(\lambda)\big), & Z(1) &= 1\,. \end{aligned}$$

However, the presence of the new length scale L now modifies the dimensional relations

$$\Gamma^{(n)}(p_i;t,M,g,L,\Lambda) = \Lambda^{d-n(d-2)/2}\Gamma^{(n)}(p_i/\Lambda;t/\Lambda^2,M\Lambda^{(2-d)/2},g,L\Lambda,1). \tag{37.4}$$

We use this relation in the r.h.s. of equation (37.3) and then choose λ such that

$$\lambda L\Lambda = 1 \quad \text{or} \quad \lambda = 1/L\Lambda\,. \tag{37.5}$$

The dilatation parameter λ goes to zero for ΛL large and, therefore, $g(\lambda)$ approaches the IR fixed point g^*. This implies for $Z(\lambda)$ and $t(\lambda)$, the behaviour ($\eta_2(g^*) = 1/\nu - 2$):

$$Z(\lambda) \propto (L\Lambda)^{-\eta}\,, \qquad \frac{t(\lambda)}{\lambda^2\Lambda^2} \propto \frac{t}{\Lambda^2}(L\Lambda)^{1/\nu}\,. \tag{37.6}$$

We have, therefore, derived the scaling of finite size correlation functions:

$$\Gamma^{(n)}(p_i;t,M,g,L,1) \propto L^{-d+n(d-2+\eta)/2}\Gamma^{(n)}(Lp_i;tL^{1/\nu},ML^{\beta/\nu},g^*,1,1). \tag{37.7}$$

It is characterized by the appearance of a new scaling variable $tL^{1/\nu} \propto (\xi/L)^{1/\nu}$, where for $t > 0$ $\xi(t)$ is the correlation length, and in general a RG invariant physical scale.

From equation (37.7), the usual infinite size scaling form is recovered by expressing that $\Gamma^{(n)}(L)$ has a limit for $L \gg \xi(t)$. In the opposite limit $\xi(t) \gg L$, correlation functions have a regular expansion in powers of t, even for zero 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 which are independent of the normalization of the field ϕ, of the temperature t, 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 immediately finds that the leading corrections to the scaling form (37.7) have near $d = 4$ the form of a scaling function multiplied by a factor $L^{-\omega}$ ($\omega = \beta'(g^*)$).

Examples. We now give two examples of quantities which have been considered in practical calculations. The quantity φ being the field average in a finite volume V:

$$\varphi = \frac{1}{V}\int \phi(x)\mathrm{d}^d x\,, \tag{37.8}$$

one calculates the ratio $\mathcal{R}_4$ of moments of the φ distribution,

$$\mathcal{R}_4 = \frac{m_4}{m_2^2} \quad \text{with} \quad m_{2p} = \left\langle |\varphi|^{2p} \right\rangle, \tag{37.9}$$

to determine the critical temperature. In zero magnetic field $\mathcal{R}_4$ has the scaling form

$$\mathcal{R}_4(t,L) \sim g(tL^{1/\nu}). \tag{37.10}$$

By calculating $\mathcal{R}_4(t,L)$ for different values of L and looking for a temperature at which it is independent of L, one finds the critical temperature $t = 0$, provided the corrections in $L^{-\omega}$ are negligible. The quantity $g(0)$ is a universal number which in principle can be calculated from the continuum field theory. We examine this problem later.

In a cylindrical geometry, the correlation length ξ_L in the infinite direction is another quantity of interest. From equation (37.7), one concludes

$$\xi_L \sim LX(tL^{1/\nu}). \tag{37.11}$$

In particular, at $t = 0$, ξ_L grows linearly with L and the ratio $\xi_L/L = X(0)$ is universal. Since ξ_L is related to the ratio of the two largest eigenvalues of the transfer matrix (equation (23.11)) λ_0 and λ_1, we learn also

$$\lambda_0/\lambda_1 - 1 \sim 1/\xi_L = 1/LX(0)\,.$$

With this knowledge, it is interesting to return to the analysis of the existence of phase transitions in Chapter 23.

Since for $t > 0$, ξ_L goes to a constant for large L, and since for $t < 0$ it grows faster than L, as one can easily verify, the ratio ξ_L/L can be used to determine the critical temperature in transfer matrix calculations.

37.1.2 Low temperature expansion and finite size effects

We have shown in Chapter 31 that in models with continuous symmetries, at low temperature the long distance behaviour is described by the effective interactions between Goldstone modes. In the case of the N-vector model, universal physical observables can be derived from the low temperature or low coupling expansion of the $O(N)$ non-linear σ-model. It is interesting to examine the problem of finite size effects also in this framework. Previous considerations concerning RG equations apply to the RG equations derived for the σ-model: equations remain unchanged, only solutions are modified by the finite size. General solutions (31.42,31.44) of the RG equations (31.23) 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 (31.34):

$$\xi(t) = \Lambda^{-1} t^{1/(d-2)} \exp\left[\int_0^t \left(\frac{1}{\beta(t')} - \frac{1}{(d-2)t'}\right) \mathrm{d}t'\right], \qquad (37.12)$$

and for $d = 2$ by

$$\xi(t) \propto \Lambda^{-1} \exp\left[\int^t \frac{\mathrm{d}t'}{\beta(t')}\right]. \qquad (37.13)$$

Alternatively, the solutions can be parametrized in terms of a size-dependent temperature t_L, obtained by solving the equation

$$\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda} t(\lambda) = \beta\left[t(\lambda)\right], \qquad t(1) = t \qquad (37.14)$$

at scale $\lambda = 1/\Lambda L$:

$$t_L \equiv t(1/\Lambda L). \qquad (37.15)$$

Then,

$$\ln(\Lambda L) = \int_{t_L}^t \frac{\mathrm{d}t'}{\beta(t')}, \qquad (37.16)$$

which shows in particular that t_L is a function of t and L only through the expected combination $L/\xi(t)$.

At one-loop order, the RG $\beta(t)$ function in a cut-off scheme has the expansion (Section 31.2.2)

$$\beta(t) = (d-2)t + \beta_2(d)t^2 + O\left(t^3\right) \qquad (37.17)$$

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 $t = 0$:

$$t_L \sim \left(\xi(t)/L\right)^{d-2} \ll 1. \qquad (37.18)$$

Therefore, finite size effects can be calculated from the low temperature expansion and renormalization group.

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 31. Since t_c is a 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 a UV fixed point, t_L goes to zero *for* $L/\xi(t)$ *small*,

$$t_L \sim \frac{2\pi}{(N-2)\ln(\xi(t)/L)}, \tag{37.19}$$

and this is the limit in which physical quantities can be calculated.

Finally, solving equation (37.16) perturbatively we find

$$\frac{1}{t_L} = \frac{(\Lambda L)^{d-2}}{t} + \frac{\beta_2(d)}{d-2}\left[(\Lambda L)^{d-2} - 1\right] + O(t). \tag{37.20}$$

Notice here that a statistical (classical) field theory in two dimensions with a finite size in one of the dimensions, and periodic boundary conditions is also a finite temperature QFT (Chapter 38). Therefore, the considerations of Section 37.5.2 are also relevant for such a quantum theory.

37.2 Momentum Quantization

The scaling properties (37.7) do not depend on the specific form 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, the characteristic feature of all finite geometries is that, in the Fourier representation, momenta which correspond to directions in which the size of the system is finite are quantized. However, the precise momentum spectrum varies with 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 37.2.2. An example has actually been worked out in Section 3.4.

37.2.1 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 $\mathbf{p}$ have the form

$$\mathbf{p} = 2\pi\mathbf{k}/L\,, \qquad \mathbf{k} \in \mathbb{Z}^d. \tag{37.21}$$

When the product $tL^{1/\nu} = (L/\xi)^{1/\nu}$ 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 like $\exp[-\text{const.}\ L/\xi]$. When the product $tL^{1/\nu}$ 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, at T_c in a finite volume the propagator has an isolated pole at $\mathbf{p}=\mathbf{0}$ 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, perturbation theory is badly behaved for $\xi \gg L$.

In the cylindrical geometry, one component, ω of the momentum varies continuously and the other components are quantized. At T_c Feynman diagrams receive divergent contributions of the form $\int \mathrm{d}\omega/\omega^2$. 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 will appear in Chapter 38 devoted to finite temperature QFT. Then, the IR problem can no longer be solved exactly.

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 separate the zero momentum Fourier component of the field. The components $\mathbf{k} \neq \mathbf{0}$ can be treated by the methods developed in the infinite geometry (perturbation theory and RG), the component $\mathbf{k} = \mathbf{0}$ whose fluctuations are damped at T_c only by interaction terms, has to be treated exactly. In the case of a finite volume, we therefore construct an effective integral over the component $\tilde{\phi}(\mathbf{p} = \mathbf{0})$ by integrating over all other components. In a cylindrical geometry, the integration over all components except $\tilde{\phi}(\mathbf{p}_{\mathrm{T}} = \mathbf{0}, \tau)$, denoting by τ the coordinate in the infinite direction, leads to an effective quantum hamiltonian. Note that similar considerations will apply to the zero modes of instanton calculations (see Chapters 39–43).

We examine in Section 37.3, the two first geometries separately, beginning with the simplest case of the periodic hypercube.

37.2.2 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 antiperiodic 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 $\boldsymbol{\phi}$, the boundary conditions

$$\boldsymbol{\phi}(x_1, x_2, \ldots, x_\mu + L, \ldots, x_d) = \mathrm{e}^{A_\mu}\, \boldsymbol{\phi}(x_1, x_2, \ldots, x_\mu, \ldots, x_d),$$

where the A_μ's are constant (i.e. space independent) commuting antihermitian 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. We set

$$\boldsymbol{\phi}(\mathbf{x}) = \mathrm{e}^{A_\mu x_\mu / L}\, \boldsymbol{\phi}'(\mathbf{x}).$$

The new field $\boldsymbol{\phi}'$ then satisfies periodic boundary conditions. The action is invariant except for the derivatives which now become covariant derivatives:

$$\mathrm{e}^{-A_\mu x_\mu / L}\, \partial_\mu \boldsymbol{\phi} = (\partial_\mu + A_\mu / L)\, \boldsymbol{\phi}' \;\Rightarrow\; p_\mu \tilde{\boldsymbol{\phi}}(p) \mapsto (p_\mu - i A_\mu / L)\, \tilde{\boldsymbol{\phi}}'(p).$$

After Fourier transformation, the derivatives thus yield quantized momenta of the form $p_\mu = (2\pi k_\mu - i A_\mu)/L$. The antihermitian matrices A_μ have imaginary eigenvalues $i\theta^\alpha_\mu$. Writing the field ϕ' in a basis in which all A_μ (which commute) are diagonal, we find that the effect of twisted boundary conditions has been to generate a set of quantized momenta of the form

$$p_\mu = (2\pi k_\mu + \theta^\alpha_\mu)/L\,, \qquad k_\mu \in \mathbb{Z}^d,$$

for each component ϕ_α of the field. We can choose all angles θ^α_μ to belong to the interval $[-\pi, \pi]$.

37.3 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 explained in Section 37.2.1, we expand $\phi(x)$ in Fourier components, separating the zero mode:

$$\begin{aligned}\phi(x) &= \varphi + \chi(x),\\ \chi(x) &= (L/2\pi)^d \sum_{\mathbf{p}\neq\mathbf{0}} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{x}}\,\tilde{\phi}(\mathbf{p}), \qquad \mathbf{p} = 2\pi\mathbf{k}/L\,, \quad \mathbf{k}\in\mathbb{Z}^d.\end{aligned} \tag{37.22}$$

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 $\mathbf{k} = \mathbf{0}$ 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 the average

$$\varphi = L^{-d}\int \phi(x)\mathrm{d}^d x = (L/2\pi)^d\tilde{\phi}(0),$$

here remains a fluctuating variable (see also the discussion of Section 7.10).

As an illustration, we calculate expectation values of the form (37.9), moments of the distribution of the average spin per unit volume in a spin system. We set

$$\exp[-\Sigma(\varphi)] = \mathcal{N}^{-1}\int [\mathrm{d}\chi]\exp[-\mathcal{S}(\varphi+\chi)], \tag{37.23}$$

where $\mathcal{S}(\phi)$ is the action (37.1) and the normalization $\mathcal{N}$ is chosen such that $\Sigma(0) = 0$ for $t = 0$. Moments then are given by

$$m_\sigma = \mathcal{Z}^{-1}\int \mathrm{d}\varphi|\varphi|^\sigma \exp[-\Sigma(\varphi)], \tag{37.24}$$

where $\mathcal{Z}$ is the partition function:

$$\mathcal{Z} = \int \mathrm{d}\varphi\exp[-\Sigma(\varphi)].$$

It follows from the discussion of Section 7.10 that in the infinite volume limit $\Sigma(\varphi) = \Gamma(\varphi) - \Gamma(\mathbf{0})$, where $\Gamma(\varphi)$ is the thermodynamic potential as obtained in perturbation theory. It satisfies the same RG equation

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{1}{2}\eta(g)\varphi\frac{\partial}{\partial\varphi} - \eta_2(g)t\frac{\partial}{\partial t}\right]\Sigma(t,\varphi,g,L,\Lambda) = R(t,g,\Lambda)\,, \tag{37.25}$$

where we recall that R is a second degree polynomial in t.

The moments m_{2s}, s integer, involve only powers of φ^2 and thus are related to zero momentum correlation functions:

$$\varphi^2 = L^{-2d}\int \mathrm{d}^d x\,\mathrm{d}^d y\,\phi(x)\cdot\phi(y) = (2\pi/L)^{2d}\tilde{\phi}^2(0). \tag{37.26}$$

Though for σ generic this is no longer the case, m_σ nevertheless satisfies a usual RG equation which can easily be derived from equation (37.25):

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} + \frac{1}{2}\sigma\eta(g) - \eta_2(g)t\frac{\partial}{\partial t}\right] m_\sigma(t,g,L,\Lambda) = 0\,. \tag{37.27}$$

Another quantity of interest is the specific heat

$$C(L,t) = L^{-d}\frac{\partial^2}{(\partial t)^2}\ln\mathcal{Z}\,. \tag{37.28}$$

37.3.1 Dimensions $d > 4$

We first discuss the case $d > 4$, which in the infinite geometry is simple. To obtain the zero mode integral at leading order, we simply neglect the fluctuations of all non-zero modes ($\chi = 0$). The function (37.23) then reduces to

$$\Sigma(\varphi) = \mathcal{S}(\varphi) = L^d \left(\frac{1}{2}t\varphi^2 + \frac{1}{4!}u\left(\varphi^2\right)^2\right), \tag{37.29}$$

where $\mathcal{S}$ is the action (37.1). After the change of variables

$$\varphi \mapsto \left(uL^d\right)^{-1/4}\varphi, \tag{37.30}$$

the moments m_σ take the form

$$m_\sigma(L,t) = \left(uL^d\right)^{-\sigma/4}\mu_\sigma(tL^{d/2}u^{-1/2}), \tag{37.31}$$

in which $\mu_\sigma(z)$ is given by

$$\mu_\sigma(z) = \frac{g_{\sigma+N}(z)}{g_N(z)} \tag{37.32}$$

with

$$g_\sigma(z) = \int_0^\infty \mathrm{d}\varphi\,\varphi^{\sigma-1}\exp\left[-\left(\frac{1}{2}z\varphi^2 + \frac{1}{4!}\varphi^4\right)\right]. \tag{37.33}$$

Equation (37.31) shows that above four dimensions the finite size scaling relations, proven for a non-trivial fixed point, and which predict instead for the moment m_σ a behaviour

$$m_\sigma = L^{-\sigma(d-2)/2}\tilde{\mu}_\sigma(tL^2)$$

do not hold. In particular, instead of the argument $tL^2 \propto (L/\xi)^2$ one finds $tL^2L^{(d-4)/2}$. 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 problem of hyperscaling in Section 27.1).

The result (37.31) leads to universal predictions. For instance dimensionless ratios of moments like

$$\mathcal{R}_\sigma(t) = m_\sigma / \left(m_2\right)^{\sigma/2}, \tag{37.34}$$

at $T = T_c$ ($t = 0$) are universal. Calculating explicitly the integral (37.33) at $z = 0$:

$$g_\sigma(0) = \tfrac{1}{4}(24)^{\sigma/4}\Gamma(\sigma/4),$$

one finds

$$\mathcal{R}_\sigma(T = T_c) = \Gamma((\sigma+N)/4)\left[\Gamma(N/4)\right]^{\sigma/2-1}\left[\Gamma((N+2)/4)\right]^{-\sigma/2}. \tag{37.35}$$

In particular, for Ising-like systems, the quantity $\mathcal{R}_4$ is

$$\mathcal{R}_4 = \Gamma^4(1/4)/(8\pi^2) = 2.1884\cdots. \tag{37.36}$$

Another quantity of interest is m_2/m_1^2:

$$\mathcal{R}_1^{-2} = m_2/m_1^2 = \sqrt{2}. \tag{37.37}$$

Note finally that at leading order the specific heat C (equation (37.28)) is simply given by

$$C(L,t) = L^d \left(m_4 - m_2^2\right)/4 = \left(\mu_4 - \mu_2^2\right)/4u + \text{const.}\,,$$

where the constant comes from the regular part of the free energy.

Remarks. The separation of the zero mode makes *a priori* only sense if the correlation length is much larger than the system size, that is, $tL^2 \ll 1$. When the scaling variable $tL^{d/2}$ which appears in expression (37.31) is finite the condition is satisfied for $d > 4$. One verifies that, in addition, the expression (37.31) has both for $t < 0$ and $t > 0$ fixed the correct $L \to \infty$ behaviour. Indeed for $t < 0$, one finds

$$m_\sigma(L,t) \to (-6t/u)^{\sigma/2} \equiv [M_0(t)]^\sigma\,,$$

where $M_0(t)$ is the infinite size spontaneous magnetization at this order. For $t > 0$, one obtains instead

$$m_\sigma(L,t) \propto \left(\chi(t)L^{-d}\right)^{\sigma/2},$$

where $\chi(t) = 1/t$ is the infinite size magnetic susceptibility.

The correction terms to this leading behaviour, however, are incorrect. Their determination involves additional higher order contributions.

Higher order corrections. It now 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.

(i) According to the analysis of Chapter 27, 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 t the divergent terms yield a contribution $\delta\Sigma_{\text{div}}(\varphi,t)$ to Σ of the form

$$\delta\Sigma_{\text{div}}(\varphi,t) = L^d \sum_{l,m} \Sigma_{lm}\Lambda^{d-2m-l(d-2)} t^m \varphi^{2l},$$

where the coefficients Σ_{lm} are numbers. After the change of variables (37.30), and taking into account that t is of order $L^{-d/2}$ (equation (37.31)) we find that a term proportional to $\varphi^{2l}t^m$ gives a contribution of order $(\Lambda L)^{d(2-m-l)/2}$. Therefore, only terms with $m+l \le 2$ survive. The terms $l > 0$ shift r_c and yield finite renormalizations of t and u. The two remaining terms are proportional to t and t^2. They cancel in the moments (37.24). The t^2 term yields a constant non-universal contribution to the specific heat. Finally, the leading corrections ($l + m = 3$) are of order $(L\Lambda)^{-d/2}$.

(ii) After renormalization, the loop corrections can be formally expanded in powers of t 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.

We first examine the one-loop corrections, generated in equation (37.23) by an integration over χ in the gaussian approximation,

$$\Sigma_{1\,\text{loop}}(\varphi,L,t,u) = \frac{1}{2}\sum_{\mathbf{k}\neq 0} \operatorname{tr}\ln\left[\delta_{ij} + \frac{1}{(2\pi\mathbf{k}/L)^2}\left(\left(t+\frac{u}{6}\varphi^2\right)\delta_{ij} + \frac{u}{3}\varphi_i\varphi_j\right)\right]$$

$$= \frac{1}{2} \sum_{\mathbf{k} \neq 0} \left[\ln \left(1 + \frac{t + \frac{1}{2} u \varphi^2}{(2\pi \mathbf{k}/L)^2} \right) + (N-1) \ln \left(1 + \frac{t + \frac{1}{6} u \varphi^2}{(2\pi \mathbf{k}/L)^2} \right) \right] . \tag{37.38}$$

After renormalization, one finds from dimensional analysis that the coefficient of $t^m (u\varphi^2)^l$ in the expansion of $\Sigma_{1\,\mathrm{loop}}$ is proportional to L^{2m+2l}. Performing the change of variables (37.30), and taking into account that t is of order $u^{1/2} L^{-d/2}$ (equation (37.31)) one obtains a contribution proportional to

$$t^m (\varphi^2)^l \propto \left(u L^{4-d} \right)^{(l+m)/2} ,$$

which goes to zero for ΛL large and $d > 4$.

The latter argument can easily 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 uL^{4-d} for dimensional reasons. It follows that a term of ℓ-loop order proportional to $\varphi^{2l} t^m$ is of order $(uL^{4-d})^\kappa$ with $\kappa = \ell - 1 + \frac{1}{2}(l+m)$.

The conclusion is that for $d > 4$ the effective action (37.29) 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 (37.24). In particular, the result (37.35) is indeed universal.

Finally, the leading term at two-loop order has $\kappa = 3/2$ and thus the only meaningful corrections at one-loop order correspond to $m + l \leq 2$. A short calculation then shows that for what concerns the dimensionless ratios (37.34), for $d < 8$ the leading corrections can be reproduced by replacing t by a quantity t_L which has the form

$$t_L = \left(1 + A_1 u L^{4-d} \right) t + A_2 u L^{2-d},$$

A_1, A_2 being two constants. The specific heat receives an additional contribution of order L^{4-d}.

37.3.2 Dimension $d = 4 - \varepsilon$

We now use the RG arguments presented in Section 37.1: instead of calculating physical quantities as function of $\{t, M, u, L\}$, we can, in the critical domain, set $\Lambda = L = 1$, $u = u^*$ the IR fixed point value, then replace t by $tL^{1/\nu}$, M (if we introduce 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

$$\Sigma(\varphi, L = 1, t, u^*) = \tfrac{1}{2} t \varphi^2 + \frac{u^*}{4!} \varphi^4 .$$

At the same order, u^* can be replaced by its value at order ε:

$$u^* = \frac{48\pi^2 \varepsilon}{N+8} + O\left(\varepsilon^2\right). \tag{37.39}$$

Then, replacing t by $tL^{1/\nu}$ and φ by $\varphi L^{\beta/\nu}$, and integrating over φ we find the moments m_σ at leading order

$$m_\sigma(L, t) = \frac{L^{-\sigma(d-2+\eta)/2}}{(u^*)^{\sigma/4}} \mu_\sigma \left(t L^{1/\nu} (u^*)^{-1/2} \right), \tag{37.40}$$

where μ_σ has been defined in (37.32). 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 t is formally assumed to be of order $\varepsilon^{1/2}$ (this condition is realized in particular at $t \propto T - T_c = 0$).

Note the appearance of powers of $(u^*)^{1/2}$ which, for ε small, is equivalent to $\varepsilon^{1/2}$. This suggests that physical quantities will have an expansion in powers of $\varepsilon^{1/2}$ rather than ε. The analysis of higher order corrections confirms this observation. Let us exhibit this phenomenon in the one-loop approximation.

One-loop calculation. We have already displayed in equation (37.38), the one-loop contribution $\Sigma_{1\,\text{loop}}$ to Σ. The divergent part of the one-loop term has first to be subtracted. We then set $L = 1$, $u = u^*$. As we have already discussed at the end of Section 37.3.1, at L fixed all terms in perturbation theory can then be expanded in powers of $\boldsymbol{\varphi}^2$ and t. After the change of variables (37.30), φ^2 has a coefficient proportional to $(u^*)^{1/2} \sim \varepsilon^{1/2}$. In the same way, t is of order $\varepsilon^{1/2}$. A term contributing to the ℓ-loop order and proportional to $\varphi^{2l} t^m$ 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 one-loop order only the terms proportional to φ^2, $\varphi^2 t$, $\varphi^4\, t$ and t^2 have to be considered. The form of $\Sigma(\boldsymbol{\varphi})$ at one-loop order thus is

$$\Sigma(\varphi, L = 1, t, u^*) - \Sigma(0, 1, t, u^*) = \frac{1}{2}\left(t(1 + a_1 u^*) + a_2 u^*\right)\varphi^2 + u^*(1 + a_3 u^*)\frac{\varphi^4}{4!}.$$

Note that the correction to the coefficient of φ^4 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 t can be absorbed into a finite change of normalization of t. The conclusion is the only correction relevant at one-loop order 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 (37.38) is

$$\tilde{a}_2 = \frac{N+2}{6} \sum_{\mathbf{k} \neq \mathbf{0}} \frac{1}{(2\pi \mathbf{k})^2}. \tag{37.41}$$

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 Appendix A37.2, in terms of the function $\vartheta_0(s)$ (related to elliptic functions),

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-\pi s n^2}. \tag{37.42}$$

we can formally write

$$\sum_{\mathbf{k} \neq \mathbf{0}} \frac{1}{(2\pi \mathbf{k})^2} = \frac{1}{4\pi} \int_0^{+\infty} \mathrm{d}s \left(\vartheta_0^d(s) - 1\right). \tag{37.43}$$

The integral in the r.h.s. converges exponentially for s large. For $s \to 0$, one finds (Appendix A37.2)

$$\vartheta_0(s) - (1/s)^{1/2} \underset{s \to 0}{\sim} 2 s^{-1/2}\, \mathrm{e}^{-\pi/s}, \tag{37.44}$$

where $s^{-1/2}$ corresponds to the infinite size limit. The integral thus diverges at $s = 0$. Subtracting the infinite size limit we obtain for $d > 2$ a finite result

$$\sum_{\mathbf{k}\neq 0} \frac{1}{(2\pi\mathbf{k})^2} - \frac{1}{(2\pi)^d}\int \frac{\mathrm{d}^d p}{\mathbf{p}^2} = \frac{1}{4\pi}\int_0^{+\infty} \mathrm{d}s \left(\vartheta_0^d(s) - 1 - s^{-d/2}\right). \tag{37.45}$$

The coefficient a_2 follows:

$$a_2 = \frac{N+2}{24\pi}\int_0^{+\infty} \mathrm{d}s \left(\vartheta_0^d(s) - 1 - s^{-d/2}\right). \tag{37.46}$$

Introducing rescaled temperature $t' \propto t$ and field ϕ we can rewrite the moments at one-loop order:

$$m_\sigma(L, t') = L^{-\sigma(d-2+\eta)/2} f_\sigma(t' L^{1/\nu} + b), \tag{37.47}$$

where the constant b is given by (37.39,37.46)

$$b = a_2 u^{*1/2} = \frac{N+2}{\sqrt{N+8}}\frac{(3\varepsilon)^{1/2}}{6}\int_0^{+\infty} \mathrm{d}s \left[\vartheta_0^4(s) - 1 - 1/s^2\right] + O\left(\varepsilon^{3/2}\right). \tag{37.48}$$

The ratio $\mathcal{R}_4(T_c)$. From expression (37.47), setting $t = 0$ we immediately derive the universal dimensionless ratio $\mathcal{R}_4(T = T_c)$ at order $\varepsilon^{1/2}$:

$$\mathcal{R}_4(T_c) = \frac{g_{4+N}(b) g_N(b)}{\left[g_{2+N}(b)\right]^2}, \tag{37.49}$$

g_σ being defined by equation (37.33). Using the value of the integral

$$\int_0^{+\infty} \mathrm{d}s \left[\vartheta_0^4(s) - 1 - 1/s^2\right] = -0.561843942\ldots,$$

we obtain in three dimensions for $N = 1$,

$$\mathcal{R}_4(T_c) = 1.800\ldots. \tag{37.50}$$

This result should 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 t^2 terms, both divergent and convergent in Σ have to be taken into account. A short calculation yields

$$C(L, t) = L^{\alpha/\nu}\left[\frac{1}{4}\left(\mu_4(t' L^{1/\nu} + b) - \mu_2^2(t' L^{1/\nu} + b)\right) + \frac{3N}{4-N}\right] + \text{const.},$$

where again the constant comes from the non-universal regular contribution.

37.4 The ϕ^4 Field Theory: The 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 now expand the fields in Fourier components in the $d-1$ space dimensions:

$$\begin{aligned}\phi(\tau,x) &= \varphi(\tau)+\chi(\tau,x),\\ \chi(\tau,x) &= (2\pi/L)^{d-1}\sum_{\mathbf{k}\in\mathbb{Z}^{d-1}\neq\mathbf{0}} e^{i2\pi\mathbf{k}\cdot\mathbf{x}/L}\,\phi_{\mathbf{k}}(\tau)\,.\end{aligned} \tag{37.51}$$

We again consider only the simple example of correlation functions of space integrals

$$\varphi(\tau)=L^{1-d}\int d^{d-1}x\,\phi(x,\tau).$$

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 then are given by simple path integrals of quantum mechanics type, with the effective action $\mathcal{S}_L$.

We illustrate the method with the calculation of the finite size correlation length ξ_L. We first examine the case $d>4$ which is described by mean field theory in the infinite volume limit.

37.4.1 Dimensions $d>4$

The leading order approximation is obtained by neglecting all corrections due the integration over the $\mathbf{k}\neq\mathbf{0}$ components of the field

$$\mathcal{S}_L(\varphi)=\mathcal{S}(\varphi)=L^{d-1}\int d\tau\left[\frac{1}{2}\left(\dot{\varphi}\right)^2+\frac{t}{2}\varphi^2+\frac{u}{4!}\left(\varphi^2\right)^2\right]. \tag{37.52}$$

To simplify the action, we rescale φ and time τ:

$$\varphi\mapsto u^{-1/6}L^{(1-d)/3}\varphi\,,\quad \tau=u^{-1/3}L^{(d-1)/3}\tau'\,. \tag{37.53}$$

The rescaling (37.53) corresponds for the Fourier (energy) variable E associated with τ to

$$E\mapsto L^{-(d-1)/3}u^{1/3}E\,. \tag{37.54}$$

We obtain an action

$$\mathcal{S}'_L(\varphi)=\int d\tau'\left[\frac{1}{2}\left(\dot{\varphi}\right)^2+\frac{1}{2}u^{-2/3}L^{2(d-1)/3}t\varphi^2+\frac{1}{4!}\left(\varphi^2\right)^2\right]. \tag{37.55}$$

To this action is associated the quantum hamiltonian $\widehat{H}$ (see Chapter 2):

$$\widehat{H}=\tfrac{1}{2}\hat{\mathbf{p}}^2+\tfrac{1}{2}z\hat{\mathbf{q}}^2+\frac{1}{4!}\left(\hat{\mathbf{q}}^2\right)^2 \tag{37.56}$$

37.4 The ϕ^4 Field Theory: The 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 now expand the fields in Fourier components in the $d-1$ space dimensions:

$$\begin{aligned}\phi(\tau,x) &= \varphi(\tau)+\chi(\tau,x),\\ \chi(\tau,x) &= (2\pi/L)^{d-1}\sum_{\mathbf{k}\in\mathbb{Z}^{d-1}\neq\mathbf{0}} \mathrm{e}^{i2\pi\mathbf{k}\cdot\mathbf{x}/L}\,\phi_{\mathbf{k}}(\tau)\,.\end{aligned} \tag{37.51}$$

We again consider only the simple example of correlation functions of space integrals

$$\varphi(\tau)=L^{1-d}\int \mathrm{d}^{d-1}x\,\phi(x,\tau).$$

These can be calculated using only the effective action $\mathcal{S}_L(\varphi)$ obtained by integrating over χ,

$$\exp[-\mathcal{S}_L(\varphi)]=\mathcal{N}^{-1}\int[\mathrm{d}\chi]\exp[-\mathcal{S}(\varphi+\chi)],$$

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 then are given by simple path integrals of quantum mechanics type, with the effective action $\mathcal{S}_L$.

We illustrate the method with the calculation of the finite size correlation length ξ_L. We first examine the case $d>4$ which is described by mean field theory in the infinite volume limit.

37.4.1 Dimensions $d>4$

The leading order approximation is obtained by neglecting all corrections due the integration over the $\mathbf{k}\neq\mathbf{0}$ components of the field

$$\mathcal{S}_L(\varphi)=\mathcal{S}(\varphi)=L^{d-1}\int \mathrm{d}\tau\left[\frac{1}{2}\left(\dot\varphi\right)^2+\frac{t}{2}\varphi^2+\frac{u}{4!}\left(\varphi^2\right)^2\right]. \tag{37.52}$$

To simplify the action, we rescale φ and time τ:

$$\varphi\mapsto u^{-1/6}L^{(1-d)/3}\varphi\,,\quad \tau=u^{-1/3}L^{(d-1)/3}\tau'\,. \tag{37.53}$$

The rescaling (37.53) corresponds for the Fourier (energy) variable E associated with τ to

$$E\mapsto L^{-(d-1)/3}u^{1/3}E\,. \tag{37.54}$$

We obtain an action

$$\mathcal{S}'_L(\varphi)=\int \mathrm{d}\tau'\left[\frac{1}{2}\left(\dot\varphi\right)^2+\frac{1}{2}u^{-2/3}L^{2(d-1)/3}t\varphi^2+\frac{1}{4!}\left(\varphi^2\right)^2\right]. \tag{37.55}$$

To this action is associated the quantum hamiltonian $\widehat{H}$ (see Chapter 2):

$$\widehat{H}=\tfrac{1}{2}\hat{\mathbf{p}}^2+\tfrac{1}{2}z\hat{\mathbf{q}}^2+\frac{1}{4!}\left(\hat{\mathbf{q}}^2\right)^2 \tag{37.56}$$

with $z = u^{-2/3}L^{2(d-1)/3}t$. The finite size correlation length ξ_L is related to the two lowest eigenvalues E_0 and E_1 of $\widehat{H}$ rescaled according to (37.54):

$$\xi_L = u^{-1/3}L^{(d-1)/3}X(tL^{2(d-1)/3}u^{-2/3}) \tag{37.57}$$

with

$$X(z) = (E_1(z) - E_0(z))^{-1}. \tag{37.58}$$

In terms of the infinite size correlation length ξ_∞, this expression can be rewritten as

$$\xi_L(t) = \xi_\infty(t)Y\left(L^{(d-4)/3}L/\xi_\infty(t)\right), \qquad Y(z) = \sqrt{z}X(z). \tag{37.59}$$

The expression exhibits a violation of the naive extension of the scaling form (37.11), proven for $d < 4$. The reason here is, as before, that in case of the gaussian fixed point, the amplitude of expected leading terms vanishes and the first non-trivial contributions correspond to what are for $d < 4$ only corrections to scaling (see Section 27.1).

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. Let us now show that loop corrections due to the integration over the non-zero modes do not modify the scaling form (37.59) at leading order. These corrections are of two types:

(i) Corrections corresponding due 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

$$\delta\mathcal{S}_{L\mathrm{div}}(\varphi) = L^{d-1}\int \mathrm{d}\tau\, \mathcal{L}(\varphi),$$

where $\mathcal{L}(\varphi)$ is a sum of monomials in $\varphi(\tau)$ and its derivatives. After the changes of variables (37.53), 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 $k+n \le 2$ simply renormalize the coefficients of the leading order action. All other terms, which are absent from the leading order action, have vanishing coefficients. The simplest example is ϕ^6 which induces a correction proportional to $L^{-2(d-1)/3}\hat{q}^6$ in $\widehat{H}$.

(ii) Corrections which are finite in the large cut-off limit. The one-loop correction $\mathcal{S}_{L\,1\,\mathrm{loop}}(\varphi)$ to the effective action, analogous to expression (37.38), is given in the initial φ, τ variables by

$$\begin{aligned}\mathcal{S}_{L\,1\,\mathrm{loop}}(\varphi) = \frac{1}{2}\sum_{\mathbf{k}\neq 0}\Big\{&\mathrm{tr}\ln\left[1 + \left(-\partial_\tau^2 + (2\pi\mathbf{k}/L)^2\right)^{-1}\left(t + u\varphi^2(\tau)/2\right)\right]\\ &+(N-1)\,\mathrm{tr}\ln\left[1 + \left(-\partial_\tau^2 + (2\pi\mathbf{k}/L)^2\right)^{-1}\left(t + u\varphi^2(\tau)/6\right)\right]\Big\}.\end{aligned} \tag{37.60}$$

After renormalization, it yields a UV finite non-local contribution. It would seem, therefore, at first sight that we can no longer use a hamiltonian formalism to evaluate corrections to the mean field approximation.

However, we first notice that φ appears only in the combination $uL^2\varphi^2$. After the change (37.53), 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 t appears only in the combination tL^2 which, according to the form (37.57), 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 t.

with $z = u^{-2/3}L^{2(d-1)/3}t$. 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 (37.54):

$$\xi_L = u^{-1/3}L^{(d-1)/3}X(tL^{2(d-1)/3}u^{-2/3}) \tag{37.57}$$

with

$$X(z) = (E_1(z) - E_0(z))^{-1}. \tag{37.58}$$

In terms of the infinite size correlation length ξ_∞, this expression can be rewritten as

$$\xi_L(t) = \xi_\infty(t)Y\left(L^{(d-4)/3}L/\xi_\infty(t)\right), \qquad Y(z) = \sqrt{z}X(z). \tag{37.59}$$

The expression exhibits a violation of the naive extension of the scaling form (37.11), proven for $d < 4$. The reason here is, as before, that in case of the gaussian fixed point, the amplitude of expected leading terms vanishes and the first non-trivial contributions correspond to what are for $d < 4$ only corrections to scaling (see Section 27.1).

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. Let us now show that loop corrections due to the integration over the non-zero modes do not modify the scaling form (37.59) at leading order. These corrections are of two types:

(i) Corrections corresponding due 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

$$\delta\mathcal{S}_{L\mathrm{div}}(\varphi) = L^{d-1}\int \mathrm{d}\tau\, \mathcal{L}(\varphi),$$

where $\mathcal{L}(\varphi)$ is a sum of monomials in $\varphi(\tau)$ and its derivatives. After the changes of variables (37.53), 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 $k+n \le 2$ simply renormalize the coefficients of the leading order action. All other terms, which are absent from the leading order action, have vanishing coefficients. 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 which are finite in the large cut-off limit. The one-loop correction $\mathcal{S}_{L\,1\,\mathrm{loop}}(\varphi)$ to the effective action, analogous to expression (37.38), is given in the initial φ, τ variables by

$$\begin{aligned}\mathcal{S}_{L\,1\,\mathrm{loop}}(\varphi) = \frac{1}{2}\sum_{\mathbf{k}\neq 0}\Big\{&\operatorname{tr}\ln\left[1 + \left(-\partial_\tau^2 + (2\pi\mathbf{k}/L)^2\right)^{-1}\left(t + u\varphi^2(\tau)/2\right)\right]\\ &+(N-1)\operatorname{tr}\ln\left[1 + \left(-\partial_\tau^2 + (2\pi\mathbf{k}/L)^2\right)^{-1}\left(t + u\varphi^2(\tau)/6\right)\right]\Big\}.\end{aligned} \tag{37.60}$$

After renormalization, it yields a UV finite non-local contribution. It would seem, therefore, at first sight that we can no longer use a hamiltonian formalism to evaluate corrections to the mean field approximation.

However, we first notice that φ appears only in the combination $uL^2\varphi^2$. After the change (37.53), 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 t appears only in the combination tL^2 which, according to the form (37.57), 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 t.

Finally, we face the problem of the non-local operator $[-L^2\partial_\tau^2 + (2\pi\mathbf{k})^2]^{-1}$. After the change of variables (37.53), the differential operator $L^2\partial_\tau^2$ becomes also of order $u^{2/3}L^{2(4-d)/3}$ and thus is small compared to $\mathbf{k}^2$. We then know from general arguments that the action (37.60) has a local expansion. Actually, this problem is related to the classical problem of expanding the resolvent of the Schrödinger operator at high energy

$$\operatorname{tr}\left[-\mathrm{d}_q^2 + U(q) - E\right]^{-1} = \sum_{n=0} E^{-1/2-n}\Pi_n(U),$$

where one finds that the $\Pi_n(U)$ are local polynomials in the potential U.

We have thus shown that for L large the effective action $\mathcal{S}_L$ has, at one-loop order, a local expansion. Similar arguments apply at higher orders in the loop expansion and, therefore, $\mathcal{S}_L(\varphi)$ has a local expansion to any finite order. Moreover, we have a kind of power counting property

$$\varphi^2(\tau) \sim t \sim -\partial_\tau^2 \sim \left(uL^{4-d}\right)^{2/3}. \tag{37.61}$$

In addition, each new loop brings a factor uL^{4-d}. Therefore, at loop-order l a term with $2k$ powers of φ, m powers of t 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 $\int \mathrm{d}\tau\, \varphi^2(\tau)$ which vanishes like $L^{-2(d-4)/3}$.

Remark. Here again, although the result (37.57) has been proven only in the scaling regime $tL^{2(d-1)/3}$ finite, it has the correct large argument behaviour. For $t > 0$,

$$X(z) \underset{z\to+\infty}{\sim} z^{-1/2} \ \Rightarrow\ \xi_L \sim t^{-1/2}.$$

For $t < 0$, the behaviour changes drastically depending whether the symmetry is continuous or discrete as we discuss in detail in Sections 37.5.2 and A37.1.2. For $N > 2$, we observe that the lowest eigenvalues of the hamiltonian (37.56) can, for $z \to -\infty$, be obtained by approximating $\widehat{H}$ by the angular moment part, fixing the radial coordinate at $|\mathbf{q}| = \sqrt{-6z}$. The corresponding eigenvalues are then $\ell(\ell+N-2)/(-12z)$. It follows

$$X(z) \underset{z\to-\infty}{\sim} -\frac{12z}{N-1} \ \Rightarrow\ \frac{\xi_L}{L} \sim -\frac{12}{u}\frac{t}{N-1}L^{d-2} \propto \frac{L^{d-2}}{\xi_\infty^2(t)},$$

a result which can be compared with (37.93).

For $N = 1$ instead, instantons are responsible for the splitting between the two lowest-lying states (Chapter 41). A WKB analysis yields

$$\ln X(z) \sim 2(-2z)^{3/2} \ \Rightarrow\ \ln \xi_L \sim 2(-2t)^{3/2}\frac{L^{d-1}}{u} = \frac{2}{u}\frac{L^{d-1}}{\xi_\infty^3(t)}, \tag{37.62}$$

a result which can be compared with (A37.13).

37.4.2 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 (37.57) and the RG scaling (37.11), we find

$$\xi_L/L = (u^*)^{-1/3} X[tL^{1/\nu}/(u^*)^{2/3}]. \tag{37.63}$$

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 above transforms into an argument about the powers of ε. An interaction term generated at loop order l with m powers of t, $2k$ 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 $\int \mathrm{d}\tau\, \varphi^2(\tau)$ 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 $\dot\varphi^2$ is generated. The one-loop terms proportional to $\int \mathrm{d}\tau (\varphi^2(\tau))^2$ and $t\int \mathrm{d}\tau\, \varphi^2(\tau)$ are of order ε. The $\int \mathrm{d}\tau(\varphi^2(\tau))^3$ 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 37.3. We call $a_1 u^*$ the coefficients of $t\int \mathrm{d}\tau\, \varphi^2(\tau)$, $a_2 u^*$ of $\int \mathrm{d}\tau\, \varphi^2(\tau)$ and $a_3 u^{*2}$ of $\int \mathrm{d}\tau(\varphi^2(\tau))^2$ in the expansion of the expression (37.60). Adding these contributions to the effective action simply amounts to the substitution

$$t \mapsto t(1 + a_1 u^*) + a_2 u^* , \qquad u^* \mapsto u^*(1 + a_3 u^*).$$

After this substitution into equation (37.63) we find the finite size correlation length ξ_L at one loop:

$$\begin{aligned} \xi_L(t') &= L\left(u^* + a_3 u^{*2}\right)^{-1/3} X(t' L^{1/\nu} + b), \\ b &= a_2 u^* \left(u^* + a_3 u^{*2}\right)^{-2/3} = a_2 \left(\frac{48\pi^2\varepsilon}{N+8}\right)^{1/3} + O\left(\varepsilon^{4/3}\right), \end{aligned} \tag{37.64}$$

where

$$t' = t\left[1 + (a_1 - \tfrac{2}{3}a_3)u^*\right]/u^{*2/3}$$

is a renormalized temperature.

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. We thus need only the coefficient $\tilde a_2$ of $u^* \int \mathrm{d}\tau\, \varphi^2/2$.

The coefficient of φ^2. The coefficient $\tilde a_2$ of $u^* \int \mathrm{d}\tau\, \varphi^2/2$ in the expansion of the expression (37.60) is

$$\tilde a_2 = \frac{N+2}{6} \int \frac{\mathrm{d}\omega}{2\pi} \sum_{\mathbf{k}\neq\mathbf{0}} \frac{1}{\omega^2 + (2\pi\mathbf{k})^2}. \tag{37.65}$$

We can integrate over ω. We then use the formal identity

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{1}{|2\pi\mathbf{k}|} = \frac{1}{4\pi}\int_0^{+\infty} \frac{\mathrm{d}s}{\sqrt{s}} \left(\vartheta_0^{d-1}(s) - 1\right), \tag{37.66}$$

in which the function $\vartheta_0(s)$ has been defined by equation (37.42). The integral has a divergence for s small which, as we have seen in Section 37.3.2, is cancelled by the critical temperature shift. We then find

$$a_2 = \frac{N+2}{24\pi} \int_0^{+\infty} \frac{\mathrm{d}s}{\sqrt{s}} \left[\vartheta_0^3(s) - 1 - s^{-3/2}\right] + O(\varepsilon). \tag{37.67}$$

The finite size correlation length at T_c. Substituting the value of a_2 into (37.64), we obtain in particular the finite size correlation length at T_c at one-loop order:

$$\frac{\xi_L}{L} = \left(\frac{48\pi^2\varepsilon}{N+8}\right)^{-1/3} X \left[K\pi^{-1/3}\frac{N+2}{12}\left(\frac{6\varepsilon}{N+8}\right)^{1/3}\right](1+O(\varepsilon)) \tag{37.68}$$

with

$$K = \int_0^{+\infty} \frac{\mathrm{d}s}{\sqrt{s}} \left[\vartheta_0^3(s) - 1 - s^{-3/2}\right] = -2.8372974\ldots . \tag{37.69}$$

37.5 Finite Size Effects in the Non-Linear σ-Model

In models with continuous symmetries, below T_c, 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 for all temperatures below T_c. We thus have to separate the zero mode and treat it non-perturbatively. We consider the action of the non-linear σ-model (see Chapter 31):

$$\mathcal{S}(\boldsymbol{\phi}) = \frac{\Lambda^{d-2}}{2t} \int \mathrm{d}^d x \left[\partial_\mu \boldsymbol{\phi}(x)\right]^2 \tag{37.70}$$

with

$$\boldsymbol{\phi}^2(x) = 1 . \tag{37.71}$$

The non-linear σ-model can be used to describe the IR behaviour at fixed temperature below T_c and fixed dimension $2 < d < 4$ in which case a cut-off Λ is required, or the whole physical range up to T_c but only within the framework of the $\varepsilon = d-2$ expansion.

We now examine separately the two cases of the hypercubic and cylindrical geometries.

37.5.1 The hypercubic geometry

The zero mode can here 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 spin distribution (equation (37.24)):

$$m_2 = L^{-d} \int \mathrm{d}^d x \left\langle \boldsymbol{\phi}(x) \cdot \boldsymbol{\phi}(0) \right\rangle . \tag{37.72}$$

In terms of the functions (31.33,31.34), m_2 has the scaling form

$$m_2 = M_0^2(t)\mu_2\bigl(L/\xi(t)\bigr). \tag{37.73}$$

For $d > 2$, m_2 is finite at the critical temperature t_c and thus

$$\mu_2(z) \underset{z\to 0}{\propto} z^{-(d-2+\eta)}.$$

To calculate m_2, we parametrize the field $\boldsymbol{\phi}(x)$ in terms of $\boldsymbol{\pi}(x)$:

$$\boldsymbol{\phi}(x) = \left\{\sqrt{1-\boldsymbol{\pi}^2(x)}\,, \quad \boldsymbol{\pi}(x)\right\}. \tag{37.74}$$

It is convenient to set

$$L^{-2d}\int \mathrm{d}^d x\, \mathrm{d}^d y\, \boldsymbol{\phi}(x)\cdot\boldsymbol{\phi}(y) = 1 - 2\mathcal{O}\,. \tag{37.75}$$

Expanding in powers of the field π one obtains

$$\mathcal{O} = \tfrac{1}{4}L^{-2d}\int \mathrm{d}^d x\, \mathrm{d}^d y\, \left[\boldsymbol{\pi}(x)-\boldsymbol{\pi}(y)\right]^2 + O(\pi^4). \tag{37.76}$$

The π-propagator Δ_{ij} is

$$\Delta_{ij}(x) = \delta_{ij}\Delta(x), \qquad \Delta(x) \equiv t\Lambda^{2-d}L^{-d}\sum_{\mathbf{k}\neq\mathbf{0}} \frac{\mathrm{e}^{i2\pi\mathbf{k}\cdot\mathbf{x}}}{(2\pi\mathbf{k})^2}. \tag{37.77}$$

The moment m_2 at one-loop order follows

$$m_2 = 1 - 2\left\langle\mathcal{O}\right\rangle = 1 + t(N-1)(\Lambda L)^{2-d}\int \mathrm{d}^d x \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\left(\mathrm{e}^{i2\pi\mathbf{k}\cdot\mathbf{x}/L}-1\right)}{(2\pi\mathbf{k})^2} + O\left(t^2\right). \tag{37.78}$$

After integration over x the expression becomes

$$m_2 = 1 - t(N-1)(\Lambda L)^{2-d}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{1}{(2\pi\mathbf{k})^2} + O\left(t^2\right), \tag{37.79}$$

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:

$$M_0 = 1 - \frac{t\Lambda^{2-d}}{2(2\pi)^d}\int\frac{\mathrm{d}^d p}{p^2} + O\left(t^2\right) = 1 - \frac{t\Lambda^{2-d}}{8\pi}\int_0^\infty \mathrm{d}s\, s^{-d/2}. \tag{37.80}$$

As usual, we transform the sum over $\mathbf{k}$ into an integral involving the function (37.42). Taking into account equation (37.80), we find for the scaling function (37.73),

$$\mu_2 = 1 - \frac{t(N-1)}{4\pi}(\Lambda L)^{2-d}\int_0^\infty \mathrm{d}s\left(\vartheta_0^d(s) - 1 - s^{-d/2}\right) + O\left(t^2\right). \tag{37.81}$$

RG arguments tell us that we 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 (37.20))

$$t_L = t(\Lambda L)^{2-d} + O\left(t^2\right).$$

We then obtain

$$\mu_2 = 1 - \frac{t_L(N-1)}{4\pi} \int_0^\infty \mathrm{d}s \left(\vartheta_0^d(s) - 1 - s^{-d/2}\right) + \cdots ,$$

expression valid for $2 < d < 4$ and $t < t_c$ fixed. Since then t_L goes to zero for L large, this expression shows how μ_2 approaches 1. Using the estimate (37.18), we find

$$\mu_2 - 1 \sim -\frac{N-1}{4\pi} \int_0^\infty \mathrm{d}s \left(\vartheta_0^d(s) - 1 - s^{-d/2}\right) \left(\frac{\xi(t)}{L}\right)^{d-2}. \qquad (37.82)$$

To generate universal quantities, we again consider the ratios (37.34) $\mathcal{R}_\sigma = m_\sigma/(m_2)^{\sigma/2}$. With the definition (37.75)

$$\left(L^{-d}\left|\int \mathrm{d}^d x\, \phi(x)\right|\right)^\sigma = 1 - \sigma\mathcal{O} + \frac{\sigma(\sigma-2)}{2}\mathcal{O}^2 + \cdots .$$

Therefore,

$$\mathcal{R}_\sigma - 1 \sim \tfrac{1}{2}\sigma(\sigma-2)\left[\left\langle \mathcal{O}^2\right\rangle - \left(\langle\mathcal{O}\rangle\right)^2\right].$$

Using equations (37.76,37.77), we obtain

$$\begin{aligned}\mathcal{R}_\sigma &= 1 + \frac{1}{8}\sigma(\sigma-2)(N-1)L^{-2d}\int \mathrm{d}^d x\, \mathrm{d}^d y\, [\Delta(x) - \Delta(y)]^2 , \\ &= 1 + \frac{1}{4}\sigma(\sigma-2)(N-1)\frac{t^2(\Lambda L)^{4-2d}}{16\pi^2}\int_0^\infty \mathrm{d}s\, s\left[\vartheta_0^d(s) - 1\right] + O\left(t^3\right).\end{aligned}$$

We then use RG arguments, set $\Lambda L = 1$ and replace t by t_L (equation (37.20)):

$$\mathcal{R}_\sigma = 1 + \sigma(\sigma-2)\frac{N-1}{4}\frac{t_L^2}{16\pi^2}\int_0^\infty \mathrm{d}s\, s\left[\vartheta_0^d(s) - 1\right]. \qquad (37.83)$$

For $d > 2$ and $t < t_c$ fixed, t_L goes to zero for L large. Using the estimate (37.18), we thus find how $\mathcal{R}_\sigma$ goes to 1:

$$\mathcal{R}_\sigma = 1 + \sigma(\sigma-2)\frac{N-1}{4}\frac{1}{16\pi^2}\int_0^\infty \mathrm{d}s\, s\left[\vartheta_0^d(s) - 1\right]\left(\frac{\xi(t)}{L}\right)^{2d-4}. \qquad (37.84)$$

The neighbourhood of the critical temperature can be studied only within the $\varepsilon = d-2$ expansion. Setting $t_L = t_c$, we obtain the universal ratio $\mathcal{R}_\sigma$ at order ε^2:

$$\mathcal{R}_\sigma = 1 + \varepsilon^2\sigma(\sigma-2)\frac{N-1}{16(N-2)^2}\int_0^\infty \mathrm{d}s\, s\left[\vartheta_0^2(s) - 1\right]. \qquad (37.85)$$

Induced magnetization in a small field. In a small magnetic field h, that is, such that hL^d is small, we can easily 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

$$M(h,t,L) = M_0(t) m\left[\frac{hM_0(t)}{t}\xi^d(t), \frac{L}{\xi(t)}\right]. \qquad (37.86)$$

At leading order, the partition function in a field is given by

$$\begin{aligned}\mathcal{Z}(h,t,L) &= \int \mathrm{d}^N\varphi\, \delta\left(\varphi^2-1\right)\exp\left(L^d\,\mathbf{h}\cdot\varphi/t\right)\\ &= \frac{1}{\pi}\int_0^\pi \mathrm{d}\theta\,(\sin\theta)^{N-2}\exp\left(\dot{L}^d h\cos\theta/t\right). \end{aligned} \qquad (37.87)$$

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

$$v = hM_0(t)L^d/t\,, \qquad (37.88)$$

and is given by

$$M(h,t,L) = M_0(t)\frac{\mathrm{d}}{\mathrm{d}v}\ln\int_0^\pi \mathrm{d}\theta(\sin\theta)^{N-2}\,\mathrm{e}^{v\cos\theta}\,. \qquad (37.89)$$

37.5.2 The cylindrical geometry

We now consider, as an illustration, the calculation of 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 3.4:

$$\mathcal{S}_L(\varphi) = \frac{\Lambda^{d-2}L^{d-1}}{2t}\int \mathrm{d}\tau\,\dot{\varphi}^2(\tau)\,. \qquad (37.90)$$

From equation (3.46), we infer the correlation length at leading order

$$\xi_L(t) = \frac{2}{N-1}\frac{\Lambda^{d-2}L^{d-1}}{t}. \qquad (37.91)$$

Combining this result with the RG arguments of Section 37.1.2, we can rewrite this equation in terms of the effective coupling at scale L (equations (37.15,37.20)):

$$\frac{\xi_L(t)}{L} \sim \frac{2}{(N-1)t_L}\,. \qquad (37.92)$$

For $d>2$ in the ordered phase t_L goes to zero and we find the scaling form

$$\frac{\xi_L(t)}{L} \underset{\text{for } 2<d<4}{\sim} \frac{2}{N-1}\left(\frac{L}{\xi_\infty(t)}\right)^{d-2}. \qquad (37.93)$$

We immediately 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 A37.1.2).

The critical temperature is a RG fixed point and thus $t_L = t_c$. We then obtain the universal ratio ξ_L/L at leading order in $\varepsilon = d-2$:

$$\frac{\xi_L(t_c)}{L} \underset{d\to 2}{\sim} \frac{N-2}{(N-1)\pi\varepsilon}.$$

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 (37.92), we evaluate the effective action for $\varphi(\tau)$ obtained by integrating over non-zero modes at one-loop order. We define $\varphi(\tau)$ as the unit vector along the direction of the average of $\int \mathrm{d}^{d-1}x\, \boldsymbol{\phi}(\tau,x)$. We then parametrize $\boldsymbol{\phi}(\tau,x)$ by a rotation $R(\tau)$ acting on a field $\boldsymbol{\phi}'(\tau,x)$ whose average $\int \mathrm{d}^{d-1}x\, \boldsymbol{\phi}'(\tau,x)$ points to a fixed direction. To simplify, instead of calculating for generic fields $\varphi(\tau)$, it is convenient to restrict φ to lie in a two-dimensional plane:

$$\varphi(\tau) = \{\cos(\alpha(\tau), \sin(\alpha(\tau)), \mathbf{0}\}.$$

The action (37.90) then reduces to

$$\mathcal{S}_L(\alpha) = \frac{\Lambda^{d-2}L^{d-1}}{2t} \int \mathrm{d}\tau\, \dot\alpha^2(\tau)\,. \tag{37.94}$$

At order $\dot\alpha^2$, the mapping between the actions (37.90) and (37.94) is unambiguous.

We then parametrize the field $\boldsymbol{\phi}(\tau,x)$ as

$$\boldsymbol{\phi}(\tau,x) = \begin{cases} \cos(\alpha(\tau))\sigma_1(\tau,x) - \sin(\alpha(\tau))\sigma_2(\tau,x), \\ \sin(\alpha(\tau))\sigma_1(\tau,x) + \cos(\alpha(\tau))\ \sigma_2(\tau,x), \\ \boldsymbol{\pi}(\tau,x), \end{cases} \tag{37.95}$$

in which the field $\boldsymbol{\pi}(\tau,x)$ here has only $N-2$ components. The fields $\boldsymbol{\pi}$ and σ_2 have no zero mode. Since the transformation (37.95) is a rotation, the three fields $\sigma_1, \sigma_2, \boldsymbol{\pi}$ still satisfy the constraint:

$$\sigma_1^2 + \sigma_2^2 + \boldsymbol{\pi}^2 = 1\,, \tag{37.96}$$

and the integration measure in the functional integral is left invariant.

The action $\mathcal{S}(\boldsymbol{\phi})$ in the new fields reads

$$\begin{aligned}\mathcal{S}(\sigma_1,\sigma_2,\boldsymbol{\pi}) = \frac{\Lambda^\varepsilon}{2t} \int \mathrm{d}\tau\, \mathrm{d}^{d-1}x \Big[& \dot\alpha^2\,(\sigma_1^2+\sigma_2^2) + \dot\sigma_1{}^2 + \dot\sigma_2{}^2 + \dot{\boldsymbol{\pi}}^2 \\ & + 2\dot\alpha(\dot\sigma_2\sigma_1 - \dot\sigma_1\sigma_2) + (\partial_i\sigma_1)^2 + (\partial_i\sigma_2)^2 + (\partial_i\boldsymbol{\pi})^2 \Big].\end{aligned} \tag{37.97}$$

To calculate the effects of non-zero modes, we eliminate the field σ_1 using the constraint (37.96):

$$\sigma_1 = \left(1 - \sigma_2^2 - \boldsymbol{\pi}^2\right)^{1/2},$$

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

$$\mathcal{S}_2(\sigma_2,\boldsymbol{\pi}) = \frac{\Lambda^{\varepsilon}}{2t}L^{d-1}\int \mathrm{d}\tau\, \dot{\alpha}^2 + \frac{\Lambda^{\varepsilon}}{2t}\int \mathrm{d}\tau\, \mathrm{d}^{d-1}x \left[-\dot{\alpha}^2\boldsymbol{\pi}^2 + \dot{\sigma_2}^2 + \dot{\boldsymbol{\pi}}^2 + 2\dot{\alpha}\dot{\sigma}_2 + (\partial_i\sigma_2)^2 + (\partial_i\boldsymbol{\pi})^2\right]. \quad (37.98)$$

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 functional integral. The integral over $\boldsymbol{\pi}$ yields a determinant to the power $(N-2)/2$. Hence, the result at one-loop order is

$$\mathcal{S}_L(\alpha) = \frac{\Lambda^{\varepsilon}}{2t}L^{d-1}\int \mathrm{d}\tau\, \dot{\alpha}^2 + \frac{1}{2}(N-2)\operatorname{tr}\ln\left[\left(-\dot{\alpha}^2 - \partial_\tau^2 - \partial_i^2\right)\left(-\partial_\tau^2 - \partial_i^2\right)^{-1}\right]. \quad (37.99)$$

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

$$\mathcal{S}_L(\alpha) = \mathfrak{S}(t,\Lambda,L)\int \mathrm{d}\tau\, \dot{\alpha}^2,$$
$$\mathfrak{S}(t,\Lambda,L) = \frac{\Lambda^{\varepsilon}}{2t}L^{d-1} - \frac{N-2}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{1}{2\pi}\int\frac{\mathrm{d}\omega}{\omega^2+(2\pi\mathbf{k}/L)^2}. \quad (37.100)$$

The sum has to be understood with a large momentum cut-off. The cut-off dependence can be is eliminated by subtracting the infinite size limit. We use the identity ($d>2$)

$$\sum_{\mathbf{k}\neq\mathbf{0}}\frac{1}{2\pi}\int\frac{\mathrm{d}\omega}{\omega^2+(2\pi\mathbf{k}/L)^2} - \frac{L^{d-1}}{(2\pi)^d}\int^{\Lambda}\frac{\mathrm{d}^d\mathbf{p}}{\mathbf{p}^2} = \int_0^{\infty}\mathrm{d}s\frac{\mathrm{d}\omega\, \mathrm{e}^{-s\omega^2}}{2\pi}\left[\sum_{\mathbf{k}\neq\mathbf{0}}\mathrm{e}^{-4\pi^2\mathbf{k}^2 s/L^2} - \frac{L^{d-1}}{(2\pi)^{d-1}}\int \mathrm{d}^{d-1}p\, \mathrm{e}^{-sp^2}\right]$$
$$= \frac{L}{4\pi}\int_0^{\infty}\frac{\mathrm{d}s}{\sqrt{s}}\left[\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2}\right],$$

where the change of variables $4\pi s/L^2 \mapsto s$ and the function (37.42) have been used. For dimensional reasons

$$\int^{\Lambda}\frac{\mathrm{d}^d\mathbf{p}}{\mathbf{p}^2} = -\frac{\tilde{\beta}_2(d)}{d-2}\Lambda^{d-2},$$

where $\tilde{\beta}_2$ is finite for $d\to 2$ but depends on the precise regularization.

Comparing with equation (37.91), we conclude

$$\frac{\xi_L}{L} = \frac{2}{N-1}\left[(L\Lambda)^{d-2}\left(\frac{1}{t}+\frac{\tilde{\beta}_2(d)}{d-2}\right) - \frac{N-2}{4\pi}\int_0^{\infty}\frac{\mathrm{d}s}{\sqrt{s}}\left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2}\right)\right] + O(t).$$

Comparing with expression (37.20), we verify $\tilde{\beta}_2 = \beta_2$ and we rewrite this expression in terms of the size-dependent coupling constant t_L:

$$\frac{\xi_L(t)}{L} = \frac{2}{N-1}\left[\frac{1}{t_L} + \frac{\beta_2(d)}{d-2} - \frac{N-2}{4\pi}\int_0^\infty \frac{\mathrm{d}s}{\sqrt{s}}\left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2}\right)\right] + O(t_L). \tag{37.101}$$

We now specialize this expression to various situations. For $d > 2$ and $t < t_c$ fixed or $d = 2$, $t \to 0$, 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 $\varepsilon = d - 2$. The subtracted integral diverges for s large for $d \to 2$:

$$\int_0^\infty \frac{\mathrm{d}s}{\sqrt{s}}\left(\vartheta_0^{d-1}(s) - 1 - s^{-(d-1)/2}\right) = -\frac{2}{\varepsilon} + \gamma - \ln 4\pi\,,$$

where γ is Euler's constant. The expression in the r.h.s. of (37.101) has, as expected, a limit for $d \to 2$. 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

$$\frac{1}{t_c} + \frac{\beta_2(d)}{d-2} = \frac{\beta_3(2)}{\beta_2(2)} + O(d-2),$$

a quantity which is scheme-independent. Using the expression (31.60), coming from the minimal subtraction scheme, we find $\beta_3/\beta_2 = 1/2\pi$ and thus the universal ratio

$$\frac{\xi_L(t_c)}{L} = \frac{N-2}{N-1}\frac{1}{\pi\varepsilon}\left[1 + \frac{\varepsilon}{N-2} - \frac{\varepsilon}{2}(\gamma - \ln 4\pi) + O\left(\varepsilon^2\right)\right]. \tag{37.102}$$

37.6 Finite Size Effects and Dynamics

We have discussed dynamics in Chapters 17 and 36 from the RG point of view. Previous considerations immediately 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 $L/\xi(T)$. For example, 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 a scaling form which generalizes equation (36.43):

$$\tau(T, L) = L^z f\left[(T - T_c)L^{1/\nu}\right]. \tag{37.103}$$

Physical quantities, for example, in a periodic hypercube, 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. We here give a few examples of calculations of correlation times in the simple case of a purely dissipative model without conservation laws, based on the static N-vector model.

Notation. Because in this section t denotes time, we denote by $\Delta T \propto T - T_c$ the deviation from the critical temperature.

37.6.1 The $(\varphi^2)^2$ theory

We consider the Langevin equation (36.9):

$$\dot{\varphi}(t,x) = -\frac{\Omega}{2}\left[\left(-\nabla^2 + r_c + \Delta T\right)\varphi(t,x) + \frac{u}{3!}\varphi\varphi^2\right] + \nu(t,x) \tag{37.104}$$

with gaussian white noise ν:

$$\langle \nu_i(t,x)\rangle = 0\,, \qquad \langle \nu_i(t,x)\nu_j(t',x')\rangle = \Omega\delta_{ij}\,\delta(t-t')\delta(x-x'). \tag{37.105}$$

We recall that to the Langevin equation (37.104) is associated a dynamic action $\mathcal{S}(\boldsymbol{\phi})$, which in terms of the superfield

$$\boldsymbol{\phi} = \boldsymbol{\varphi} + \theta\bar{\mathbf{c}} + \mathbf{c}\bar{\theta} + \theta\bar{\theta}\boldsymbol{\lambda}\,, \tag{37.106}$$

takes a supersymmetric form

$$\mathcal{S}(\boldsymbol{\phi}) = \int \mathrm{d}t\,\mathrm{d}\bar{\theta}\,\mathrm{d}\theta\left[\int \mathrm{d}^d x \frac{2}{\Omega_0}\bar{\mathrm{D}}\phi\mathrm{D}\phi + \mathcal{H}(\boldsymbol{\phi})\right], \tag{37.107}$$

where

$$\bar{\mathrm{D}} = \frac{\partial}{\partial\bar{\theta}}, \quad \mathrm{D} = \frac{\partial}{\partial\theta} - \bar{\theta}\frac{\partial}{\partial t}, \tag{37.108}$$

and $\mathcal{H}(\varphi)$ is the static euclidean action:

$$\mathcal{H}(\varphi) = \int \mathrm{d}^d x\left[\frac{1}{2}\left(\partial_\mu\varphi\right)^2 + \frac{1}{2}(r_c + \Delta T)\varphi^2 + \frac{u}{4!}\left(\varphi^2\right)^2\right]. \tag{37.109}$$

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 (37.107) to a space-independent field. We then recognize the action associated with a stochastic differential equation of a type studied in Chapter 4. The corresponding Fokker–Planck hamiltonian H_{FP}, in hermitian form, is (equation (4.41))

$$H_{\mathrm{FP}} = \frac{\Omega}{2L^d}\left[\mathbf{p}^2 + \tfrac{1}{4}\left(\nabla E(\mathbf{q})\right)^2 - \tfrac{1}{2}\nabla^2 E(\mathbf{q})\right], \tag{37.110}$$

with, at leading order

$$E(\mathbf{q}) = L^d\left(\frac{1}{2}\Delta T\mathbf{q}^2 + \frac{u}{4!}\left(\mathbf{q}^2\right)^2\right). \tag{37.111}$$

Rescaling $\mathbf{q}$:

$$\mathbf{q} \mapsto \mathbf{q}\left(L^d u\right)^{-1/4}, \tag{37.112}$$

we find that the eigenvalues E_i of H_{FP}, as functions of L and the deviation of the critical temperature ΔT, take the form

$$E_i = \Omega u^{1/2} L^{-d/2} e_i\bigl(L^{d/2}\Delta T u^{-1/2}\bigr). \tag{37.113}$$

The scaling for $d > 4$ of the relaxation time τ, which is the inverse of the difference of the two first eigenvalues, follows

$$\tau(T,L) = \Omega^{-1}u^{-1/2}L^{d/2} f\bigl(L^{d/2}\Delta T u^{-1/2}\bigr). \tag{37.114}$$

As for the finite size correlation length (see equation (37.57)), the naive extrapolation of the scaling form valid for $d < 4$ is incorrect. An analysis similar to the one performed in the static case shows that loop corrections do not modify the scaling form for $d > 4$.

As for the correlation length, let us examine the behaviour of τ when $\Delta T L^{d/2}$ is large. For $\Delta T > 0$, we find the expected limit $\tau^{-1} \to \Omega \Delta T$. For $\Delta T < 0$ and $N > 1$, we obtain

$$\tau \sim -\frac{12}{(N-1)u} \Omega^{-1} L^d \Delta T .$$

Note, however, that for $N = 1$ the behaviour found in this way is incorrect, as we discuss in Section A37.1.3.

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 minimal subtraction scheme, using the supersymmetric formulation of Appendix A36.1. Following the lines of Section 37.4, we perform a rescaling in the action, equivalent to (37.112) (we again use the notation $\boldsymbol{\theta} = (\bar{\theta}, \theta)$):

$$\boldsymbol{\phi} \mapsto \boldsymbol{\phi} \left(L^d u\right)^{-1/4}, \quad t \mapsto t \left(L^d/u\right)^{1/2}, \quad \boldsymbol{\theta} \mapsto \boldsymbol{\theta} \left(L^d/u\right)^{1/4}, \tag{37.115}$$

and count the powers of ε. We again verify that only the coefficients of $\boldsymbol{\phi}^2$, $\Delta T \boldsymbol{\phi}^2$ and $(\boldsymbol{\phi}^2)^2$ are relevant at one-loop order. In addition, the scaling of the time and Grassmann variables shows that we need only the contributions proportional to $(1, \Delta T) \int \mathrm{d}t\, \mathrm{d}\bar{\theta}\, \mathrm{d}\theta\, \boldsymbol{\phi}^2$ and $\int \mathrm{d}t\, \mathrm{d}\bar{\theta}\, \mathrm{d}\theta (\boldsymbol{\phi}^2)^2$. It is easy to verify that the calculation then becomes identical to the static calculation in a finite volume. The main result is given by equation (37.48). The relaxation time, at the IR fixed point, in the one-loop approximation, is then

$$\tau(\Delta T, L) = \Omega'^{-1} L^z f(L^{1/\nu} \Delta T' + b), \tag{37.116}$$

in which $f(z)$ is the function implicitly defined by equation (37.114), and Ω'^{-1} and $\Delta T'$ are a renormalized time scale and a renormalized deviation from the critical temperature, respectively.

37.6.2 The non-linear σ-model

With the non-linear σ-model, we can calculate, for example, the relaxation time at fixed temperature below T_c, close to T_c in a $d-2$ expansion or in two dimensions. The form of the dynamic action has been given in Appendix A36.2:

$$\mathcal{S}(\boldsymbol{\phi}) = \frac{\Lambda^\varepsilon}{g} \int \mathrm{d}\bar{\theta}\, \mathrm{d}\theta\, \mathrm{d}t \int \mathrm{d}^d x \left[\frac{2}{\Omega} \bar{\mathrm{D}} \boldsymbol{\phi} \cdot \mathrm{D} \boldsymbol{\phi} + \frac{1}{2} \left(\partial_\mu \boldsymbol{\phi}\right)^2\right] \tag{37.117}$$

with

$$\boldsymbol{\phi}^2 = 1 . \tag{37.118}$$

To avoid confusions, we here denote the coupling constant or temperature by g.

The RG equations of this model have been discussed in Section 36.2. The relaxation time, defined in the infinite volume by equation (36.56), satisfies the RG equation:

$$\left(\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} + \eta_\omega(g) \Omega \frac{\partial}{\partial \Omega}\right) \tau(\Lambda, L, g, \Omega) = 0 . \tag{37.119}$$

The finite size relaxation time τ satisfies the dimensional relation

$$\tau(\Lambda, L, g, \Omega) = \Omega^{-1}\Lambda^{-2}\tau(1, \Lambda L, g, 1) . \tag{37.120}$$

The RG equation (37.119) can thus be rewritten as

$$\left(L\frac{\partial}{\partial L} + \beta(g)\frac{\partial}{\partial g} - 2 - \eta_\omega(g)\right)\tau(\Lambda, L, g, \Omega) = 0 . \tag{37.121}$$

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

$$\tau(L, g) = L^2\zeta(L, g)\tau(1, g_L) \tag{37.122}$$

with the notation $g_L \equiv g(1/L)$ (see Section 37.1.2), and

$$\ln L = \int_{g_L}^{g} \frac{\mathrm{d}g'}{\beta(g')}, \tag{37.123}$$

$$\zeta(L, g) = \exp\left[\int_{g_L}^{g} \frac{\eta_\omega(g')}{\beta(g')}\mathrm{d}g'\right] . \tag{37.124}$$

Equation (37.122) can also be rewritten as

$$\tau(L, g) = L^2\zeta(L, g)\mathcal{T}\big(L/\xi(g)\big), \tag{37.125}$$

in which $\xi(g)$ is the infinite volume correlation length. As we have already shown in Section 37.1, at fixed coupling $g < g_c$, the effective temperature g_L goes to zero. Therefore, τ can be derived from perturbation theory.

Note finally that the function $\eta_\omega(g)$ begins at order g^2 (see Section A36.2). Since we calculate only at one-loop order, the function $\zeta(L, g)$ can be replaced in what follows by a constant renormalization factor

$$\zeta(g) = \exp\left[\int_{0}^{g} \frac{\eta_\omega(g')}{\beta(g')}\mathrm{d}g'\right] . \tag{37.126}$$

To calculate τ, we use the method explained in Section 37.5.2. We parametrize the field $\boldsymbol{\phi}(t, \boldsymbol{\theta}, x)$ as

$$\boldsymbol{\phi}(t, \boldsymbol{\theta}, x) = \begin{cases} \cos\alpha(t, \boldsymbol{\theta})\sigma_1(t, \boldsymbol{\theta}, x) - \sin\alpha(t, \boldsymbol{\theta})\sigma_2(t, \boldsymbol{\theta}, x), \\ \sin\alpha(t, \boldsymbol{\theta})\sigma_1(t, \boldsymbol{\theta}, x) + \cos\alpha(t, \boldsymbol{\theta})\sigma_2(t, \boldsymbol{\theta}, x), \\ \boldsymbol{\pi}(t, \boldsymbol{\theta}, x) , \end{cases} \tag{37.127}$$

where σ_2 and $\boldsymbol{\pi}$ have no zero-mode, and determine the one-loop effective action for the field $\alpha(t, \boldsymbol{\theta})$. 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

$$\begin{aligned} \mathcal{S}(\alpha, \boldsymbol{\pi}) = {} & \frac{2\Lambda^\varepsilon}{\Omega g}\int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t\,\bar{\mathrm{D}}\alpha\mathrm{D}\alpha\left[L^d - \int \mathrm{d}^d x\,\boldsymbol{\pi}^2(t, \boldsymbol{\theta}, x)\right] \\ & + \frac{\Lambda^\varepsilon}{g}\int \mathrm{d}\bar{\theta}\,\mathrm{d}\theta\,\mathrm{d}t\int \mathrm{d}^d x\left[\frac{2}{\Omega}\bar{\mathrm{D}}\boldsymbol{\pi}\cdot\mathrm{D}\boldsymbol{\pi} + \frac{1}{2}\left(\partial_\mu\boldsymbol{\pi}\right)^2\right] . \end{aligned} \tag{37.128}$$

Neglecting the non-zero modes, and using renormalization group, we immediately obtain the form of the relaxation time, at leading order for large L:

$$\tau(L,g) \sim \Omega'^{-1}(g)\,\frac{L^2}{g_L} \quad \text{with } \Omega'(g) = \Omega\zeta^{-1}(g)\frac{N-1}{2}, \tag{37.129}$$

where $\zeta(g)$ is defined by equation (37.126).

The integral over π expanded for α small yields simply a renormalization of the leading order α action:

$$\frac{L^d}{2\Omega g} \mapsto \frac{L^2}{2\Omega}\left(\frac{L^{d-2}}{g} - \frac{N-2}{4\pi^2}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{1}{\mathbf{k}^2}\right). \tag{37.130}$$

The relaxation time follows

$$\frac{\Omega\tau(g,L)}{L^2} = \frac{2}{N-1}\left(\frac{L^{d-2}}{g} - \frac{N-2}{4\pi^2}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{1}{\mathbf{k}^2}\right). \tag{37.131}$$

The sum has to be understood with a cut-off. As we have explained in Section 37.5.2, if we subtract to the sum its infinite size limit we obtain a finite result (equation (37.45)). We then introduce the size-dependent temperature g_L (equation (37.20) with $t \mapsto g$ and $\Lambda = 1$) and find

$$\frac{\tau(g,L)}{L^2} = \Omega'^{-1}(g)\left[\frac{1}{g_L} + \frac{\beta_2(d)}{d-2} - \frac{N-2}{4\pi}\int_0^\infty \mathrm{d}s\left(\vartheta_0^d(s) - 1 - s^{-d/2}\right)\right] \tag{37.132}$$

with $\Omega'(g)$ defined in (37.129). We have thus obtained the first correction to the leading term for $g < g_c$. Note again that this expression has a finite limit when $d \to 2$. Indeed,

$$\int_0^\infty \mathrm{d}s\left(\vartheta_0^d(s) - 1 - s^{-d/2}\right) = -\frac{2}{d-2} + \int_0^\infty \mathrm{d}s\left(\vartheta_0^2(s) - 1 - \theta(1-s)/s\right). \tag{37.133}$$

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.

Bibliographical Notes

Early articles and reviews about finite size effects are

M.E. Fisher in *Critical Phenomena, Proceedings of the International School of Physics Enrico Fermi*, Varenna 1971, M.S. Green ed. (Academic Press, New York 1972); M.E. Fisher and M.N. Barber, *Phys. Rev. Lett.* 28 (1972) 1516; Y. Imry and D. Bergman, *Phys. Rev.* A3 (1971) 1416.

See also

M.N. Barber in *Phase Transitions and Critical Phenomena* vol. 8, C. Domb and J. Lebowitz eds. (Academic Press, New York 1983),

and the papers collected in

Finite Size Scaling, J.L. Cardy ed. (North-Holland, Amsterdam 1988).

The field theoretical analysis has been performed in

E. Brézin, *J. Physique (Paris)* 43 (1982) 15; E. Brézin and J. Zinn-Justin, *Nucl. Phys.* B257 [FS14] (1985) 867; J. Rudnick, H. Guo and D. Jasnow, *J. Stat. Phys.* 41 (1985) 353; M. Lüscher, *Phys. Lett.* 118B (1982) 391; *Nucl. Phys.* B219 (1983) 233.
The correlation length is calculated with twisted boundary conditions in
E.G. Floratos and D. Petscher, *Nucl. Phys.* B252 (1985) 689.
A RG analysis is also given in
V. Privman and M.E. Fisher, *J. Stat. Phys.* 33 (1983) 385.
The almost planar interface model is studied in
D.J. Wallace and R.K.P. Zia, *Phys. Rev. Lett.* 43 (1979) 808; D.J. Wallace in *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, R. Stora and J.-B. Zuber eds. (North-Holland, Amsterdam 1984).
Finite size scaling laws in Critical Dynamics have been postulated in
M. Suzuki, *Prog. Theor. Phys.* 58 (1977) 1142.
A systematic RG analysis is given in
J.C. Niel and J. Zinn-Justin, *Nucl. Phys.* B280 [FS18] (1987) 335; Y.Y. Goldschmidt, *Nucl. Phys.* B280[FS18] (1987) 340.

APPENDIX A37

A37.1 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. We now consider, for completeness, 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.

A37.1.1 Finite volume

We here consider first order phase transitions in which the order parameter jumps from one constant value to another. We restrict ourselves to the calculation of homogeneous quantities like the magnetization. Then, the basic function we need is the free energy $-\mathfrak{S}(\varphi, L)$ at fixed field average

$$\exp\left[-\mathfrak{S}(\varphi, L)\right] = \int \left[\mathrm{d}\phi(x)\right] \delta\left(\varphi - L^{-d}\int \mathrm{d}^d x\, \phi(x)\right) \exp\left[-\mathcal{S}(\phi)\right], \qquad (A37.1)$$

a quantity whose physical meaning has already been discussed in Section 7.10. In the case of first order transitions, fluctuations are not critical because the correlation length remains finite at the transition. Therefore, the integration over the zero-momentum component does not solve a non-existent zero-mode problem, but instead restores the symmetry.

The free energy, in the presence of a constant magnetic field h, is given by

$$\mathrm{e}^{W(h)} = \int \mathrm{d}\varphi \, \exp\left[-\mathfrak{S}(\varphi, L) + L^d \beta h \varphi\right]. \qquad (A37.2)$$

In a translation invariant finite system, $\mathfrak{S}$ at large size L behaves like

$$\mathfrak{S}(\varphi, L) \sim L^d \Sigma(\varphi), \qquad (A37.3)$$

where $\Sigma(\varphi)$ is independent of L.

For L large, the integral can thus be calculated by steepest descent. In the case of a unique saddle point one finds

$$W(h)/L^d = -\Sigma(\varphi) + \beta h \varphi \qquad (A37.4)$$

with

$$\Sigma'(\varphi) = \beta h .$$

Note that $\mathfrak{S}(\varphi, L)$ is such that the corrections for L large to ($A37.4$) are exponentially small in L because the infinite size correlation length remains finite.

Degenerate minima. When instead 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 call $\pm M_0$ the two minima of $\Sigma(\varphi)$ (the generalization to any discrete set of minima is straightforward):

$$\Sigma(\varphi) = \Sigma(M_0) + \tfrac{1}{2}\Sigma''(M_0)\,(\varphi - M_0)^2 + O\,(\varphi - M_0)^3 . \qquad (A37.5)$$

The free energy is now the sum of the two saddle point contributions:

$$W(h) - W(0) = \ln\cosh(\beta h L^d M_0) + \tfrac{1}{2}\beta^2 h^2 L^d/\Sigma''(M_0). \qquad (A37.6)$$

The magnetization M and the zero-field susceptibility χ are then given by

$$M = \langle\varphi\rangle = (\beta L^d)^{-1} W'(h) = M_0 \tanh(\beta h L^d M_0) + \beta h/\Sigma''(M_0), \qquad (A37.7)$$
$$\chi = \beta/\Sigma''(M_0), \qquad (A37.8)$$

where χ is here defined as $\chi = \langle\varphi^2\rangle - (\langle|\varphi|\rangle)^2$.

In equation ($A37.7$), for hL^d finite, the second term is negligible, and the finite volume magnetization takes the universal form:

$$M = M_0 \tanh(\beta h L^d M_0). \qquad (A37.9)$$

We note that 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 $1-\rho$ instead $\varphi = M_0$. 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

$$\tfrac{1}{2}L^d\Sigma''(M_0)(\varphi - M_0)^2 = 2\Sigma''(M_0)M_0^2\rho^2 L^d > \sigma(T)(\rho^{1/d}L)^{d-1},$$

and, therefore, $\rho >$ const. $(\xi(T)/L)^{d/(d+1)}$, where the correlation length $\xi(T)$ is implied by dimensional considerations. For larger values of ρ, $\mathfrak{S}(\varphi, L)$ is at leading order of the form

$$\mathfrak{S}(\varphi, L) = L^d\Sigma(M_0) + \text{const.}\ \sigma(T)L^{d-1}\left(1 - (\varphi/M_0)^2\right)^{(d-1)/d}.$$

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 $2\sigma(T)L^{d-1}$.

A37.1.2 Finite size correlation length in Ising-like systems below T_c

In Chapter 41, we shall relate the restoration of discrete symmetries in one dimension to the existence of instantons. Assuming the results of Chapter 41 known, we here indicate how these arguments extend to scalar field theories in higher dimensions in order to demonstrate the complete parallel between the low temperature analysis of Section 23.2 for the Ising model and the instanton analysis. The form of the finite size correlation length below T_c will follow.

We consider, therefore, the example of an effective euclidean action in d dimensions of the form:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\left(\partial_\mu \phi(x)\right)^2 + V\left(\phi(x)\right)\right], \qquad (A37.10)$$

where $V(\phi)$ is a potential which is invariant in the reflection

$$\phi(x) \mapsto -\phi(x)\,, \tag{A37.11}$$

and has degenerate minima. It could for instance have the form of the double-well potential

$$V(\phi) \sim \left(\phi^2 - M_0^2\right)^2. \tag{A37.12}$$

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 37.4 and expand the field in Fourier components. At leading order, we keep only the zero-mode and thus

$$\mathcal{S}(\varphi) = L^{d-1} \int \mathrm{d}t \left[\tfrac{1}{2}(\dot{\varphi})^2 + V(\varphi)\right].$$

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 $L \to \infty$ corresponds to the classical limit $\hbar \to 0$ of quantum mechanics. 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

$$\ln \xi_L(T) \sim \sigma(T) L^{d-1}. \tag{A37.13}$$

This behaviour of the finite size correlation length, characteristic of the breaking of discrete symmetries, has to be contrasted with the power law behaviour (37.93) found in the case of continuous symmetries. Equation ($A37.13$) is valid for temperatures $T < T_c$, and the surface tension vanishes at T_c. The behaviour near T_c is given by equation (37.62) 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. We have found that for $T < T_c$, the finite size correlation length is much larger than the size of the system. This justifies a mode expansion. Integrating over non-zero modes, we obtain 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

$$\mathcal{S}_{\text{eff}}(\theta) \sim \int \mathrm{d}^{d-1}x \left[(\partial_i \theta(x))^2 + O\left(|\partial_i \theta(x)|^4\right)\right]. \tag{A37.14}$$

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 tells us that the effective action begins with a term proportional to the area of the wall and has thus the form

$$\mathcal{S}_{\text{eff}}(\theta) \sim \int \mathrm{d}^{d-1}x \left[1 + \left(\partial_i \theta(x)\right)^2\right]^{1/2} . \tag{A37.15}$$

It has been conjectured by Wallace and Zia 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 remind us the non-linear σ-model.

A37.1.3 Dynamics in the ordered phase

We consider a purely dissipative dynamics associated with a static action of the form (A37.10), for example,

$$\mathcal{A}(\varphi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\left(\partial_\mu \varphi(x)\right)^2 + g\left(\varphi^2 - M_0^2\right)^2\right]. \tag{A37.16}$$

We assume that our system evolves inside a hypercube of linear size L with periodic boundary conditions.

It is shown in Section 41.3 that the difference between the two first eigenvalues of the corresponding Fokker–Planck hamiltonian is of the order of

$$\mathrm{e}^{-\Delta\mathcal{A}}, \quad \text{with} \quad \Delta\mathcal{A} = \mathcal{A}_{\text{max}} - \mathcal{A}_{\text{min}} ,$$

in which $\mathcal{A}_{\text{min}}$ is the value of $\mathcal{A}(\varphi)$ at the degenerate minima and $\mathcal{A}_{\text{max}}$ the value at the saddle point which separates them. We have already briefly analysed this problem in Section 23.4. 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 will correspond 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. Calling 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

$$\tau_L \propto \mathrm{e}^{2\sigma L^{d-1}} . \tag{A37.17}$$

Comparing with equation (A37.13), we find a simple relation between the relaxation time in a cubic geometry and the finite size correlation length ξ_L in a cylindrical geometry:

$$\tau_L \propto \xi_L^2 . \tag{A37.18}$$

become important. General euclidean invariance tells us that the effective action begins with a term proportional to the area of the wall and has thus the form

$$\mathcal{S}_{\text{eff}}(\theta) \sim \int \mathrm{d}^{d-1}x \left[1 + \left(\partial_i\theta(x)\right)^2\right]^{1/2} . \tag{A37.15}$$

It has been conjectured by Wallace and Zia 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 remind us the non-linear σ-model.

A37.1.3 Dynamics in the ordered phase

We consider a purely dissipative dynamics associated with a static action of the form (A37.10), for example,

$$\mathcal{A}(\varphi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\left(\partial_\mu\varphi(x)\right)^2 + g\left(\varphi^2 - M_0^2\right)^2\right]. \tag{A37.16}$$

We assume that our system evolves inside a hypercube of linear size L with periodic boundary conditions.

It is shown in Section 41.3 that the difference between the two first eigenvalues of the corresponding Fokker–Planck hamiltonian is of the order of

$$\mathrm{e}^{-\Delta\mathcal{A}}, \quad \text{with} \quad \Delta\mathcal{A} = \mathcal{A}_{\text{max}} - \mathcal{A}_{\text{min}} ,$$

in which $\mathcal{A}_{\text{min}}$ is the value of $\mathcal{A}(\varphi)$ at the degenerate minima and $\mathcal{A}_{\text{max}}$ the value at the saddle point which separates them. We have already briefly analysed this problem in Section 23.4. 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 will correspond 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. Calling 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

$$\tau_L \propto \mathrm{e}^{2\sigma L^{d-1}} . \tag{A37.17}$$

Comparing with equation (A37.13), we find a simple relation between the relaxation time in a cubic geometry and the finite size correlation length ξ_L in a cylindrical geometry:

$$\tau_L \propto \xi_L^2 . \tag{A37.18}$$

A37.2 Perturbation Theory in a Finite Volume

Let us add a few remarks concerning the calculation of Feynman diagrams in a finite volume. It is convenient to introduce Schwinger's parameters and write the momentum space propagator

$$\Delta(p) = \frac{1}{p^2 + \mu^2} = \int_0^{\infty} \mathrm{d}s \, \mathrm{e}^{-s(p^2+\mu^2)},$$

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

$$D_\gamma \equiv L^{-d} \sum_{\mathbf{p}=2\pi\mathbf{k}/L} \left(p^2 + \mu^2\right)^{-\sigma} = \frac{L^{-d}}{\Gamma(\sigma)} \int_0^{\infty} \mathrm{d}s \, s^{\sigma-1} \sum_{\mathbf{p}=2\pi\mathbf{k}/L} \mathrm{e}^{-s(p^2+\mu^2)} .$$

In terms of the function $\vartheta_0(s)$ defined by (37.42),

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-\pi s n^2}, \tag{A37.19}$$

and related to Jacobi's elliptic θ_3 function

$$\vartheta_0(s) = \theta_3(0, \mathrm{e}^{-\pi s}),$$

the sums can be written as

$$D_\gamma = \frac{L^{-d}}{\Gamma(\sigma)} \int_0^{\infty} \mathrm{d}s \, s^{\sigma-1} \, \mathrm{e}^{-s\mu^2} \, \vartheta_0^d(4s\pi/L^2) = \frac{L^{2\sigma-d}}{(4\pi)^\sigma \Gamma(\sigma)} \int_0^{\infty} \mathrm{d}s \, s^{\sigma-1} \, \mathrm{e}^{-sL^2\mu^2/(4\pi)} \, \vartheta_0^d(s).$$

Poisson's formula is useful in this context. Let $f(x)$ be a function which has a Fourier transform

$$\tilde{f}(k) = \int \mathrm{d}x \, f(x) \, \mathrm{e}^{i2\pi kx} .$$

Then from

$$\sum_{k=-\infty}^{+\infty} \mathrm{e}^{i2\pi kx} = \sum_{l=-\infty}^{+\infty} \delta(x-l),$$

follows Poisson's formula

$$\sum_{k=-\infty}^{+\infty} \tilde{f}(k) = \sum_{l=-\infty}^{+\infty} f(l). \tag{A37.20}$$

Applying this relation to the function $\mathrm{e}^{-\pi s x^2}$, one finds the identity

$$\vartheta_0(s) = s^{-1/2} \vartheta_0\left(1/s\right). \tag{A37.21}$$

This identity shows, in particular, that the infinite size limit is approached exponentially when the mass μ is finite:

$$\vartheta_0(s) - s^{-1/2} \sim 2s^{-1/2} \, \mathrm{e}^{-\pi/s} \Rightarrow D_\gamma(L) - D_\gamma(L=\infty) \underset{L\to\infty}{\propto} \mu^{2d-\sigma} (\mu L)^{\sigma-(d+1)/2} \, \mathrm{e}^{-\mu L} .$$

38 QUANTUM FIELD THEORY AT FINITE TEMPERATURE: EQUILIBRIUM PROPERTIES

In this chapter, we review some *equilibrium properties* in Statistical Quantum Field Theory, that is, relativistic Quantum Field Theory (QFT) at finite temperature, a relativistic extension of the statistical quantum theories discussed in Sections 5.5, 5.6. Study of QFT at finite temperature was initially motivated by cosmological problems and more recently has gained additional attention in connection with high energy heavy ion collisions and speculations about possible phase transitions, as found in numerical simulations. Since we are interested here only in equilibrium physics the euclidean (or imaginary) time formalism will be 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 more convenient path integral formulation.

We discuss, in particular, the limit of high temperature or the situation of finite temperature phase transitions. Note that high temperature now refers to an ultra-relativistic limit where the temperature, in energy unit, is much larger than the physical masses of particles. There, the concept of *dimensional reduction* emerges: in many cases, statistical properties of finite temperature QFT in $(1, d-1)$ dimensions can be described by an effective classical statistical field theory in $d-1$ dimensions. Dimensional reduction generalizes a property already observed in the non-relativistic example of the Bose gas in Section 28.3, and indicates that quantum effects are not 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 specially 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 finite size in one dimension. This identification allows, in particular, an analysis of finite temperature QFT in terms of the renormalization group and the theory of finite size effects of the classical theory.

We illustrate these ideas with several standard examples, the ϕ^4 field theory, the non-linear σ model, the Gross–Neveu model, 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 a specially convenient tool to study the complete crossover between low and high temperature, and, therefore, dimensional reduction.

38.1 Finite (and High) Temperature Field Theory

In this section, we present 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 described by an effective local classical statistical field theory in $d-1$ dimensions, and indicate how to construct it explicitly.

3 QUANTUM FIELD THEORY AT FINITE TEMPERATURE: EQUILIBRIUM PROPERTIES

In this chapter, we review some *equilibrium properties* in Statistical Quantum Field Theory, that is, relativistic Quantum Field Theory (QFT) at finite temperature, a relativistic extension of the statistical quantum theories discussed in Sections 5.5, 5.6. Study of QFT at finite temperature was initially motivated by cosmological problems and more recently has gained additional attention in connection with high energy heavy ion collisions and speculations about possible phase transitions, as found in numerical simulations. Since we are interested here only in equilibrium physics the euclidean (or imaginary) time formalism will be 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 more convenient path integral formulation.

We discuss, in particular, the limit of high temperature or the situation of finite temperature phase transitions. Note that high temperature now refers to an ultra-relativistic limit where the temperature, in energy unit, is much larger than the physical masses of particles. There, the concept of *dimensional reduction* emerges: in many cases, statistical properties of finite temperature QFT in $(1, d-1)$ dimensions can be described by an effective classical statistical field theory in $d-1$ dimensions. Dimensional reduction generalizes a property already observed in the non-relativistic example of the Bose gas in Section 28.3, and indicates that quantum effects are not 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 specially 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 finite size in one dimension. This identification allows, in particular, an analysis of finite temperature QFT in terms of the renormalization group and the theory of finite size effects of the classical theory.

We illustrate these ideas with several standard examples, the ϕ^4 field theory, the non-linear σ model, the Gross–Neveu model, 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 a specially convenient tool to study the complete crossover between low and high temperature, and, therefore, dimensional reduction.

38.1 Finite (and High) Temperature Field Theory

In this section, we present 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 described by an effective local classical statistical field theory in $d-1$ dimensions, and indicate how to construct it explicitly.

38.1.1 Finite temperature quantum field theory

The equilibrium properties of QFT at finite temperature can be derived from the partition function $\mathcal{Z} = \operatorname{tr} \mathrm{e}^{-H/T}$, where H is the quantum hamiltonian and T the temperature. For a simple theory with boson fields ϕ and euclidean action $\mathcal{S}(\phi)$, the partition function is given by the functional integral

$$\mathcal{Z} = \int [\mathrm{d}\phi] \exp\left[-\mathcal{S}(\phi)\right], \tag{38.1}$$

where $\mathcal{S}(\phi)$ is the integral of the euclidean lagrangian density $\mathcal{L}(\phi)$:

$$\mathcal{S}(\phi) = \int_0^{1/T} \mathrm{d}\tau \int \mathrm{d}^{d-1}x\, \mathcal{L}(\phi),$$

and the field ϕ satisfies periodic boundary conditions in the (euclidean or imaginary) time direction

$$\phi(\tau = 0, x) = \phi(\tau = 1/T, x).$$

The QFT may also involve fermions. Fermion fields $\psi(\tau, x)$ instead satisfy anti-periodic boundary conditions

$$\psi(\tau = 0, x) = -\psi(\tau = 1/T, x).$$

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,

$$\phi(t, x) = \sum_{n \in \mathcal{Z}} \mathrm{e}^{i\omega_n t} \phi_n(x), \quad \omega_n = 2n\pi T\,. \tag{38.2}$$

In the case of fermions, anti-periodic boundary conditions lead to the expansion

$$\psi(t, x) = \sum_{\omega_n = (2n+1)\pi T} \mathrm{e}^{i\omega_n t} \psi_n(x). \tag{38.3}$$

Remark. The mode expansion (38.2) 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 (38.1) 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 (38.1) corresponds to a finite size $\beta = 1/T$ in one direction for the classical partition function. The zero temperature limit of the quantum theory corresponds to the usual infinite volume limit of the classical theory.

An important parameter is the correlation length $\xi_T = 1/m_T$ which characterizes the decay of correlations in space directions. A crossover is expected between a d-dimensional behaviour when the correlation length ξ_T is small compared to β, that is, the thermal mass m_T is large compared to the temperature T, to the $(d-1)$-dimensional behaviour when m_T is small compared to T. For momenta much smaller than the temperature T

or distances much larger than β 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 finally that in QFT the initial microscopic scale Λ^{-1}, where Λ is the QFT cut-off, always appears in the bare theory. Even at high temperature, the ratio Λ/T is assumed to remain large.

Renormalization group. General results obtained in the study of finite size effects then apply. For example, renormalization group (RG) equations are only sensitive to short distance singularities, and, therefore, finite sizes do not modify RG equations, as we have discussed in Chapter 37. Correlation functions satisfy the RG equations of the corresponding d-dimensional theory. Finite size affects only the solutions of the RG equations, because a new dimensionless, RG invariant, variable appears which can be written as the ratio m_T/T.

In the tree approximation, the thermal mass m_T coincide with m the physical mass scale. At high temperature, the ratio m/T goes to zero. If m_T remains of order m, beyond leading order, the same will apply 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 zero at high temperature then the ratio m_T/T really becomes small. Two examples will be met: the first one corresponds to theories where the free field theory is an IR fixed point, like ϕ^4_4 scalar field theory or QED$_4$, the second to UV asymptotically free field theories. 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.

38.1.2 Zero-mode, dimensional reduction and effective local field theory

The role of the zero-mode. We consider first the example of a free scalar field theory with the action

$$\mathcal{S}(\phi)=\frac{1}{2}\int_0^{1/T}\mathrm{d}t\int \mathrm{d}^{d-1}x\left[(\partial_t\phi)^2+(\nabla_x\phi)^2+m^2\phi^2\right].$$

If we introduce the mode expansion (38.2) in the action and integrate over time, we obtain an euclidean $(d-1)$-dimensional field theory with an infinite number of fields, the modes $\phi_n(x)$:

$$\mathcal{S}(\phi)=\frac{1}{2T}\int \mathrm{d}^{d-1}x\sum_n\left[|\nabla_x\phi_n(x)|^2+(m^2+4\pi^2n^2T^2)|\phi_n(x)|^2\right].$$

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. The large distance, low momentum, physics can entirely described by an effective $(d-1)$-dimensional field theory involving only the zero-mode.

Note that due to anti-periodic boundary conditions the thermal masses of fermion modes (38.3) become all very large and fermion modes decouple at high temperature.

In an interacting theory, we expect all scalar non-zero modes and fermion modes to have thermal masses at least of order T. If m_T the thermal mass of the zero-mode remains much smaller than the temperature, then it is still possible to describe low momentum physics 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. As we have already mentioned, in the reduced $(d-1)$-dimensional theory, the temperature T acts as a large momentum cut-off.

Note that in the following when we speak about 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.

The effective field theory. We now outline the construction of the $(d-1)$-dimensional effective theory 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

$$\phi(t,x) = \varphi(x) + \chi(t,x), \tag{38.4}$$

where φ is the zero-mode and χ the sum of all other modes (equation (38.2)):

$$\chi(t,x) = \sum_{n\neq 0} \mathrm{e}^{i\omega_n t}\,\phi_n(x), \quad \omega_n = 2n\pi T\,. \tag{38.5}$$

The action $\mathcal{S}_T(\varphi)$ of the reduced theory then is defined by

$$\mathrm{e}^{-\mathcal{S}_T(\varphi)} = \int [\mathrm{d}\chi]\exp[-\mathcal{S}(\varphi+\chi)]. \tag{38.6}$$

At leading order in perturbation theory one sets $\chi = 0$ and simply finds

$$\mathcal{S}_T(\varphi) = \frac{1}{T}\int \mathrm{d}^{d-1}x\,\mathcal{L}(\varphi), \tag{38.7}$$

an action that is automatically local. We note that T plays, in this leading approximation, the formal role of $\hbar$, and the small T expansion corresponds 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 non-renormalizable. However, another dimensionless combination can be found, m/T, 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 integration over non-zero modes the effective action contains non-local interactions. To study long wave length 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 will 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 which render the perturbative expansion of the effective field theory 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.

Note, however, that new apparent divergences can be generated by the local expansion. Determining the finite parts may involve some non-trivial calculations.

38.2 The Example of the $\phi^4_{1,d-1}$ Field Theory

We first study the example of the $(\phi^2)^2$ scalar field theory where field ϕ is a N-component vector and the hamiltonian $\mathcal{H}(\mathbf{\Pi}, \phi)$ is $O(N)$ symmetric:

$$\mathcal{H}(\mathbf{\Pi}, \phi) = \frac{1}{2}\int \mathrm{d}^{d-1}x\, \mathbf{\Pi}^2(x) + \Sigma(\phi) \tag{38.8}$$

with

$$\Sigma(\phi) = \int \mathrm{d}^{d-1}x \left\{\frac{1}{2}\left[\nabla\phi(x)\right]^2 + \frac{1}{2}(r_c + r)\phi^2(x) + \frac{1}{4!}u\left(\phi^2(x)\right)^2\right\}. \tag{38.9}$$

A cut-off Λ as usual is implied that renders the field theory UV finite. The quantity $r_c(u)$ has the form of a mass renormalization. It is determined by the condition that at $T = 0$ (zero temperature) when r vanishes the physical mass m of the field ϕ vanishes. At $r = 0$ a transition occurs between a symmetric phase, $r > 0$, and a broken phase, $r < 0$. 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) $|r| \ll \Lambda^2$ or, for $N \neq 1$, $r < 0$ which corresponds to a spontaneously broken symmetry with massless Goldstone modes. The latter situation will be examined in Section 38.4 within the more suitable formalism of the non-linear σ-model.

It is also convenient to introduce a dimensionless coupling λ:

$$u = \Lambda^{4-d}\lambda\,, \tag{38.10}$$

which later will be assumed to take generic (i.e. not very small) values.

The quantum partition function (38.1) at finite temperature T then is calculated with the action

$$\mathcal{S}(\phi) = \int_0^{1/T} \mathrm{d}t \left[\int \mathrm{d}^{d-1}x \frac{1}{2}(\mathrm{d}_t\phi)^2 + \Sigma(\phi)\right], \tag{38.11}$$

where the field ϕ satisfies periodic boundary conditions in the euclidean time direction.

38.2.1 Renormalization group at finite temperature

As we have already mentioned, some useful information can be obtained from renormalization group 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, since then the ϕ_4^4 theory 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 non-renormalizable, which means that 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, $m \propto r^{1/2}$. 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,

$$m_T \propto (r + \text{const.}\ \lambda\Lambda^{4-d}T^{d-2})^{1/2}.$$

This general form has several consequences.

If at zero temperature, the symmetry is broken ($r < 0$), a phase transition occurs at a temperature T_c which scales like

$$T_c \propto \Lambda \left(-r/\Lambda^2\right)^{1/(d-2)} \gg (-r)^{1/2}.$$

This means that the critical temperature is large with respect to the $T = 0$ crossover mass scale (see Section 31.3.2) of the massless phase for $N > 1$, and with respect to the physical mass $m \propto (-r)^{1/2}$ in the case $N = 1$.

At high temperature or in the massless theory ($r = 0$), 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 take the form

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(\lambda)\frac{\partial}{\partial\lambda} - \frac{n}{2}\eta(\lambda) - \eta_2(\lambda)r\frac{\partial}{\partial r}\right]\Gamma^{(n)}(p_i;\Lambda,\lambda,r,T) = 0. \tag{38.12}$$

The ratio $m_T/T = F(\Lambda,\lambda,r,T)$ is dimensionless and RG invariant, and thus satisfies

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(\lambda)\frac{\partial}{\partial\lambda} - \eta_2(\lambda)r\frac{\partial}{\partial r}\right]F(\Lambda,\lambda,r,T) = 0.$$

The solution can be written as

$$m_T/T = F(\Lambda/T,\lambda,r/T^2,1) = F\big(\ell\Lambda/T,\lambda(\ell),r(\ell)/T^2,1\big), \tag{38.13}$$

where ℓ is a scale parameter, and $\lambda(\ell), r(\ell)$ the corresponding running parameters (or effective parameters at scale ℓ):

$$\ell\frac{\mathrm{d}\lambda(\ell)}{\mathrm{d}\ell} = \beta\big(\lambda(\ell)\big), \quad \ell\frac{\mathrm{d}r(\ell)}{\mathrm{d}\ell} = -r(\ell)\eta_2\big(\lambda(\ell)\big).$$

The form of the RG β-function

$$\beta(\lambda) = \frac{(N+8)}{48\pi^2}\lambda^2 + O\left(\lambda^3\right), \tag{38.14}$$

implies that the theory is IR free, that is, that $\lambda(\ell) \to 0$ for $\ell \to 0$. 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

$$\lambda(T/\Lambda) \sim \frac{48\pi^2}{(N+8)\ln(\Lambda/T)}. \tag{38.15}$$

From

$$\eta_2(\lambda) = -\frac{N+2}{48\pi^2}\lambda + O\left(\lambda^2\right),$$

one also finds

$$r(T/\Lambda) \propto \frac{r}{(\ln \Lambda/T)^{(N+2)/(N+8)}}. \tag{38.16}$$

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 will be discussed in Section 38.3.2.

Dimension $d = 3$. The three-dimensional classical theory has an IR fixed point λ^*. Then, finite size scaling (equation (38.13)) predicts, in the symmetric phase,

$$m_T/T = f(r/T^{1/\nu}),$$

where is ν the correlation exponent of the three-dimensional system. Therefore, m_T in general 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 ($r < 0$). If $f(x)$ vanishes at a value $x = x_0 < 0$, the critical temperature T_c has a scaling behaviour of the form

$$T_c = (-x_0)^{-\nu}(-r)^{\nu} \propto m\,,$$

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 will be examined in the framework of the non-linear σ-model starting with Section 38.4.

38.2.2 One-loop effective action

We now construct the effective $(d-1)$-dimensional theory and discuss its validity. Note, however, that this construction is useful mainly if the IR contributions are large enough to invalidate perturbation theory, that is, in dimensions $d \leq 4$. Instead for dimensions $d > 4$, the reduced $(d-1)$-dimensional theory has a finite perturbation expansion even in the massless limit. This is a property we will check by discussing 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 (38.2)). One then calculates the effective action (38.6) by integrating perturbatively over all non-zero modes. In the notation (38.9), the result at leading order simply is

$$\mathcal{S}_T(\varphi)=\frac{1}{T}\Sigma(\varphi)=\frac{1}{T}\int \mathrm{d}^{d-1}x\left\{\frac{1}{2}\left[\nabla\phi(x)\right]^2+\frac{1}{2}r\phi^2(x)+\frac{1}{4!}u\left(\phi^2(x)\right)^2\right\}, \qquad (38.17)$$

because $r_c(u)$ 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 $\lambda T/\Lambda^{d-4}$. A true expansion parameter is dimensionless and, therefore, $\lambda T/\Lambda^{d-4}$ has to be multiplied by a mass function of T and r 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 at zero temperature, the symmetry is unbroken, at high temperature $T\gg m=\sqrt{r}$ one expects the corrections to the mass m_T of the zero-mode to be dominated by the one-loop tadpole diagram, 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 some 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 is nevertheless perturbative.

If instead r 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. We now calculate the one-loop contribution to the effective action generated by integrating over the non-zero modes. It takes the form ($\ln\det=\operatorname{tr}\ln$)

$$\mathcal{S}_T^{(1)}(\varphi)=\tfrac{1}{2}\operatorname{tr}\ln\left[(-\mathrm{d}_t^2-\nabla_{d-1}^2+r+\tfrac{1}{6}u\varphi^2)\delta_{ij}+\tfrac{1}{3}u\varphi_i\varphi_j\right]-\ (\varphi=0).$$

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.

To evaluate the one-loop contribution $\mathcal{S}_T^{(1)}$ to the reduced action one can, as in the general finite size calculations, introduce Schwinger's representation and Jacobi's related ϑ_0 function (see Section 37.3.2), or use directly the partition function of the harmonic oscillator

$$\begin{aligned}\operatorname{tr}\ln(-\mathrm{d}_t^2-\nabla_{d-1}^2+M^2)&=V_{d-1}\int\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}\sum_{n\neq 0}\ln(\omega_n^2+k^2+M^2)\\&=2V_{d-1}\int\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}\ln\left[2\sinh(\beta\omega(k)/2)/\omega(k)\right],\qquad(38.18)\end{aligned}$$

where V_{d-1} is the $d-1$ volume and $\omega(k)=\sqrt{k^2+M^2}$. Applied to the present example one finds

$$\begin{aligned}\mathcal{S}_T^{(1)}(\varphi)=\int\mathrm{d}^{d-1}x\int\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}&\left\{(N-1)\ln\left[2\sinh(\beta\omega_{\mathrm{T}}(k)/2)/\omega_{\mathrm{T}}(k)\right]\right.\\&\left.+\ln\left[2\sinh(\beta\omega_{\mathrm{L}}(k)/2)\omega_{\mathrm{L}}(k)\right]\right\}-(\varphi=0)\qquad(38.19)\end{aligned}$$

with

$$\omega_{\mathrm{T}}(k)=\sqrt{k^2+r+\tfrac{1}{6}u\varphi^2(x)},\quad \omega_{\mathrm{L}}(k)=\sqrt{k^2+r+\tfrac{1}{2}u\varphi^2(x)}. \tag{38.20}$$

We now perform an expansion in powers of φ. At this order, this expansion makes sense only if $-r/T^2<4\pi^2$, 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 meaningful only if the field expectation value is small enough.

Order φ^2. Note that for the quadratic term the local approximation is not needed because the corresponding one-loop diagram is a constant:

$$\left[\mathcal{S}_T^{(1)}\right]_2=\frac{1}{12}(N+2)\bar{G}_2(r,T)\frac{u}{T}\int \mathrm{d}^{d-1}x\,\varphi^2(x)\,, \tag{38.21}$$

where the constant $\bar{G}_2$ is given by

$$\bar{G}_2(r,T)=\int\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}\omega(k)}\left(\frac{1}{2}+\frac{1}{\mathrm{e}^{\beta\omega(k)}-1}-\frac{T}{\omega(k)}\right) \tag{38.22}$$

with now $\omega(k)=\sqrt{k^2+r}$. 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 (30.12,30.13)),

$$\Omega_d(m)=\frac{1}{(2\pi)^d}\int^{\Lambda}\frac{\mathrm{d}^d k}{k^2+m^2}\,, \tag{38.23}$$

and the UV finite function

$$\begin{aligned} f_d(s)&=N_{d-1}\int_0^{\infty}\frac{x^{d-2}\mathrm{d}x}{\sqrt{x^2+s}}\frac{1}{\exp[\sqrt{x^2+s}]-1}\\ &=N_{d-1}\int_{\sqrt{s}}^{\infty}(y^2-s)^{(d-3)/2}\frac{\mathrm{d}y}{\mathrm{e}^y-1}\,, \end{aligned} \tag{38.24}$$

where N_d is the loop factor defined in (11.29):

$$N_d=\frac{2}{(4\pi)^{d/2}\Gamma(d/2)}, \tag{38.25}$$

we can rewrite $\bar{G}_2$ as

$$\bar{G}_2(r,T)=\Omega_d(\sqrt{r})-T\Omega_{d-1}(\sqrt{r})+T^{d-2}f_d(r/T^2). \tag{38.26}$$

In particular,

$$f_d(0)=N_{d-1}\Gamma(d-2)\zeta(d-2),\quad f_d'(0)=-\tfrac{1}{2}(d-3)N_{d-1}\Gamma(d-4)\zeta(d-4), \tag{38.27}$$

where $\zeta(s)$ is Riemann's ζ function.

Order $(\varphi^2)^2$. The quartic term is proportional to the initial interaction

$$\left[\mathcal{S}_T^{(1)}\right]_4=-\frac{1}{144}(N+8)\bar{G}_4(r,T)\frac{u^2}{T}\int \mathrm{d}^{d-1}x\,(\varphi^2(x))^2. \tag{38.28}$$

One verifies

$$\bar{G}_4(r,T) = -\frac{\partial}{\partial r}\bar{G}_2(r,T). \tag{38.29}$$

The one-loop reduced action. We first keep only the 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 $[\bar{G}_2]_\mathrm{r}$. One obtains ($d>2$)

$$\begin{aligned}[\bar{G}_2]_\mathrm{r}(r,T) &= \bar{G}_2(r,T) - \Omega_d(0)\\ &= -rD_d(\sqrt{r}) - T\Omega_{d-1}(\sqrt{r}) + T^{d-2}f_d(r/T^2),\end{aligned}$$

where we define

$$D_d(m) = \frac{1}{(2\pi)^d}\int^\Lambda \frac{\mathrm{d}^d k}{k^2(k^2+m^2)} = \frac{1}{m^2}\left[\Omega_d(0) - \Omega_d(m)\right]. \tag{38.30}$$

After the rescaling $\varphi \mapsto \varphi T^{1/2}$, the effective action can be written as

$$\mathcal{S}_T(\varphi) = \int \mathrm{d}^{d-1}x \left\{\frac{1}{2}\left[\nabla\varphi(x)\right]^2 + \frac{1}{2}\sigma_2\varphi^2(x) + \frac{1}{4!}\sigma_4\left(\varphi^2(x)\right)^2\right\} \tag{38.31}$$

with

$$\sigma_2 = r + \tfrac{1}{6}(N+2)u[\bar{G}_2]_\mathrm{r}\,, \quad \sigma_4/T = u - \tfrac{1}{6}(N+8)u^2\bar{G}_4\,.$$

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 (38.31)):

$$[\mathcal{S}_T]_{(2n)} \propto \lambda^n (T/\Lambda)^{n(d-4)} T^{d-1-n(d-3)} \int \mathrm{d}^{d-1}x \left(\varphi^2(x)\right)^n, \tag{38.32}$$

and, therefore, increasingly negligible at high temperature at least for $d \ge 4$.

The local expansion of the one-loop determinant also generates monomials with derivatives. No term proportional to $(\partial_\mu\varphi)^2$ 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 $r \ne 0$ but $r/T^2 \ll 1$, we can expand in powers of r and the previous arguments immediately generalize.

38.3 High Temperature and Critical Limits

We now examine two interesting situations. First we discuss $r \to 0$ 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 $r\varphi^2$.

The constants $[\bar{G}_2]_\mathrm{r}$ and $\bar{G}_4$ for $r=0$ become

$$[\bar{G}_2]_\mathrm{r} = -T\Omega_{d-1}(0) + T^{d-2}N_{d-1}\Gamma(d-2)\zeta(d-2), \tag{38.33}$$

$$\bar{G}_4 = D_d(0) - TD_{d-1}(0) + \tfrac{1}{2}N_{d-1}(d-3)\Gamma(d-4)\zeta(d-4)T^{d-4}, \tag{38.34}$$

where the values (38.27) have been used. The expression for $[\bar{G}_2]_{\mathrm{r}}$ 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 (38.29) and the value of $\bar{G}_4$, we obtain the term linear in r in $\bar{G}_2$:

$$\bar{G}_2(r,T) = \bar{G}_2(0,T) - r\bar{G}_4(0,T) + O(r^2). \tag{38.35}$$

Then, we 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. Note that from the point of view of the classical d-dimensional theory, it is more natural to determine the parameter r, 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.

38.3.1 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.

Let us examine more precisely the $d = 5$ case, keeping the contribution of order r in $\bar{G}_2$. Then,

$$[\bar{G}_2]_{\mathrm{r}} = \frac{T^3}{4\pi^2}\zeta(3) - T\Omega_4(0) - r\left[D_5(0) - \frac{T}{8\pi^2}\left(\ln(\Lambda/T) + \kappa_5\right)\right] + O(r^2),$$

$$\bar{G}_4 = D_5(0) - \frac{T}{8\pi^2}\left(\ln(\Lambda/T) + \kappa_5\right),$$

where κ_5 is a renormalization scheme dependent constant, $\Omega_4(0) \propto \Lambda^2$, $D_5(0) \propto \Lambda$. The infinite volume terms proportional to $D_5(0)$ induce finite renormalizations $\lambda \mapsto \lambda_{\mathrm{r}}$ of the dimensionless ϕ^4 coupling constant λ, and $r \mapsto r_{\mathrm{r}}$ of the ϕ^2 coefficient,

$$u = \lambda/\Lambda\,, \quad \lambda_{\mathrm{r}} = \lambda - \frac{D_5(0)}{\Lambda}\frac{N+8}{6}\lambda^2, \quad r_{\mathrm{r}}/r = 1 - \frac{D_5(0)}{\Lambda}\frac{N+2}{6}\lambda\,.$$

The remaining cut-off dependent terms in $[\bar{G}_2]_{\mathrm{r}}$ and $\bar{G}_4$ render the effective four-dimensional theory one-loop finite. Using the expression (38.31) and introducing the small dimensionless (effective) coupling constant

$$\lambda_T = \lambda_{\mathrm{r}} T/\Lambda\,,$$

we can write the effective action at one-loop order ($\varphi \mapsto \varphi T^{1/2}$) as

$$\mathcal{S}_T(\varphi) = \int \mathrm{d}^4 x \left\{\frac{1}{2}\left[\nabla\varphi(x)\right]^2 + \frac{1}{2} r_T \varphi^2(x) + \frac{1}{4!}\lambda_T \left(\varphi^2(x)\right)^2\right\} + \delta\mathcal{S}_{T,\Lambda}(\varphi) \tag{38.36}$$

with

$$r_T = r_{\mathrm{r}} + \frac{1}{6}(N+2)\frac{\zeta(3)}{4\pi^2}T^2\lambda_T\,, \tag{38.37}$$

and where $\delta S_{T,\Lambda}$ is the sum of one-loop counter-terms:

$$\delta S_{T,\Lambda}(\varphi) = \int \mathrm{d}^4 x \left[\frac{1}{2}\delta r_T \varphi^2(x) + \frac{1}{4!}\delta\lambda_T \left(\varphi^2(x)\right)^2\right]$$

with

$$\begin{aligned}
\delta r_T &= -\frac{N+2}{6}\left\{\Omega_4(0) + \frac{r_{\mathrm{r}}}{8\pi^2}\left[\ln(\Lambda/T) + \kappa_5\right]\right\}\lambda_T\,, \\
\delta\lambda_T &= \frac{N+8}{48\pi^2}\left[\ln(\Lambda/T) + \kappa_5\right]\lambda_T^2\,.
\end{aligned}$$

Note that for dimensions $d \geq 5$ we can study the effective theory by perturbation theory and renormalization group. The temperature T plays the role of the cut-off in the reduced theory. The perturbative expansion of the reduced four-dimensional theory contains large logarithms of the form $\ln(r_{\mathrm{r}}/T^2)$ 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 $r = 0$ the mass m_T is

$$m_T/T = \left[\tfrac{1}{24\pi^2}(N+2)\zeta(3)\lambda_T\right]^{1/2} \ll 1\,.$$

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 (38.37), one finds in the tree approximation

$$T_c^2 = -\frac{24\pi^2}{(N+2)\zeta(3)}\frac{r_{\mathrm{r}}}{\lambda_T}\,,$$

an expression that makes sense only if $T_c \ll \Lambda$ and thus $|r_{\mathrm{r}}| \ll T\Lambda$, that is, very small. This condition justifies a small r expansion and implies that the critical temperature is large in the zero-temperature mass scale $\sqrt{-r_{\mathrm{r}}}$ (the mass scale corresponding to the crossover between critical and Goldstone mode behaviours). The equation also confirms that a phase transition is possible only if at $T = 0$ (zero QFT temperature), the system is in the ordered phase.

38.3.2 Dimension $d \leq 4$

Four dimensions. For $d = 4$ and at order r, we now obtain

$$[\bar{G}_2]_{\mathrm{r}} = \frac{T^2}{12} - T\Omega_3(0) - \frac{1}{8\pi^2}\left[\ln(\Lambda/T) + \kappa_4\right] r. \tag{38.38}$$

$$\bar{G}_4 = \frac{1}{8\pi^2}\left[\ln(\Lambda/T) + \kappa_4\right], \tag{38.39}$$

where $\Omega_3(0) \propto \Lambda$ and κ_4 is a renormalization scheme dependent constant.

The coupling constant u is dimensionless $u \equiv \lambda$. Then, $\bar{G}_4$ just yields the one-loop contribution to the perturbative expansion of the running (or effective) coupling constant

$$\lambda(T/\Lambda) = \lambda - \frac{N+8}{48\pi^2}\left[\ln(\Lambda/T) + \kappa_4\right]\lambda^2 + O(\lambda^3).$$

In fact, we know from RG arguments that all quantities can be expressed entirely in terms of the running coupling constant.

In the same way, $\bar{G}_2$ contains a one-loop contribution to the perturbative expansion of $r(T/\Lambda)$, the running coefficient of ϕ^2:

$$r(T/\Lambda)/r = 1 - \frac{N+2}{48\pi^2}\left[\ln(\Lambda/T) + \kappa_4\right]\lambda\,.$$

The three-dimensional effective theory is super-renormalizable, and thus requires only a mass renormalization. In $\bar{G}_2$, we find two terms, $-T\Omega_3(0)$ that is cut-off dependent and yields the one-loop counter-term, and $T^2/12$ that contributes to the mass of the zero-mode. The one-loop effective action takes the form

$$\begin{aligned} \mathcal{S}_T(\varphi\sqrt{T}) = \int \mathrm{d}^3x \left\{\frac{1}{2}\left[\nabla\varphi(x)\right]^2 + \frac{1}{2} r_T \varphi^2(x) + \frac{1}{4!}\lambda_T \left(\varphi^2(x)\right)^2\right\} \\ + \text{one-loop counter-terms} \end{aligned} \tag{38.40}$$

with

$$r_T = r(T/\Lambda) + \frac{N+2}{72} T^2 \lambda(T/\Lambda), \quad \lambda_T = T\lambda(T/\Lambda).$$

The massless theory. For $r = 0$ (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 (38.15)). Therefore,

$$(m_T/T)^2 \sim \frac{2\pi^2(N+2)}{3(N+8)} \frac{1}{\ln(\Lambda/T)}\,. \tag{38.41}$$

The mass of the zero-mode is smaller, even though 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. The convergence, however, is expected to be extremely slow and, therefore, summation techniques have been proposed. General summation methods, which have been used in the calculation of 3D critical exponents, could also be useful here.

The critical temperature. If at zero temperature, the system is in an ordered phase ($r < 0$), at higher temperature a phase transition occurs at a critical temperature T_c, which at leading order is solution of the equation

$$r_T = r(T/\Lambda) + \frac{N+2}{72} T^2 \lambda(T/\Lambda) = 0\,.$$

This relation can be rewritten in various forms, for example,

$$\sqrt{(N+2)/12}\,T_c \propto m\sqrt{\ln(\Lambda/m)} \propto (-r)^{1/2}\big(\ln(-r)\big)^{3/(N+8)}, \qquad (38.42)$$

where m is the low temperature physical (crossover) mass scale. We note that 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 have devoted several chapters, starting with Chapter 25.

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 $d < 4$. Renormalization group arguments tell us 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.

38.4 The Non-Linear σ-Model in the Large N Limit

We now discuss another related example, the non-linear σ-model, because it is better suited to describe the ordered phase when the symmetry is continuous, and the influence of Goldstone modes. Moreover, due the non-linear character of the group representation, one is confronted with difficulties which also appear 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 dealing with the non-linear σ-model in the perturbative framework, we discuss the finite temperature properties 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. We recall that it has been proven in Section 30.6 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 ϕ^4 coupling), both QFTs corresponding to two perturbative expansions, in different parameters, of the same physical model.

The non-linear σ-model. The non-linear σ-model has been studied at zero temperature in Section 31.2. It is an $O(N)$ symmetric QFT, with an N-component scalar field $\mathbf{S}(t, x)$ which belongs to a sphere, that is, satisfies the constraint $\mathbf{S}^2(t, x) = 1$.

The partition function of the non-linear σ-model can be written as

$$\mathcal{Z} = \int [\mathrm{d}\mathbf{S}(t,x)\mathrm{d}\lambda(t,x)] \exp\left[-\mathcal{S}(\mathbf{S},\lambda)\right] \tag{38.43}$$

with

$$\mathcal{S}(\mathbf{S},\lambda) = \frac{1}{2g}\int_0^{1/T} \mathrm{d}t\,\mathrm{d}^{d-1}x\left[\left(\partial_t \mathbf{S}(t,x)\right)^2 + \left(\nabla\mathbf{S}(t,x)\right)^2 + \lambda(t,x)\left(\mathbf{S}^2(t,x)-1\right)\right]. \tag{38.44}$$

The λ integration runs along the imaginary axis and enforces the constraint $\mathbf{S}^2(x)=1$. The parameter g is the coupling constant of the quantum model as well as the temperature of the corresponding classical theory in d dimensions.

As we have already stressed, finite temperature T corresponds to one finite size $\beta = 1/T$ with periodic boundary conditions in the corresponding d-dimensional classical theory.

The large N limit: Finite temperature saddle point equations. The non-linear σ-model has been discussed in the large N limit at zero temperature in Section 30.6 with a slightly different notation ($T \to g$). At finite temperature, the saddle point equation (30.71a) remains unchanged. The saddle point equation (30.71b) is modified because the frequencies in the time direction are quantized. In the symmetric phase $\langle \mathbf{S}(T)\rangle = 0$, it becomes

$$1 = (N-1)\, g\, G_2(m_T,T), \tag{38.45}$$

where

$$\begin{aligned} G_2(m_T,T) &= \frac{T}{(2\pi)^{d-1}}\sum_{n\in\mathbb{Z}}\int^{\Lambda}\frac{\mathrm{d}^{d-1}k}{(2\pi nT)^2+k^2+m_T^2} \\ &= \int^{\Lambda}\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}\frac{1}{\omega(k)}\left(\frac{1}{2}+\frac{1}{\mathrm{e}^{\beta\omega(k)}-1}\right). \end{aligned} \tag{38.46}$$

Introducing the functions (38.24,38.23), we can write equation (38.45) (for N large)

$$1/Ng = \Omega_d(m_T) + T^{d-2} f_d(m_T^2/T^2). \tag{38.47}$$

Here, $\xi_T = m_T^{-1}$ has the meaning of a correlation length in space directions.

A phase transition is possible only if $f_d(0)$ is finite for $m_T = 0$. We see from equation (38.27) that this implies $d > 3$. The result has a simple interpretation: at $d = 3$ 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 result 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. Instead the large N expansion allows to study the problem because it exists for any dimension.

Dimension $d=2$. We first examine the case $d=2$, which is also a special case of the situation discussed in Section 37.5.2. This corresponds to a situation where even at zero temperature, the symmetry remains always 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 (30.81):

$$1/Ng = \Omega_2(m) \;\Rightarrow\; m = \xi_0^{-1} \propto \Lambda\,\mathrm{e}^{-2\pi/Ng}.$$

The partition function of the non-linear σ-model can be written as

$$\mathcal{Z} = \int [\mathrm{d}\mathbf{S}(t,x)\mathrm{d}\lambda(t,x)] \exp[-\mathcal{S}(\mathbf{S},\lambda)] \tag{38.43}$$

with

$$\mathcal{S}(\mathbf{S},\lambda) = \frac{1}{2g}\int_0^{1/T} \mathrm{d}t\,\mathrm{d}^{d-1}x \left[\left(\partial_t \mathbf{S}(t,x)\right)^2 + \left(\nabla \mathbf{S}(t,x)\right)^2 + \lambda(t,x)\left(\mathbf{S}^2(t,x) - 1\right)\right]. \tag{38.44}$$

The λ integration runs along the imaginary axis and enforces the constraint $\mathbf{S}^2(x) = 1$. The parameter g is the coupling constant of the quantum model as well as the temperature of the corresponding classical theory in d dimensions.

As we have already stressed, finite temperature T corresponds to one finite size $\beta = 1/T$ with periodic boundary conditions in the corresponding d-dimensional classical theory.

The large N limit: Finite temperature saddle point equations. The non-linear σ-model has been discussed in the large N limit at zero temperature in Section 30.6 with a slightly different notation ($T \to g$). At finite temperature, the saddle point equation (30.71a) remains unchanged. The saddle point equation (30.71b) is modified because the frequencies in the time direction are quantized. In the symmetric phase $\langle \mathbf{S}(T)\rangle = 0$, it becomes

$$1 = (N-1)\,g\,G_2(m_T, T), \tag{38.45}$$

where

$$\begin{aligned} G_2(m_T,T) &= \frac{T}{(2\pi)^{d-1}} \sum_{n\in\mathcal{Z}} \int^{\Lambda} \frac{\mathrm{d}^{d-1}k}{(2\pi n T)^2 + k^2 + m_T^2} \\ &= \int^{\Lambda} \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\omega(k)} \left(\frac{1}{2} + \frac{1}{\mathrm{e}^{\beta\omega(k)} - 1}\right). \end{aligned} \tag{38.46}$$

Introducing the functions (38.24,38.23), we can write equation (38.45) (for N large)

$$1/Ng = \Omega_d(m_T) + T^{d-2} f_d(m_T^2/T^2). \tag{38.47}$$

Here, $\xi_T = m_T^{-1}$ has the meaning of a correlation length in space directions.

A phase transition is possible only if $f_d(0)$ is finite for $m_T = 0$. We see from equation (38.27) that this implies $d > 3$. The result has a simple interpretation: at $d = 3$ 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 result 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. Instead the large N expansion allows to study the problem because it exists for any dimension.

Dimension $d = 2$. We first examine the case $d = 2$, which is also a special case of the situation discussed in Section 37.5.2. This corresponds to a situation where even at zero temperature, the symmetry remains always 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 (30.81):

$$1/Ng = \Omega_2(m) \;\Rightarrow\; m = \xi_0^{-1} \propto \Lambda\,\mathrm{e}^{-2\pi/Ng}.$$

By subtracting this equation from equation (38.47) (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 (38.24) then is dominated by the contribution of the zero-mode and, therefore,

$$T/m_T = \frac{1}{\pi}\ln(m_T/m) \sim \frac{1}{\pi}\ln(T/m). \tag{38.48}$$

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

$$\frac{1}{Ng} - \frac{1}{Ng_c} = \Omega_d(m_T) - \Omega_d(0) + T^{d-2} f_d(m_T^2/T^2). \tag{38.49}$$

For $g > g_c$, the equation can be also be written in terms of the physical mass m (equation (30.75)):

$$\left[\Omega_d(m) - \Omega_d(m_T)\right]/T^{d-2} = f_d(m_T^2/T^2). \tag{38.50}$$

The behaviour of the system then depends on the ratio T/m. To obtain more detailed results, we have to specify dimensions.

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

$$f_3(s) = -\frac{1}{2\pi}\ln\left(1 - \mathrm{e}^{-\sqrt{s}}\right),$$

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

$$2\sinh(m_T/2T) = \mathrm{e}^{m/2T}.$$

One verifies for m/T large (low temperature) $m_T \to m$, and at high temperature $T \gg m$, m_T becomes proportional to T:

$$m_T/T \sim 2\ln\left((1+\sqrt{5})/2\right).$$

For $g < g_c$, that is, when the symmetry is broken at zero temperature, one has to return to the general form

$$2\sinh(m_T/2T) = \exp\left[-\frac{2\pi}{NT}\left(\frac{1}{g} - \frac{1}{g_c}\right)\right]. \tag{38.51}$$

One can also introduce the mass scale $m_{\mathrm{cr}}(g)$ (see equations (30.79) and (38.59)):

$$m_{\mathrm{cr}}(g) = \frac{1}{g} - \frac{1}{g_c},$$

which at zero temperature characterizes the crossover between critical and Goldstone mode behaviours. Then,

$$2\sinh(m_T/2T) = \mathrm{e}^{-2\pi m_{\mathrm{cr}}/NT}.$$

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 $m_{\mathrm{cr}} = O(\Lambda) \gg T$: 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 $T \ll m_{\mathrm{cr}}$, that is, at low (but non-zero) temperature. Then,

$$m_T \sim T\,\mathrm{e}^{-2\pi m_{\mathrm{cr}}/NT} = T\exp\left[-\frac{2\pi}{NT}\left(\frac{1}{g} - \frac{1}{g_c}\right)\right]. \tag{38.52}$$

Note that the mass m_T has, when the coupling constant g or the temperature T go to zero, the 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.

$d > 3$: *Critical temperature.* The quantity $f_d(0)$ then 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 (38.49), one infers

$$T_c^{d-2} = \frac{1}{N f_d(0)}\left(\frac{1}{g} - \frac{1}{g_c}\right). \tag{38.53}$$

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 $m_{\mathrm{cr}}(g)$ between critical and Goldstone behaviours, which in all dimensions has near g_c the same scaling property as the physical mass above g_c.

Dimension $d = 4$. Since $f_4(0) = 1/12$, one finds

$$T_c^2 = \frac{12}{N}\left(\frac{1}{g} - \frac{1}{g_c}\right) \propto m_{\mathrm{cr}}^2 \ln(\Lambda/m_{\mathrm{cr}}) \gg m_{\mathrm{cr}}^2. \tag{38.54}$$

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 $f_4(0) = 1/12$, we obtain

$$(m_T/T)^2 = \frac{2\pi^2}{3\ln(\Lambda/T)},$$

in agreement with equation (38.41).

Dimension $d = 5$. From $f_5(0) = \zeta(3)/4\pi^2$, we infer the critical temperature T_c:

$$T_c^3 \sim \frac{4\pi^2}{N\zeta(3)}\left(\frac{1}{g} - \frac{1}{g_c}\right) \propto \Lambda m_{\rm cr}^2 \gg m_{\rm cr}^3.$$

Thus for $d \geq 4$, the critical temperature T_c is large in the relevant physical scale.

In the massless limit $g = g_c$,

$$(m_T/T)^2 \sim \frac{\zeta(3)}{4\pi^2}\frac{T}{D_5(0)}, \tag{38.55}$$

($D_5(0) \propto \Lambda$) in agreement with the behaviour found in Section 38.3.1.

The $((\phi)^2)^2$ field theory for N large. To compare with the situation in the $((\phi)^2)^2$ theory of Section 38.3.2 it is interesting to also write the corresponding gap equation for $d = 5$ in the large N limit. One finds

$$m_T^2 = \frac{N\lambda}{6\Lambda}\left(G_2(m_T, T) - \Omega_5(0)\right).$$

For m_T/T small, one finds

$$(m_T/T)^2 \sim \frac{\zeta(3)}{4\pi^2}\frac{T}{(6\Lambda/N\lambda) + D_5(0)} = \frac{N}{6}\frac{\zeta(3)}{4\pi^2}\lambda_T\,,$$

with

$$\lambda_T = \lambda\frac{T}{\Lambda}\frac{1}{1 + N\lambda D_5(0)/6\Lambda}\,.$$

This behaviour is consistent with the behaviour found in Section 38.3.1 and the behaviour (38.55).

38.5 The Non-Linear σ-Model: Dimensional Reduction

We want now to 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 $\mathbf{S}(t, x)$

$$\mathbf{S}(t, x) = \{\sigma(t, x), \boldsymbol{\pi}(t, x)\},$$

and eliminate locally the field $\sigma(x)$ by

$$\sigma(t, x) = \left(1 - \boldsymbol{\pi}^2(t, x)\right)^{1/2}.$$

One then performs a mode expansion on $\pi(t, x)$, 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 some complications. 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. We explore here 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 37.5.2. Here, instead, we describe a method based on the introduction of an auxiliary field. This allows us to use a more physical momentum cut-off regularization of Pauli–Villars type.

38.5.1 Renormalization group at finite temperature

We again start from the action in the form (30.68), but rescale the coupling constant $g \mapsto g\Lambda^{2-d}$ in such a way that g is now dimensionless. The correlation functions of the $\mathbf{S}$ field satisfy RG equations (Sections 37.1.2, 31.2.2):

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\zeta(g)\right]\Gamma^{(n)}(p_i;\Lambda,g,T) = 0\,. \tag{38.56}$$

For $g < g_c$ (and thus $d > 2$), the solution can be written as

$$\Gamma^{(n)}(p_i;\Lambda,g,T) = m^d(g)M_0^{-n}(g)F^{(n)}\left(p_i/m(g),T/m(g)\right) \tag{38.57}$$

with (using $\beta(g) = (d-2)g + O(g^2)$)

$$M_0(g) = \exp\left[-\frac{1}{2}\int_0^g \frac{\zeta(g')}{\beta(g')}\mathrm{d}g'\right], \tag{38.58}$$

$$m(g) = \frac{1}{\xi(g)} = \Lambda g^{-1/(d-2)}\exp\left[-\int_0^g\left(\frac{1}{\beta(g')} - \frac{1}{(d-2)g'}\right)\mathrm{d}g'\right]. \tag{38.59}$$

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 is 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 $d = 2$,

$$\beta(g) = -\frac{N-2}{2\pi}g^2 + O(g^3)\,,\quad \zeta(g) = \frac{N-1}{2\pi}g + O(g^2), \tag{38.60}$$

and $m(g)$, which now is proportional to the physical mass, is defined by

$$m(g) \propto \Lambda\exp\left[-\int^g\frac{\mathrm{d}g'}{\beta(g')}\right] \;\Rightarrow\; \ln(m/\Lambda) = -\frac{2\pi}{(N-2)g} + O(\ln g). \tag{38.61}$$

At finite temperature, RG equations can also be solved in another way, by introducing the coupling constant g_T at mass scale T:

$$\ln(\Lambda/T) = \int_{g_T}^g \frac{\mathrm{d}g'}{\beta(g')}\,, \tag{38.62}$$

where g_T is a function of g and T only through the combination $m(g)/T$:

$$\ln\left(m(g)/T\right) = -\int^{g_T}\frac{\mathrm{d}g'}{\beta(g')}. \tag{38.63}$$

For $d > 2$ and $g < g_c$ fixed (which implies $m(g) = O(\Lambda)$), equation (38.62) implies that g_T approaches the IR fixed point $g = 0$ at fixed temperature,

$$g_T \propto \left(T/m(g)\right)^{d-2}. \tag{38.64}$$

This is a low temperature regime, where finite temperature effects can be calculated from perturbation theory and renormalization group.

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.

Finally, in two dimensions ($d = 2$), we see from equation (38.63) that g_T goes to zero for $m(g)/T$ small, that is, at high temperature, because $g = 0$ then is a UV fixed point,

$$g_T \sim \frac{2\pi}{(N-2)\ln(m(g)/T)}, \tag{38.65}$$

and this is the limit in which two-dimensional perturbation theory is useful.

38.5.2 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 $\boldsymbol{\varphi}, \rho$ the zero momentum modes and $\mathcal{S}_T(\boldsymbol{\varphi}, \rho)$ the reduced $(d-1)$-dimensional action. At leading order, we find

$$\mathcal{S}_T(\boldsymbol{\varphi}, \rho) = \mathcal{S}(\boldsymbol{\varphi}, \rho)/T. \tag{38.66}$$

The one-loop contribution is

$$\begin{aligned}\mathcal{S}_T^{(1)}(\boldsymbol{\varphi}, \rho) &= \tfrac{1}{2}N \operatorname{tr}\ln(-\nabla^2 + \rho) + \tfrac{1}{2}\operatorname{tr}\ln\left[\boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}\boldsymbol{\varphi}\right] \\ &= \tfrac{1}{2}(N-1)\operatorname{tr}\ln(-\nabla^2+\rho) + \tfrac{1}{2}\operatorname{tr}\ln\left[\boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}\boldsymbol{\varphi}(-\nabla^2+\rho)\right].\end{aligned} \tag{38.67}$$

The form of the last term may surprise, until one remembers that the perturbative expansion is performed around a non-vanishing value of $\boldsymbol{\varphi}$.

We use the identity, obtained after some commutations,

$$\begin{aligned}\boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}\boldsymbol{\varphi}(-\nabla^2+\rho) &= \boldsymbol{\varphi}\cdot\boldsymbol{\varphi} + \boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}[\nabla^2, \boldsymbol{\varphi}] \\ &= \boldsymbol{\varphi}\cdot\boldsymbol{\varphi} + \boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}\left[(\nabla^2\boldsymbol{\varphi}) + 2\nabla\boldsymbol{\varphi}\cdot\nabla\right].\end{aligned}$$

At this order $\boldsymbol{\varphi}\cdot\boldsymbol{\varphi} = 1$ and we expect that ρ can be neglected because it yields interactions of higher dimensions, and thus irrelevant at large distance. The expansion of the $\operatorname{tr}\ln$ then yields a first term with two derivatives and higher orders yield additional derivatives which also are sub-leading at large distance. The first term yields

$$\operatorname{tr}\boldsymbol{\varphi}(-\nabla^2+\rho)^{-1}\left[(\nabla^2\boldsymbol{\varphi}) + 2\nabla\boldsymbol{\varphi}\cdot\nabla\right] \sim \operatorname{tr}(\nabla\boldsymbol{\varphi})^2(-\nabla^2)^{-1},$$

where the relations

$$\boldsymbol{\varphi}\cdot\nabla\boldsymbol{\varphi} = 0, \quad (\nabla\boldsymbol{\varphi})^2 + \boldsymbol{\varphi}\cdot\nabla^2\boldsymbol{\varphi} = 0,$$

valid at leading order, have been used.

In the same way, we expand the first term in (38.67) in powers of the field ρ. At leading order only one term is relevant, and we thus obtain

$$\mathcal{S}_T^{(1)} = \frac{1}{2T}\bar{G}_2(0,T)\int \mathrm{d}^{d-1}x\left[(\partial_\mu\boldsymbol{\varphi}(x))^2 + (N-1)\rho(x)\right], \tag{38.68}$$

where the constant $\bar{G}_2$ defined in (38.22), is given by (equation (38.26))

$$\bar{G}_2(0,T) = f_d(0)T^{d-2} - T\Omega_{d-1}(0) + \Omega_d(0). \tag{38.69}$$

We conclude that at one-loop order

$$\mathcal{S}_T(\varphi,\rho)=\frac{\Lambda^{d-2}}{2gT}\int \mathrm{d}^{d-1}x\,\left[(Z_\varphi/Z_g)(\partial_\mu\varphi(x))^2+\rho(x)\left(\varphi^2(x)-Z_\varphi^{-1}\right)\right] \tag{38.70}$$

with

$$Z_g=1+(N-2)\Lambda^{2-d}\bar{G}_2\,g+O(g^2), \tag{38.71a}$$
$$Z_\varphi=1+(N-1)\Lambda^{2-d}\bar{G}_2\,g+O(g^2). \tag{38.71b}$$

Dimension $d=3$. For $d=3$, the constant $\bar{G}_2$ in equation (38.69) 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:

$$\bar{G}_2=-\frac{T}{2\pi}\left(\ln(\Lambda/T)+\kappa_3\right)+\Omega_3(0),$$

where κ_3 is a constant. The term $\Omega_3(0)$ generates finite renormalizations of g:

$$g_{\mathrm{r}}=g+(N-2)\left(\Omega_3(0)/\Lambda\right)g^2,$$

and of the field φ

$$\varphi=\left[1-\tfrac{1}{2}(N-1)\left(\Omega_3(0)/\Lambda\right)g\right]\varphi_{\mathrm{r}}\,.$$

We now introduce the effective coupling constant

$$g_T=g_{\mathrm{r}}T/\Lambda\,.$$

The effective action becomes

$$\mathcal{S}_T(\varphi_{\mathrm{r}},\rho_{\mathrm{r}})=\frac{1}{2g_T}\int \mathrm{d}^2x\,\left[(\tilde{Z}_\varphi/\tilde{Z}_g)(\partial_\mu\varphi_{\mathrm{r}}(x))^2+\rho_{\mathrm{r}}(x)\left(\varphi_{\mathrm{r}}^2(x)-\tilde{Z}_\varphi^{-1}\right)\right]. \tag{38.72}$$

We verify that the remaining factors $Z_g,\tilde{Z}_\varphi$ render the reduced theory one-loop finite:

$$\tilde{Z}_g=1-\frac{N-2}{2\pi}\left(\ln(\Lambda/T)+\kappa_3\right)g_T+O(g_T^2),$$
$$\tilde{Z}_\varphi=1-\frac{N-1}{2\pi}\left(\ln(\Lambda/T)+\kappa_3\right)g_T+O(g_T^2).$$

The solution of the two-dimensional non-linear σ-model then requires non-perturbative techniques, but the two-dimensional RG tells us

$$\ln(m_T/T)\propto-\frac{2\pi}{(N-2)g_T}=-\frac{2\pi\Lambda}{(N-2)gT}=-\frac{2\pi}{N-2}\frac{m(g)}{T}\,, \tag{38.73}$$

where the last equation involves the three-dimensional RG. The result is consistent with equation (38.52).

Dimension $d=2$. Then,

$$\bar{G}_2=\frac{1}{2\pi}\left[\ln(\Lambda/T)+\kappa_2\right].$$

The reduced one-dimensional theory is of course 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

$$\varphi_{\mathrm{r}} = Z_\varphi^{1/2}\varphi = \left[1 + (N-1)\bar{G}_2 g/2\right]\varphi\,, \tag{38.74a}$$

$$\frac{1}{g_T} = \frac{1}{gZ_g} = \frac{1}{g} - (N-2)\bar{G}_2 + O(g)\,. \tag{38.74b}$$

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 (38.60).

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

$$m_T = \tfrac{1}{2}(N-1)Tg_T\,.$$

Expressing g_T by equation (38.65) in terms of the mass scale (38.61), which is proportional to the physical mass, we obtain

$$\frac{T}{m_T} \sim -\frac{1}{\pi}\frac{N-2}{N-1}\ln(m/T), \tag{38.75}$$

a result consistent with equation (38.48). 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 $m/T \to 0$.

38.5.3 Matching conditions

If the explicit form of the reduced theory can be guessed, another strategy is available, based on matching conditions. 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 for what concerns 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 what follows dimensional regularization is used to avoid the problem of the functional measure: in the absence of a Lagrange multiplier, a Pauli–Villars cut-off does not regularize the $O(N)$ invariant measure. The more "physical" lattice regularization is also available, but explicit calculations are more difficult.

Finally, 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

$$\mathcal{S}(\mathbf{S}) = \frac{\Lambda^{d-2} Z_S}{2gZ_g} \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left(\partial_\mu \mathbf{S}(t,x)\right)^2 - \frac{\Lambda^{d-2}}{g} \int \mathrm{d}t\, \mathrm{d}^{d-1}x\, \mathbf{h}\cdot\mathbf{S}(t,x), \tag{38.76}$$

where the $\overline{\mathrm{MS}}$ scheme is used to define renormalization constants, and

$$\mathbf{S}^2(t,x) = Z_S^{-1}. \tag{38.77}$$

Therefore, g is the effective coupling constant at scale Λ.

RG equations in an external field $\mathbf{h}$ take the form

$$\left[\Lambda\frac{\partial}{\partial\Lambda} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\zeta(g) + \rho(g)h\frac{\partial}{\partial h}\right]\Gamma^{(n)}(p_i;\Lambda,g,h,T) = 0\,, \tag{38.78}$$

where $h = |\mathbf{h}|$ and the new RG function is not independent:

$$\rho(g) = 1 - d + \tfrac{1}{2}\zeta(g) + \beta(g)/g\,. \tag{38.79}$$

Dimensional reduction. We compare the finite temperature correlation functions with those of a zero-temperature $(d-1)$-dimensional euclidean theory

$$\mathcal{S}(\varphi) = \frac{T^{d-3} Z_\varphi}{2g_T\tilde{Z}_g} \int \mathrm{d}^{d-1}x \left(\partial_\mu \varphi(x)\right)^2 - \frac{T^{d-3}}{g_T} \int \mathrm{d}^{d-1}x\, \mathbf{h}\cdot\varphi(x), \tag{38.80}$$

where the $\overline{\mathrm{MS}}$ scheme again is used to define renormalization constants, and

$$\varphi^2(x) = Z_\varphi^{-1}. \tag{38.81}$$

The coupling constant g_T now is the effective coupling constant at the temperature scale T.

We expect that between the two fields $\mathbf{S}$ and φ some finite renormalization will be required.

The one-loop diagrams are listed in figure 31.1. In the reduced model, at one-loop order the two-point function is (Section 31.2.2)

$$\Gamma^{(2)}_{d-1}(p) = \frac{T^{d-3}}{g_T}\left(p^2 Z_\varphi/\tilde{Z}_g + hZ_\varphi^{1/2}\right) + \left[p^2 + \frac{1}{2}(N-1)h\right]\Omega_{d-1}(\sqrt{h}) + O(g_T)\,.$$

At finite temperature, in the d-dimensional theory, one finds instead

$$\Gamma^{(2)}_{d}(p_0=0,p) = \frac{\Lambda^{d-2}}{g}\left(p^2 Z_{\mathbf{S}}/Z_g + hZ_{\mathbf{S}}^{1/2}\right) + \left[p^2 + \frac{1}{2}(N-1)h\right]G_2(\sqrt{h},T) + O(g), \tag{38.82}$$

where the function G_2 is defined in (38.46). In the limit $h=0$,

$$G_2(0,T) = f_d(0)T^{d-2}. \tag{38.83}$$

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

$$\mathcal{S}(\mathbf{S})=\frac{\Lambda^{d-2}Z_S}{2gZ_g}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\big(\partial_\mu \mathbf{S}(t,x)\big)^2-\frac{\Lambda^{d-2}}{g}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\,\mathbf{h}\cdot\mathbf{S}(t,x), \tag{38.76}$$

where the $\overline{\mathrm{MS}}$ scheme is used to define renormalization constants, and

$$\mathbf{S}^2(t,x)=Z_S^{-1}. \tag{38.77}$$

Therefore, g is the effective coupling constant at scale Λ.

RG equations in an external field $\mathbf{h}$ take the form

$$\left[\Lambda\frac{\partial}{\partial\Lambda}+\beta(g)\frac{\partial}{\partial g}-\frac{n}{2}\zeta(g)+\rho(g)h\frac{\partial}{\partial h}\right]\Gamma^{(n)}(p_i;\Lambda,g,h,T)=0\,, \tag{38.78}$$

where $h=|\mathbf{h}|$ and the new RG function is not independent:

$$\rho(g)=1-d+\tfrac{1}{2}\zeta(g)+\beta(g)/g\,. \tag{38.79}$$

Dimensional reduction. We compare the finite temperature correlation functions with those of a zero-temperature $(d-1)$-dimensional euclidean theory

$$\mathcal{S}(\varphi)=\frac{T^{d-3}Z_\varphi}{2g_T\tilde{Z}_g}\int \mathrm{d}^{d-1}x\big(\partial_\mu\varphi(x)\big)^2-\frac{T^{d-3}}{g_T}\int \mathrm{d}^{d-1}x\,\mathbf{h}\cdot\varphi(x), \tag{38.80}$$

where the $\overline{\mathrm{MS}}$ scheme again is used to define renormalization constants, and

$$\varphi^2(x)=Z_\varphi^{-1}. \tag{38.81}$$

The coupling constant g_T now is the effective coupling constant at the temperature scale T.

We expect that between the two fields $\mathbf{S}$ and φ some finite renormalization will be required.

The one-loop diagrams are listed in figure 31.1. In the reduced model, at one-loop order the two-point function is (Section 31.2.2)

$$\Gamma^{(2)}_{d-1}(p)=\frac{T^{d-3}}{g_T}\left(p^2Z_\varphi/\tilde{Z}_g+hZ_\varphi^{1/2}\right)+\left[p^2+\frac{1}{2}(N-1)h\right]\Omega_{d-1}(\sqrt{h})+O(g_T)\,.$$

At finite temperature, in the d-dimensional theory, one finds instead

$$\Gamma^{(2)}_d(p_0=0,p)=\frac{\Lambda^{d-2}}{g}\left(p^2Z_\mathrm{S}/Z_g+hZ_\mathrm{S}^{1/2}\right)+\left[p^2+\frac{1}{2}(N-1)h\right]G_2(\sqrt{h},T)+O(g), \tag{38.82}$$

where the function G_2 is defined in (38.46). In the limit $h=0$,

$$G_2(0,T)=f_d(0)T^{d-2}. \tag{38.83}$$

Dimension $d = 3$. For $d \to 3$, the renormalization constants at one-loop order in the $\overline{\text{MS}}$ scheme are

$$\tilde{Z}_g = 1 + (N-2)\frac{N_d}{d-3} g_T\,, \quad Z_\varphi = 1 + (N-1)\frac{N_d}{d-3} g_T\,. \tag{38.84}$$

In particular,

$$Z_\varphi / \tilde{Z}_g = 1 + \frac{N_d}{d-3} g_T\,.$$

Therefore, the renormalized $(d-1)$-dimensional two-point function reads

$$\Gamma^{(2)}_{d-1}(p) = \frac{1}{g_T}(p^2 + h) + \left[p^2 + \tfrac{1}{2}(N-1)h\right] I_\mathrm{r}(h)$$

with

$$I_\mathrm{r}(h) = \lim_{d\to 3} \Omega_{d-1}(\sqrt{h}) + \frac{N_d}{d-2} T^{d-3} = -\frac{1}{4\pi}\ln(h/T^2)\,. \tag{38.85}$$

In the finite temperature theory no renormalization is required because the theory is non-renormalizable, and dimensional regularization cancels all power divergences. Thus,

$$\Gamma^{(2)}_d(p) \underset{d\to 3}{=} \frac{\Lambda}{g}\,(p^2 + h) + \left[p^2 + \frac{1}{2}(N-1)h\right] G_2(\sqrt{h}, T) + O(g) \tag{38.86}$$

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

$$\Gamma^{(2)}_{d-1}(p) = \frac{1}{Tg_T}(p^2 + h) + \left[p^2 + \tfrac{1}{2}(N-1)h\right]\frac{1}{2\sqrt{h}} + O(g_T).$$

We compare this expression with the finite temperature two-point function, calculated in the $\overline{\text{MS}}$ scheme (with renormalization scale Λ). For this purpose, we have to subtract to expression (38.83) the $\overline{\text{MS}}$ counter-term. For $d \to 2$, we find

$$[G_2]_\mathrm{r}(0, T) = \frac{1}{2\pi}\big(\ln(\Lambda/T) + \gamma - \ln(4\pi)\big), \tag{38.87}$$

and, therefore,

$$\begin{aligned}\Gamma^{(2)}_d(p) = \frac{1}{g}\,(p^2 + h) &+ \left[p^2 + \tfrac{1}{2}(N-1)h\right] \\ &\times \left[\Omega_{d-1}(\sqrt{h}) + \frac{1}{2\pi T}\big(\ln(\Lambda/T) + \gamma - \ln(4\pi)\big)\right] + O(g).\end{aligned} \tag{38.88}$$

In this case, a field renormalization is required. We set

$$\sqrt{T}\varphi(x) = \mathbf{S}(x)\sqrt{Z_{\varphi\mathbf{S}}}\,, \qquad Z_{\mathbf{S}\varphi} = 1 + (N-1)\bigl(\ln(\Lambda/T) + \gamma - \ln(4\pi)\bigr)\frac{g}{2\pi},$$

and

$$\frac{1}{g} = \frac{1}{g_T Z_{gt}}\,, \qquad Z_{gt} = 1 - \bigl(\ln(\Lambda/T) + \gamma - \ln(4\pi)\bigr)\frac{g}{2\pi},$$

or inverting the relation

$$\frac{1}{g_T} = \frac{1}{g} - \frac{(N-2)}{2\pi}\bigl(\ln(\Lambda/T) + \gamma - \ln(4\pi)\bigr) + O(g),$$

a result that can also be obtained by the method of Section 3.4. The results for Z_{gt} and g_T are consistent with equations (38.74). .

38.6 The Gross–Neveu 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 Gross–Neveu (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 31.7):

$$\mathcal{S}(\bar{\boldsymbol{\psi}}, \boldsymbol{\psi}) = -\int \mathrm{d}t\,\mathrm{d}^{d-1}x\left[\bar{\boldsymbol{\psi}}(t,x)\cdot \not{\partial}\boldsymbol{\psi}(t,x) + \frac{1}{2}G\left(\bar{\boldsymbol{\psi}}(t,x)\cdot\boldsymbol{\psi}(t,x)\right)^2\right]. \qquad (38.89)$$

The GN model has in all dimensions a discrete symmetry (with the notation $x_0 \equiv t$)

$$\mathbf{x} = \{x_0, \ldots, x_\mu, \ldots, x_{d-1}\} \mapsto \tilde{\mathbf{x}} = \{x_0, \ldots, -x_\mu, \ldots, x_{d-1}\}, \quad \begin{cases} \psi(\mathbf{x}) \mapsto \gamma_\mu \psi(\tilde{\mathbf{x}}), \\ \bar{\psi}(\mathbf{x}) \mapsto -\bar{\psi}(\tilde{\mathbf{x}})\gamma_\mu\,, \end{cases}$$

which prevents the addition of a mass term. In even dimensions, it implies a discrete chiral symmetry, and in odd dimensions, it corresponds to space reflection. Below, 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 the chiral symmetry is always broken at zero temperature.

It has been proven in Section 31.9.2 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 elementary 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 are simply integrated out. Instead, we study here the GN model within the framework of the $1/N$ expansion. After integration over fermions we obtain the non-local action for a periodic scalar field σ that we have already discussed in the zero-temperature limit in Section 31.9.1.

Additional effects due to the addition of a chemical potential will not be considered here.

In the situations in which the σ mass is small compared with 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 the Ising 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 Ising type. If a transition exists and is continuous, the σ-mass vanishes at the transition and a potential IR problem appears.

After integration over fermions, we obtain a non-local action $\mathcal{S}_N$ for the field σ,

$$\mathcal{S}_N(\sigma) = \frac{1}{2G}\int_0^{1/T} \mathrm{d}t \int \mathrm{d}^{d-1}x\, \sigma^2(t,x) - N \operatorname{tr}\ln\left(\not{\partial} + \sigma(\cdot)\right), \tag{38.90}$$

where T is the temperature, and the σ field satisfies periodic boundary conditions in the euclidean time variable. As we have seen a non-trivial perturbation theory is obtained by expanding for N large.

If a phase transition occurs at finite temperature and if it is second order, IR divergences generated by the σ zero-mode will 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 characterized 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 $d<5$. Instead 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 effective field theory relevant for long 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 renormalization group 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.

38.6.1 The gap equation

Calling M_T the saddle point value of the field $\sigma(x)$ we obtain the action density at finite temperature and large N:

$$\mathcal{E}(M_T) = \frac{M_T^2}{2G} - \frac{NT}{V_{d-1}} \operatorname{tr}\ln(\not{\partial} + M_T), \tag{38.91}$$

where V_{d-1} is the $(d-1)$-dimensional volume,

$$N \operatorname{tr}\ln(\not{\partial} + M_T) = N' \frac{V_{d-1}}{2} \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \sum_{n\in\mathbb{Z}} \ln(\omega_n^2 + k^2 + M_T^2),$$

$N' = N \operatorname{tr}\mathbf{1}$ is the total number of fermions and $\omega_n = (2n+1)\pi T$. The sum over frequencies follows from the identity $(A38.2)$, and one obtains

$$N \operatorname{tr}\ln(\not{\partial} + M_T) = N' V_{d-1} \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\left[2\cosh\bigl(\omega(k)/2T\bigr)\right], \tag{38.92}$$

and $\omega(k) = \sqrt{k^2+M_T^2}$. Alternatively, one could use Schwinger's representation of the propagator and another function of elliptic type

$$\vartheta_1(s) = \sum_n \mathrm{e}^{-(n+1/2)^2\pi s}. \tag{38.93}$$

The gap equation at finite temperature, obtained by differentiating $\mathcal{E}$ with respect to M_T, again splits into two equations $M_T = 0$ and

$$1/G = N'\mathcal{G}_2(M_T, T) \tag{38.94}$$

with

$$\mathcal{G}_2(M_T, T) = \int^\Lambda \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\omega(k)} \left(\frac{1}{2} - \frac{1}{\mathrm{e}^{\omega(k)/T} + 1}\right), \tag{38.95}$$

and $\omega(k) = \sqrt{k^2 + M_T^2}$. This is the fermion analogue of equations (38.45,38.46) 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

$$\begin{aligned} g_d(s) &= N_{d-1} \int_0^\infty \frac{x^{d-2}\,\mathrm{d}x}{\sqrt{x^2+s}} \frac{1}{\exp[\sqrt{x^2+s}] + 1} \\ &= N_{d-1} \int_{\sqrt{s}}^\infty (y^2 - s)^{(d-3)/2} \frac{\mathrm{d}y}{\mathrm{e}^y + 1}, \end{aligned} \tag{38.96}$$

where N_d is given in (38.25), we write the gap equation

$$1/N'G = \Omega_d(M_T) - T^{d-2} g_d(M_T^2/T^2). \tag{38.97}$$

If $d > 2$ we can introduce the critical value G_c where $M \equiv M_{T=0}$ vanishes at zero temperature

$$\frac{1}{N'}\left(\frac{1}{G} - \frac{1}{G_c}\right) = \Omega_d(M_T) - \Omega_d(0) - T^{d-2} g_d(M_T^2/T^2). \tag{38.98}$$

For $G > G_c$, we can introduce the fermion physical mass $M \equiv m_\psi$ solution of

$$\frac{1}{N'}\left(\frac{1}{G} - \frac{1}{G_c}\right) = \Omega_d(M) - \Omega_d(0). \tag{38.99}$$

The gap can then be written as

$$\Omega_d(M_T) - \Omega_d(M) = T^{d-2} g_d(M_T^2/T^2). \tag{38.100}$$

Critical temperature. We note that $\Omega_d(0) - \Omega_d(M_T)$ is always negative. Therefore, the gap equation has no solution for $G \le G_c$, 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 at $N = \infty$.

For $d > 2$, $g_d(0)$ is finite

$$g_d(0) = N_{d-1}\left(1 - 2^{3-d}\right)\Gamma(d-2)\zeta(d-2),$$

and thus, if $G > G_c$, one instead finds always a transition temperature T_c where M_T vanishes, between two Ising-like phases: a low temperature broken phase and a symmetric high temperature phase,

$$T_c = \left[\frac{1}{N' g_d(0)}\left(\frac{1}{G_c} - \frac{1}{G}\right)\right]^{1/(d-2)} = \left[\frac{\Omega_d(0) - \Omega_d(M)}{g_d(0)}\right]^{1/(d-2)}. \tag{38.101}$$

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 $\Delta_\sigma(p) \equiv \Delta_\sigma(p_0 = 0, p)$ (see equation (31.102)). For $M_T = 0$, we find

$$\Delta_\sigma^{-1}(p) = \frac{1}{G} - N'\mathcal{G}_2(0,T) + \frac{N'T}{2(2\pi)^{d-1}} p^2 \sum_n \int^\Lambda \frac{\mathrm{d}^{d-1}k}{(k^2+\omega_n^2)\left[(p+k)^2+\omega_n^2\right]}. \tag{38.102}$$

For $d > 2$, the propagator can be expressed in terms of the constant $g_d(0)$:

$$\begin{aligned}\Delta_\sigma^{-1}(p) = {}& \frac{1}{G} - \frac{1}{G_c} + N'T^{d-2}g_d(0) \\ & + \frac{N'T}{2(2\pi)^{d-1}} p^2 \sum_n \int^\Lambda \frac{\mathrm{d}^{d-1}k}{(k^2+\omega_n^2)\left[(p+k)^2+\omega_n^2\right]}.\end{aligned} \tag{38.103}$$

For $M_T \neq 0$, using the gap equation, we write the propagator

$$\Delta_\sigma^{-1}(p) = \frac{N'T}{2(2\pi)^{d-1}} \left(p^2 + 4M_T^2\right) \sum_n \int^\Lambda \frac{\mathrm{d}^{d-1}k}{(k^2+\omega_n^2+M_T^2)\left[(p+k)^2+\omega_n^2+M_T^2\right]}. \tag{38.104}$$

Therefore, when the symmetry is broken the correlation length is $1/2M_T$, generalizing the zero-temperature result.

More detailed properties require specifying dimensions.

Local expansion. When the σ mass or expectation value are small compared to T we can perform a mode expansion of the field σ, retaining only the zero-mode and then a local expansion of the action (38.90), 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

$$\mathcal{S}_{d-1}(\sigma) = \int \mathrm{d}^{d-1}x \left[\frac{1}{2} Z_\sigma (\partial_\mu \sigma)^2 + \frac{1}{2} r\sigma^2 + \frac{1}{4!} u\sigma^4\right], \tag{38.105}$$

where terms of order σ^6 and $\partial^2\sigma^4$ and higher have been neglected, and the three parameters are given by

$$Z_\sigma = \tfrac{1}{2} N'\mathcal{G}_4(0,T)/T\,, \quad r = \left[1/G - N'\mathcal{G}_2(0,T)\right]/T\,, \quad u = 6N'\mathcal{G}_4(0,T)/T\,,$$

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

$$Tr = \frac{1}{G} - \frac{1}{G_c} - N'T^{d-2}g_d(0) = N'g_d(0)\left(T_c^{d-2} - T^{d-2}\right). \tag{38.106}$$

As 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.

38.6.2 Phase structure

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 σ mass behaves like

$$m_\sigma^2 \propto T^2(T/\Lambda)^{d-4}\left[1-(T_c/T)^{d-2}\right],$$

and thus is small with respect 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 m_σ (inverse correlation length) is

$$m_\sigma^2 \propto \frac{1}{\ln(\Lambda/T)}\left(\frac{1}{G}-\frac{1}{G_c}+N'T^2 g_4(0)\right).$$

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 (38.101))

$$m_\sigma^2 \propto \frac{1}{\ln(\Lambda/T)}(T^2 - T_c^2).$$

Eventually, m_σ vanishes as $(T - T_c)^{1/2}$ a typical mean-field behaviour, and a phase transition occurs. Equation (38.101) 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

$$g_3(s) = \frac{1}{2\pi}\ln\left(1+\mathrm{e}^{-\sqrt{s}}\right),$$

the gap equation can be written as

$$2\cosh(M_T/2T) = \mathrm{e}^{M/2T},$$

an equation that has a 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:

$$T_c/M = \frac{1}{2\ln 2}\,.$$

Dimension $d = 2$. The situation $d = 2$ is doubly special, since 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 $\Lambda(G)$,

$$\Lambda(G) \propto \Lambda \exp\left[-\int^G \frac{\mathrm{d}G'}{\beta(G')}\right].$$

In particular, the fermion physical mass M has the form

$$M \propto \Lambda\, \mathrm{e}^{-2\pi/N'G}\,.$$

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

$$\int_G^{G_T} \frac{\mathrm{d}G'}{\beta(G')} = -\ln(\Lambda/T).$$

At high temperature G_T decreases like

$$G_T \sim \frac{2\pi}{N'\ln(T/M)}.$$

We, therefore, 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 not different from other modes. Eventually, it decreases when T approaches $T_c = M/\pi$. At leading order one finds

$$m_\sigma^2 \propto T^2 \ln(\pi T/M). \tag{38.107}$$

This result does not imply a phase transition since for $m_\sigma/T \ll 1$ dimensional reduction is justified and we are lead to an essentially one-dimensional statistical system with short range interactions that can have no phase transition. Due to 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 $s \to 0$ and goes to 0 for $s \to \infty$. At low temperature M_T/M converges to one exponentially in M/T. At high temperature, M_T/T goes to zero and

$$g_2(s) \sim -\frac{1}{4\pi} \ln s\,.$$

The equation implies $M \propto T$ and thus has no solution for $T \to \infty$, but instead solutions at finite temperature, in agreement with equation (38.107) which shows that m_σ vanishes for some value $T_c \propto M$. The existence of non-trivial solutions to the gap equation here implies only 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 (38.105) to the $d = 2$ example. We find a simple model in 1D 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 happens only 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. Calculating the difference between the two lowest eigenvalues of the hamiltonian one obtains a behaviour of the form

$$m_\sigma/T \propto (\ln M/T)^{5/4}\, \mathrm{e}^{-\mathrm{const.}(\ln M/T)^{3/2}}\,.$$

Again, the property that m_σ/T is small at low temperature is a precursor of the symmetry breaking at zero temperature.

38.7 Abelian Gauge Theories

We first discuss the abelian case which is much simpler, because the mode decomposition is consistent with the gauge structure. Some additional problems arising in non-abelian gauge theories will be considered in Section 38.8. Because the gauge field has a number of components which depends on the number of space dimensions, the mode expansion have some new properties and affects 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.

Notation. In what follows greek indices refer to space–time components while latin indices refer to space components only.

38.7.1 Mode expansion and gauge transformations

We decompose a general gauge field $A_\mu(t,x)$ into the sum of a zero-mode $B_\mu(x)$ and the sum of all non-zero modes $Q_\mu(t,x)$:

$$A_\mu(t,x) = B_\mu(x) + Q_\mu(t,x).$$

At finite temperature $T > 0$, $Q_\mu(t,x)$ thus satisfies ($\beta = 1/T$)

$$\int_0^\beta \mathrm{d}t\, Q_\mu(t,x) = 0\,. \tag{38.108}$$

With this decomposition, gauge transformations

$$\delta A_\mu(t,x) = \partial_\mu \varphi(t,x)$$

become

$$\delta B_\mu(x) = \partial_\mu \varphi_0(t,x), \quad \delta Q_\mu(t,x) = \partial_\mu \varphi_1(t,x)\,, \quad \varphi = \varphi_0 + \varphi_1\,. \tag{38.109}$$

Since δB_μ does not depend on t we conclude that $\varphi_0(t,x)$ must have the special form

$$\varphi_0(t,x) = F(x) + \Omega t\,, \tag{38.110}$$

where Ω is a constant. 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

$$\delta B_0(x) = \Omega\,. \tag{38.111}$$

Invariance under the translation (38.111) implies, in the absence of matter fields, that the scalar field B_0 is massless.

The condition (38.108) then implies

$$\partial_i \int_0^\beta \mathrm{d}t\, \varphi_1(t,x) = 0\,, \quad \int_0^\beta \mathrm{d}t\, \partial_t \varphi_1(t,x) = 0 \ \Rightarrow\ \varphi_1(0,x) = \varphi_1(\beta,x).$$

The transformations of the gauge field Q_μ are thus specified by periodic functions $\varphi_1(t,x)$ with a constant zero-mode, which can be set to zero.

Finally, we verify that the function $\varphi = \varphi_1 + \varphi_0$ is such that $\partial_\mu \varphi$ is periodic as it should, since A_μ is periodic.

Matter fields. We now 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. To the gauge transformation (38.109) corresponds for the fermions:

$$\psi(t,x) = \mathrm{e}^{i\varphi(t,x)}\, \psi'(t,x), \quad \bar\psi(t,x) = \mathrm{e}^{-i\varphi(t,x)}\, \bar\psi'(t,x).$$

Anti-periodicity implies that

$$\varphi(\beta,x) = \varphi(0,x) \pmod{2\pi}\,.$$

Since φ_1 is periodic, this condition implies for the constant Ω in (38.110),

$$\Omega = 2n\pi T\,. \tag{38.112}$$

This restriction on the transformation (38.111) of the scalar 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 remain massless. Instead, the thermodynamic potential for constant fields is a periodic function of B_0 with period $2\pi T$.

38.7.2 Gauge field coupled to fermions: quantization

We now consider a gauge field coupled to an N-component charged fermion:

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\tfrac{1}{4e^2} F_{\mu\nu}^2(t,x) - \bar{\psi}(t,x) \cdot (\not{\partial} + i\not{A})\, \psi(t,x)\right]. \tag{38.113}$$

We have neglected the fermion mass because we are interested only in high temperature physics. The theory has RG properties which bear some similarities with 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.

The temporal gauge. To calculate the partition function, we first quantize in the temporal gauge $A_0(t,x) = 0$ because the corresponding hamiltonian formalism is simple. The action becomes

$$\mathcal{S}(\bar{\psi}, \psi, A_\mu) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\tfrac{1}{4e^2}\left(2\dot{A}_i^2 + F_{ij}^2(t,x)\right) - \bar{\psi}(t,x) \cdot (\not{\partial} + i\not{A})\, \psi(t,x)\right]. \tag{38.114}$$

In the calculation of the partition function $\operatorname{tr} \mathrm{e}^{-H/T}$, 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:

$$\begin{aligned} A_i(\beta, x) &= A_i(0,x) - \beta \partial_i \varphi(x), \\ \psi(\beta, x) &= \mathrm{e}^{i\beta\varphi(x)}\, \psi(0,x), \end{aligned}$$

and integrate over the gauge transformation $\varphi(x)$. We then set

$$A_i(t,x) = A_i'(t,x) - t\partial_i \varphi(x),$$

and correspondingly

$$\psi(t,x) = \mathrm{e}^{it\varphi(x)}\, \psi'(t,x), \quad \bar{\psi}(t,x) = \mathrm{e}^{-it\varphi(x)}\, \bar{\psi}'(t,x),$$

where the fields A_i', $\psi', \bar{\psi}'$ are now periodic and anti-periodic, respectively. This induces two modifications in the action

$$\begin{aligned} \int \mathrm{d}t\, \mathrm{d}x\, (\partial_t A_i)^2 &\mapsto \int \mathrm{d}t\, \mathrm{d}x\, (\partial_t A_i)^2 + \beta \int \mathrm{d}x\, (\partial_i \varphi(x))^2, \\ \int \mathrm{d}t\, \mathrm{d}x\, \bar{\psi}(t,x)\gamma_0 \partial_t \psi(t,x) &\mapsto \int \mathrm{d}t\, \mathrm{d}x\, \bar{\psi}(t,x)\gamma_0(\partial_t + i\varphi(x))\psi(t,x). \end{aligned}$$

Therefore, $\varphi(x)$ is simply the residual zero-mode of the A_0 component:

$$\varphi(x) \equiv B_0(x).$$

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 which appear in perturbation calculations with the temporal gauge (the gauge field propagator is singular at $k_0 = 0$, see Section 18.4.2) 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 introduce a renormalizable gauge.

Covariant gauge. The change of gauges follows the standard zero-temperature method (Section 19.3). We first introduce a time component A_0 for the gauge field (periodic in time) and multiply the functional measure by the corresponding δ-function

$$1 = \int [\mathrm{d}Q_0(t,x)] \prod_{t,x} \delta(Q_0).$$

The action can then be written in a gauge invariant form. To change to a covariant gauge, we then introduce a second identity in the functional integral:

$$1 = \det(-\partial^2) \int [\mathrm{d}\varphi_1] \delta\big(\partial_\mu Q_\mu + \partial^2 \varphi_1 - n(t,x)\big), \tag{38.115}$$

where φ_1 and $n(t,x)$ are two periodic functions without zero-mode. We perform the gauge transformation

$$Q_\mu + \partial_\mu \varphi_1 \mapsto Q_\mu \,.$$

The φ_1 dependence remains only in $\delta(Q_0 - \partial_t \varphi_1)$, and the integration over φ_1 yields a constant. Integration over $n(t,x)$ with a gaussian weight yields the standard covariant gauge

$$\mathcal{S}_{\text{gauge}} = \frac{1}{2\xi T} \int \mathrm{d}x \big(\partial_i B_i(x)\big)^2 + \frac{1}{2\xi} \int \mathrm{d}t\, \mathrm{d}x (\partial_\mu Q_\mu)^2 \equiv \frac{1}{2\xi} \int \mathrm{d}t\, \mathrm{d}x (\partial_\mu A_\mu)^2 .$$

since

$$\partial_\mu A_\mu = \partial_t Q_0(t,x) + \partial_i B_i(x) + \partial_i Q_i(t,x).$$

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.

Note that the transformation from the temporal gauge to the covariant gauge generates a determinant (equation (38.115)) which is field-independent, but contributes to the free energy.

38.7.3 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 we find the effective action

$$\mathcal{S}_T(B) = \frac{1}{T} \int \mathrm{d}^{d-1}x \left[\frac{1}{2e^2} (\partial_i B_0)^2 + \frac{1}{4e^2} F_{ij}^2(B)\right] - N \operatorname{tr} \ln (\not\partial + i \not B) \,. \tag{38.116}$$

An important issue is the behaviour of the induced mass of the time component $B_0 = \varphi$ of the gauge field. We thus first calculate the action density for constant φ.

The action density. The action density as a function of a constant field $\varphi \equiv B_0$ then is given by

$$\mathcal{E}(\varphi) = -\frac{1}{2}N'T\sum_n \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\left[k^2 + \left(\varphi + (2n+1)\pi T\right)^2\right],$$

where $N' = N \operatorname{tr} \mathbf{1}$ is the total number of fermion degrees of freedom.

The sum over n can be performed with the help of the identity ($A38.2$) and one obtains

$$\mathcal{E}(\varphi) = -\frac{1}{2}N'T\int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\bigl(\cosh(\beta k) + \cos(\beta\varphi)\bigr). \tag{38.117}$$

We verify that the difference $\mathcal{E}(\varphi) - \mathcal{E}(0)$ is UV finite and has a scaling form $T^d f(\varphi/T)$. This is not surprising since in the zero-temperature limit no gauge field mass or quartic φ potential are generated.

The derivative

$$\mathcal{E}'(\varphi) = \frac{1}{2}N' \sinh(\beta\varphi)\frac{1}{(2\pi)^{d-1}}\int \frac{\mathrm{d}^{d-1}k}{\cosh(\beta k) + \cos(\beta\varphi)},$$

is negative for $-\pi < \beta\varphi < 0$ and positive for $0 < \beta\varphi < \pi$. The action density has a unique minimum at $\varphi = 0$ in the interval $-\pi < \beta\varphi < \pi$.

A special case is $d = 2$ for which one finds

$$\mathcal{E}'(\varphi) = \tfrac{1}{2}N'\varphi/\pi \quad \text{for } |\varphi| < \pi\,, \quad \text{and thus} \quad \mathcal{E}(\varphi) = \tfrac{1}{4}N'\varphi^2/\pi\,.$$

Neglecting all φ derivatives, one obtains a contribution to the action $\mathcal{S}_T$:

$$-N \operatorname{tr}\ln(\not{\partial} + i\not{B}) \sim -\frac{1}{2}N'\int \mathrm{d}^{d-1}x \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\bigl(\cosh(\beta k) + \cos(\beta\varphi(x))\bigr).$$

The coefficient $K_2(d)$ of $\frac{1}{2}\int \mathrm{d}x\, \varphi^2(x)$ follows

$$\begin{aligned} K_2(d) &= N' N_{d-1}\Gamma(d-1)\left(1 - d^{3-d}\right)\zeta(d-2)T^{d-3} \\ &= N'\frac{8}{(4\pi)^{d/2}}\Gamma(d/2)\left(2^{d-3} - 1\right)\zeta(d-2)T^{d-3}. \end{aligned}$$

Discussion. At leading order we thus obtain a mass term which is 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,

$$m_\varphi^2 \propto \frac{T^2}{\ln(\Lambda/T)}.$$

The separation between zero and non-zero modes remains justified. High temperature QED shares some properties of high temperature ϕ^4 field theory, and a perturbative expansion for the same reason remains meaningful.

Note that if one is interested in IR physics only, one can in a second step integrate over the massive scalar field φ.

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, we still have to fix the gauge, using for instance a covariant gauge. For what concerns the massive modes, they are here quantized 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 we can make an independent gauge transformation on each vector field. Furthermore, we note that the vector masses are not renormalized.

For more details and more systematic QED calculations we refer to the literature.

38.7.4 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 to study properties of the electro-weak phase transition at finite temperature. The gauge action reads

$$\mathcal{S}(A_\mu,\phi) = \int \mathrm{d}t\,\mathrm{d}^{d-1}x \left[\frac{1}{4e^2}F_{\mu\nu}^2 + |\mathrm{D}_\mu\phi|^2 + U(|\phi|^2)\right], \tag{38.118}$$

where the quartic potential $U(|\phi|^2)$:

$$U(z) = rz + \tfrac{1}{6}gz^2, \tag{38.119}$$

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. Below we use instead the temporal gauge because the unitary gauge becomes singular near the phase transition.

One limitation of the model is that RG shows that in $3+1$ dimensions the hypothesis of second order phase transition is inconsistent, and, therefore, the transition is most likely weak first order. Indeed, in a more general model with N charged scalars for $d=4$ the RG β-functions are (equation (18.134))

$$\beta_g = \frac{1}{24\pi^2}\left[(N+4)g^2 - 18ge^2 + 54e^4\right], \quad \beta_{e^2} = \frac{1}{24\pi^2}Ne^4. \tag{38.120}$$

Studying the RG flow, one verifies that the origin $e^2 = g = 0$ is a stable IR fixed point only for $N \geq 183$. For N small, the continuum model remains meaningful 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. The transition then is weak first order.

Presumably, the same result applies for small values of N to the three-dimensional classical 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 (38.112) 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

$$A_i = B_i + Q_i\,, \quad \phi = \varphi + \chi\,.$$

Neglecting all non-zero modes, we obtain the reduced action at leading order:

$$\mathcal{S}_T(B_\mu, \varphi) = \frac{1}{T}\int \mathrm{d}^{d-1}x \left[\frac{1}{2e^2}(\partial_i B_0)^2 + \frac{1}{4e^2}F_{ij}^2(B)\right.$$
$$\left. + |\mathrm{D}_i\varphi|^2 + |\varphi|^2 B_0^2 + U(|\varphi|^2)\right], \tag{38.121}$$

where covariant derivative and curvature now refer to the gauge field B_i,

$$\mathrm{D}_i = \partial_i + iB_i \quad \text{and} \quad F_{ij}(B) = [D_i, D_j].$$

If φ 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 φ constant.

For the contribution to the B_0 mass the relevant quadratic action is

$$\mathcal{S}_2(B_0) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x\, |\mathrm{D}_\mu \chi|^2,$$

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:

$$\int \mathrm{d}^{d-1}x \sum_{n\neq 0} \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\left[k^2 + \left(2\pi nT + B_0(x)\right)^2\right] = \int \mathrm{d}^{d-1}x \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}$$
$$\times \left[\ln\left(\cosh(\beta k) - \cos(\beta B_0)\right) - \ln(k^2 + B_0^2)\right].$$

The zero-mode subtraction is a one-loop counter-term that can be omitted. The contribution $\delta m_{B_0}^2$ to the coefficient of $\frac{1}{2}B_0^2$ then is

$$\delta m_{B_0}^2 = 2N_{d-1}T^{d-3}\Gamma(d-1)\zeta(d-2)\int \mathrm{d}^{d-1}x\, B_0^2(x).$$

In particular, for $d = 4$, one finds

$$\delta m_{B_0}^2 = \tfrac{1}{3} e^2 T^2 .$$

The quadratic terms in the action relevant for the φ mass shift are

$$\mathcal{S}_2(\varphi) = \int \mathrm{d}t\, \mathrm{d}^{d-1}x \left[\frac{1}{2e^2}(\partial_t Q_i)^2 + \frac{1}{4e^2} F_{ij}^2(Q) + |\partial_t \chi|^2 + |\partial_i \chi + iQ_i \varphi|^2 + \tfrac{2}{3} g |\varphi|^2 |\chi|^2 \right].$$

The integration over χ then yields two contributions, one proportional to $g|\varphi|^2$, and another one that adds to the Q action,

$$|\varphi|^2 \sum_{n \neq 0} \int \mathrm{d}^{d-1}k\, Q_i(\omega_n, k) \left(\delta_{ij} - \frac{k_i k_j}{k^2 + \omega_n^2}\right) Q_j(-\omega_n, -k),$$

($\omega_n = 2\pi n T$). Finally, the integration over Q yields a contribution proportional to e^2 to the coefficient of $|\varphi|^2$. The total contribution δr is

$$\delta r = \bar{G}_2(0, T)\big((d-1)e^2 + 2g/3\big),$$

where $\bar{G}_2$ is defined by equation (38.22). For $d = 4$, the finite part of $\bar{G}_2$ is $T^2/12$ (Section 38.3.2) and, therefore,

$$\delta r = \frac{T^2}{36}(9e^2 + 2g).$$

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, let us point out that the coefficient of $|\varphi|^2 B_0^2$ gets renormalized. At one-loop one finds

$$1 \mapsto 1 + \frac{e^2 + g}{12\pi^2}.$$

(iii) We note that the coefficient of $|\varphi^2|$ 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 (38.119)) 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,

$$\tilde{\mathcal{S}}_T(B_i, \varphi) = \frac{1}{T} \int \mathrm{d}^{d-1}x \left[\frac{1}{4\tilde{e}^2} F_{ij}^2 + |\mathrm{D}_i \varphi|^2 + \tilde{U}(|\varphi|^2)\right], \qquad (38.122)$$

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.

38.8 Non-Abelian Gauge Theories

Non-abelian gauge theories with a small number of fermions are UV asymptotically free in four dimensions (Section 20.2.2). 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.

Notation. We consider fields $\mathbf{A}_\mu$ written as antihermitian or antisymmetric 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 compact group G, in some representation

$$\operatorname{tr} \tau^\alpha \tau^\beta = \delta_{\alpha\beta} , \quad [\tau^\alpha, \tau^\beta] = i f_{\alpha\beta\gamma} \tau^\gamma , \tag{38.123}$$

and the structure constants $f_{\alpha\beta\gamma}$ are chosen antisymmetric.

Gauge transformations take the form

$$\mathbf{A}'_\mu(t,x) = \mathbf{g}(t,x)\mathbf{A}_\mu(t,x)\mathbf{g}^{-1}(t,x) + \mathbf{g}(t,x)\partial_\mu \mathbf{g}^{-1}(t,x), \tag{38.124}$$

where $\mathbf{g}$ is a group element in the matrix representation.

Covariant derivatives $\mathbf{D}_\mu$ act on fields φ transforming under the adjoint representation as

$$\mathbf{D}_\mu \varphi = \partial_\mu \varphi + [\mathbf{A}_\mu, \varphi] . \tag{38.125}$$

The corresponding curvature tensor $\mathbf{F}_{\mu\nu}(t,x)$ is

$$\mathbf{F}_{\mu\nu}(t,x) = [\mathbf{D}_\mu, \mathbf{D}_\nu] = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + [\mathbf{A}_\mu, \mathbf{A}_\nu] . \tag{38.126}$$

In what follows, we discuss for simplicity only the pure gauge action which reads

$$\mathcal{S}(\mathbf{A}) = -\frac{1}{4g^2} \operatorname{tr} \int \mathrm{d}t\, \mathrm{d}^{d-1}x\, \mathbf{F}_{\mu\nu}^2(t,x). \tag{38.127}$$

38.8.1 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 which enforces Gauss's law:

$$\mathbf{A}_i(\beta,x) = \mathbf{g}(x)\mathbf{A}_i(0,x)\mathbf{g}^{-1}(x) + \mathbf{g}(x)\partial_i \mathbf{g}^{-1}(x).$$

We parametrize the group element $\mathbf{g}$ in terms of an element φ of the Lie algebra:

$$\mathbf{g}(x) = \mathrm{e}^{\varphi(x)/T},$$

and introduce

$$\mathbf{g}(t,x) = \mathrm{e}^{t\varphi(x)},$$

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 $\mathbf{A}'_0$ is simply

$$\mathbf{A}_0 = 0 \ \Rightarrow\ \mathbf{A}'_0(t,x) = \mathbf{g}^{-1}(t,x)\partial_t\mathbf{g}(t,x) = \varphi(x).$$

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 (38.127) then reads (omitting now the primes)

$$\begin{aligned}\mathcal{S}(\mathbf{A},\varphi) &= -\frac{1}{2g^2}\operatorname{tr}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\left(\partial_i\varphi(x) - \partial_t\mathbf{A}_i + [\mathbf{A}_i,\varphi]\right)^2 - \frac{1}{4g^2}\operatorname{tr}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\,\mathbf{F}_{ij}^2\\ &= -\frac{1}{2g^2}\operatorname{tr}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\,(\mathbf{D}_i\varphi - \partial_t\mathbf{A}_i)^2 - \frac{1}{4g^2}\operatorname{tr}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\,\mathbf{F}_{ij}^2.\end{aligned}$$

Mode expansion. We now 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

$$\mathbf{Q}_i(t,x) = \sum_{n\neq 0} \mathrm{e}^{2i\pi nTt}\,\mathbf{Q}_{n,i}(x).$$

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 $4\pi^2n^2T^2$, 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. A way to solve this problem is to go over to a covariant gauge for the field $\mathbf{A}_\mu$. 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 19.3.2). 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 which satisfy periodic boundary conditions, unlike ordinary fermions.

38.8.2 Dimensional reduction

Neglecting all non-zero modes, we obtain the reduced action at leading order:

$$\mathcal{S}_T(\mathbf{B},\varphi) = -\frac{1}{4Tg^2}\operatorname{tr}\int \mathrm{d}^{d-1}x\,\left[2(\mathbf{D}_i(\mathbf{B})\varphi)^2 + \mathbf{F}_{ij}^2(\mathbf{B})\right],$$

that is, the action of a massless scalar field coupled to a gauge field.

One-loop calculation of the effective φ potential. Corrections generated by the integration over non-zero modes give a mass to the scalar field φ, as in the abelian example. We verify here this property at one-loop order. Since we need only the terms without derivatives we can treat φ as a constant. Then, omitting the massless gauge field we find a simplified action.

We then expand the fields, using the generators (38.123),

$$\mathbf{Q}_i = iQ_i^\alpha\tau^\alpha\,,\quad \varphi = i\varphi^\alpha\tau^\alpha.$$

The relevant Q action can be written as

$$\mathcal{S}_2(Q) = \frac{1}{2g^2}\int \mathrm{d}t\,\mathrm{d}^{d-1}x\left[\frac{1}{2}\left(\partial_i Q_j^\alpha - \partial_j Q_i^\alpha\right)^2 + \left(\partial_t Q_i^\alpha + f_{\alpha\beta\gamma}Q_i^\beta\varphi^\gamma\right)^2\right].$$

The integration yields a determinant which generates an additive contribution to the effective action. Introducing the antisymmetric matrix Φ with elements (the element φ of the Lie algebra in the adjoint representation)

$$\Phi^{\alpha\beta} = f_{\alpha\beta\gamma}\varphi^\gamma,$$

we can write it as

$$\mathcal{S}_T^{(1)}(\varphi) = \frac{1}{2}\sum_{n\neq 0}\operatorname{tr}\ln\left[\left(k^2\delta_{ij} - k_ik_j\right)\mathbf{1} + \delta_{ij}\left(\omega_n\mathbf{1} + i\Phi\right)^2\right] - (\varphi = 0) \quad (38.128)$$

with $\omega_n = 2\pi nT$. The result is the sum of two contributions, along k and transverse to k. The longitudinal contribution $\Sigma(\varphi)$ is

$$\Sigma(\varphi) = \sum_x\sum_{n\neq 0}\operatorname{tr}\ln\left(\mathbf{1} + i\Phi/\omega_n\right) = \ln\prod_x\det\left[2T\Phi^{-1}\sinh(\Phi/2T)\right]. \quad (38.129)$$

This term contributes to the φ integration measure and yields a factor at each point x:

$$\prod_x \mathrm{d}\varphi(x)\det\frac{\Phi(x)/2T}{\sinh\left(\Phi(x)/2T\right)},$$

which cancels the invariant group measure in the φ parametrization (see Section A38.2).

The transverse contribution to the action density is

$$\mathcal{E}(\varphi) = \frac{1}{2}(d-2)T\sum_{n\neq 0}\int\frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}}\ln\det\left[k^2\mathbf{1} + (\omega_n\mathbf{1} + i\Phi)^2\right] - (\varphi = 0).$$

In terms of the imaginary eigenvalues of the antisymmetric matrix Φ one obtains a sum of terms of the form (38.117), 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 $\varphi = |\varphi|$,

$$\begin{aligned}\mathcal{E}(\varphi) &= (d-2)T\sum_{n\neq 0}\int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \ln\left[k^2 + (\omega_n+\varphi)^2\right] - (\varphi = 0),\\ &= (d-2)T\int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \left[\ln\bigl(\cosh(\beta k) - \cos(\beta\varphi)\bigr) - \ln(k^2+\varphi^2)\right] - (\varphi = 0).\end{aligned}$$

Thus,

$$\mathcal{E}'(\varphi) = (d-2)\int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \left[\frac{\sin(\beta\varphi)}{\cosh(\beta k) - \cos(\beta\varphi)} - \frac{2T\varphi}{k^2+\varphi^2}\right].$$

The subtracted zero-mode contribution acts as a one-loop counter-term, and can be omitted. Then, as in the QED example $\varphi = 0$ is the minimum of the potential (which is also periodic in φ), and the generated mass term is UV finite:

$$m_T^2 = 2(d-2)\frac{\Gamma(d/2)}{\pi^{d/2}}\zeta(d-2)g^2T^{d-2},$$

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 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.

Bibliographical Notes

Cosmological problems have initially motivated the study of finite temperature quantum field theory

D.A. Kirznits *JETP Lett.* 15 (1972) 529; D.A. Kirznits and A.D. Linde, *Phys. Lett.* B42 (1972) 471; *Ann. Phys.* 101 (1976) 195; A.D. Linde *Rep. Prog. Phys.* 42 (1979) 389.

The possibility of phase transitions in heavy ion collisions has generated additional interest

C.W. Bernard, *Phys. Rev.* D9 (1974) 3312 S. Weinberg, *Phys. Rev.* D9 (1974) 3357; L. Dolan and R. Jackiw, *Phys. Rev.* D9 (1974) 3320; D.J. Gross, R.D. Pisarski and L.G. Yaffe, *Rev. Mod. Phys.* 53 (1981) 43.

Among the various reviews and textbooks see, for example,
N.P. Landsman and C. van Weert, *Phys. Rep.* 145 (1987) 141; J. I. Kapusta, *Finite Temperature Field Theory*, Cambridge Univ. Press (Cambridge 1989); M. Le Bellac, *Thermal Field Theory*, Cambridge Univ. Press (Cambridge 1996); H. Meyer-Ortmanns, *Rev. Mod. Phys.* 68 (1996) 473; J.P. Blaizot, E. Iancu, *Phys. Rep.* 359 (2002) 355, hep-ph/0101103.

Schwinger's closed path formalism in the path integral formulation is presented in
K.-C. Chou, Z.-B. Su, B.-L. Hao and L. Yu, *Phys. Rep.* 118 (1985) 1.

RG equations in the context of finite temperature dimensional reduction are discussed in
N.P. Landsman, *Nucl. Phys.* B322 (1989) 498;

Early articles on dimensional reduction include
P. Ginsparg, *Nucl. Phys.* B170 (1980) 388; T. Appelquist and R.D. Pisarski, *Phys. Rev.* D23 (1981) 2305; S. Nadkarni, *Phys. Rev.* D27 (1983) 917, *ibidem* D38 (1988) 3287; E. Braaten, *Phys. Rev. Lett.* 74 (1995) 2164; K. Kajantie, K. Rummukainen and M.E. Shaposhnikov, *Nucl. Phys.* B407 (1993) 356.

For a review see for example
M.E. Shaposhnikov, *Proc. Int. School of Subnuclear Phys.* (Erice 1996), World Scientific, hep-ph/9610247.

The use of matching conditions is advocated in
E. Braaten and A. Nieto, *Phys. Rev.* D51 (1995) 6990; K. Kajantie, M. Laine, K. Rummukainen, M. Shaposhnikov, *Nucl. Phys.* B458 (1996) 90.

Finite temperature calculations in QED and QCD are reported in
P. Arnold and C. Zhai, *Phys. Rev.* D51 (1995) 1906,
and in the framework of the electro-weak transition in
Z. Fodor and A. Hebecker, *Nucl. Phys.* B432 (1994) 127; P. Arnold and O. Espinosa, *Phys. Rev.* D47 (1993) 3546.

The Schwinger model is solved in
A.V. Smilga, *Phys. Lett.* B278 (1992) 371.

For a review of thermal QCD see, for instance,
A.V. Smilga, *Phys. Rep.* 291 (1997) 1.

For the abelian Higgs model see for instance
K. Farakos, K. Kajantie, K. Rummukainen, M. Shaposhnikov, *Nucl. Phys.* B425 (1994) 67; *ibidem* 442 (1995) 317.

Partial summation scheme are discussed in
J.-P. Blaizot, E. Iancu, A. Rebban, *Phys. Rev.* D63 (2001) 065003; J.O. Andersen, E. Braaten, M. Strickland, *Phys. Rev.* D63 (2001) 105008.

APPENDIX A38
FEYNMAN DIAGRAMS AT FINITE TEMPERATURE

In this 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.

A38.1 One-Loop Calculations

We give here some technical details about explicit calculations of one-loop diagrams.

A38.1.1 General remarks

Let us add a few remarks concerning the calculation of Feynman diagrams. General methods explained in the framework of finite size scaling can also be used here, involving 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 $\Delta(t,p)$ of the harmonic oscillator with frequency $\omega(p)=\sqrt{p^2+m^2}$ and time interval $\beta=1/T$:

$$\frac{1}{p^2+\omega^2+m^2}\mapsto\Delta(t,p)=\frac{1}{2\omega(p)}\frac{\cosh\left[(\beta/2-|t|)\omega(p)\right]}{\sinh\left(\beta\omega(p)/2\right)}.$$

Summing over all frequencies is equivalent to set $t=0$. For the simple one-loop diagram, one finds

$$\frac{1}{2\pi}\int\frac{d\omega}{\omega^2+p^2+m^2}\mapsto\frac{1}{2\omega(p)}\frac{\cosh\left(\beta\omega(p)/2\right)}{\sinh\left(\beta\omega(p)/2\right)},$$

This expression can be written in a way that separates quantum and thermal contributions:

$$\frac{1}{2\omega(p)\tanh(\beta\omega(p)/2)}=\frac{1}{2\omega(p)}+\frac{1}{\omega(p)(e^{\beta\omega(p)}-1)},$$

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,

$$\sum_n\frac{x}{(n+\nu)^2+x^2}=\pi\frac{\sinh(2\pi x)}{\cosh(2\pi x)-\cos(2\pi\nu)},\qquad (A38.1)$$

and thus after integration,

$$\ln\left(\cosh(2\pi x)-\cos(2\pi\nu)\right)-\ln 2=\lim_{N\to\infty}\sum_{n=-N}^{N}\left[\ln\left((n+\nu)^2+x^2\right)-2\ln(n+1/2)\right].\qquad (A38.2)$$

A38.1.2 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

$$\sqrt{\pi}\,\Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z+1/2)\,, \quad \Gamma(z)\Gamma(1-z)\sin(\pi z) = \pi\,. \tag{A38.3}$$

These identities imply for the $\psi(z)$ function, $\psi(z) = \Gamma'(z)/\Gamma(z)$,

$$2\psi(2z) = 2\ln 2 + \psi(z) + \psi(z+1/2), \quad \psi(z) - \psi(1-z) + \pi/\tan(\pi z) = 0\,. \tag{A38.4}$$

We also need Riemann's ζ function

$$\zeta(s) = \sum_{n\ge 1} \frac{1}{n^s}\,. \tag{A38.5}$$

It satisfies the reflection formula

$$\zeta(s)\Gamma(s/2) = \pi^{s-1/2}\Gamma\big((1-s)/2\big)\zeta(1-s), \tag{A38.6}$$

which can be written in different forms using Γ-function relations. Moreover,

$$\sum_{n=1} \frac{(-1)^n}{n^s} = (2^{1-s}-1)\zeta(s), \tag{A38.7}$$

Finally ($\gamma = -\psi(1)$),

$$\zeta(1+\varepsilon) = 1/\varepsilon + \gamma + O(\varepsilon), \tag{A38.8}$$

$$\zeta(\varepsilon) = -\tfrac{1}{2}(2\pi)^{\varepsilon} + O(\varepsilon^2). \tag{A38.9}$$

Jacobi's θ*-functions.* In Chapter 37, we have already defined the function $\vartheta_0(s)$ (equation (37.42)), related to Jacobi's elliptic functions (see below):

$$\vartheta_0(s) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-\pi s n^2},$$

and proven in Appendix A37.2 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

$$\vartheta_2(s;\nu,\lambda) = \mathrm{e}^{-\pi s\nu^2}\,\theta_3(\lambda + i\nu s, \mathrm{e}^{-\pi s}) = \sum_n \mathrm{e}^{-\pi s(n+\nu)^2 + 2i\pi n\lambda}, \tag{A38.10}$$

where θ_3 is an elliptic Jacobi's function, which, applying Poisson's formula, can be shown to satisfy

$$\vartheta_2(s;\nu,\lambda) = \vartheta_2(s;-\nu,-\lambda) = s^{-1/2}\vartheta_2(1/s;\lambda,-\nu). \tag{A38.11}$$

For fermions the relevant function is $\vartheta_1(s)$

$$\vartheta_1(s) \equiv \vartheta_2(s;1/2,0) = \sum_n \mathrm{e}^{-(n+1/2)^2\pi s}\,. \tag{A38.12}$$

From ($A38.11$), we obtain

$$\vartheta_1(s) = \frac{1}{\sqrt{s}}\sum_n (-1)^n\,\mathrm{e}^{-\pi n^2/s}\,. \tag{A38.13}$$

A38.2 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 38.8.

We set

$$\mathbf{g} = \mathrm{e}^{\boldsymbol{\xi}},$$

and we will determine the invariant measure in terms of the components ξ^α:

$$\boldsymbol{\xi} = i\tau^\alpha \xi^\alpha .$$

We thus introduce a time-dependent group element $\mathbf{g}$ as

$$\mathbf{g}(t) = \mathrm{e}^{t\boldsymbol{\xi}}, \quad \mathbf{g}(1) = \mathbf{g} .$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{L}^\alpha = i\tau^\alpha + [\boldsymbol{\xi}, \mathbf{L}^\alpha], \quad \mathbf{L}^\alpha(0) = 0 . \qquad (A38.14)$$

This equation can also be written in component form, setting

$$\mathbf{L}^\alpha = i L^{\alpha\beta} \tau^\beta .$$

Then,

$$\frac{\mathrm{d}}{\mathrm{d}t} L^{\alpha\beta} = \delta_{\alpha\beta} + f_{\gamma\beta\delta} \xi^\gamma L^{\alpha\delta} . \qquad (A38.15)$$

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 (A38.15) then can be written as

$$\Lambda(t) = \int_0^t \mathrm{d}t' \, \mathrm{e}^{X(t'-t)} = X^{-1}\left(1 - \mathrm{e}^{-tX}\right).$$

The metric tensor corresponding to the group is

$$g_{\alpha\beta} = -\operatorname{tr} \mathbf{L}^\alpha(1)\mathbf{L}^\beta(1) = L^{\alpha\gamma}(1) L^{\beta\gamma}(1) = (\Lambda\Lambda^T)_{\alpha\beta}, \qquad (A38.16)$$

and the group invariant measure is

$$\mathrm{d}\mathbf{g} \equiv (\det g_{\alpha\beta})^{1/2} \prod_\alpha \mathrm{d}\xi^\alpha = \left(\det \Lambda\Lambda^T\right)^{1/2} \prod_\alpha \mathrm{d}\xi^\alpha . \qquad (A38.17)$$

Then,

$$\begin{aligned} \Lambda\Lambda^T &= -X^{-2}\left(1 - \mathrm{e}^{-X}\right)\left(1 - \mathrm{e}^{X}\right) = -4X^{-2}\sinh^2(X/2) \\ &= -\prod_{n\neq 0}\left(1 + X^2/(2\pi n)^2\right), \end{aligned}$$

where we recognize an expression which appears in equation (38.129).

39 INSTANTONS IN QUANTUM MECHANICS

Up to now we have calculated euclidean functional integrals perturbatively, 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 have always a larger action than minimal constant solutions because the gradient term gives an additional positive contribution.

In what follows we will mainly be interested in the structure of the ground state and thus in the infinite volume limit. For a given constant solution we will consider 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. In this chapter we consider the example of simple Quantum Mechanics where instanton calculus is an alternative to the WKB method, but in the coming chapters we show that the instanton method can easily be generalized to Field Theory.

Dyson had conjectured that the behaviour of perturbation theory at large orders in QED was related to the vacuum instability for negative values of e^2 (e being the electric charge). Bender and Wu proved the analogous conjecture for the quartic anharmonic oscillator in quantum mechanics, showing that the behaviour of perturbation theory at large orders can indeed be obtained from a semi-classical calculation of barrier penetration effects occurring when the sign of the coupling constant is changed. We derive in Chapter 42 these results from instanton considerations, by a method that can be generalized to a large class of potentials (provided one considers also complex finite action solutions). Following Lipatov, we then extend this analysis to field theory, and characterize in Chapter 42 the large order behaviour of the perturbative expansion of several field theories. We finally show how this knowledge can be used to sum efficiently perturbation series. Remarkably enough one important application of these summation methods is a precise estimation of critical exponents. The results have been given in Chapter 29.

In this chapter we explain the role of instantons in unstable systems in quantum mechanics. We first discuss the quartic anharmonic oscillator with negative coupling and calculate the contributions of instantons at leading order. 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.

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 describe how analogous results can be obtained from calculations based on the WKB method, and comment about the average action in path integrals.

Let us finally mention that, although we shall deal only with euclidean theories, many aspects of the techniques we shall develop apply also to the calculation of effects coming from finite energy solutions of the real time field equations, called *soliton* solutions in the literature.

39 INSTANTONS IN QUANTUM MECHANICS

Up to now we have calculated euclidean functional integrals perturbatively, 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 have always a larger action than minimal constant solutions because the gradient term gives an additional positive contribution.

In what follows we will mainly be interested in the structure of the ground state and thus in the infinite volume limit. For a given constant solution we will consider 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. In this chapter we consider the example of simple Quantum Mechanics where instanton calculus is an alternative to the WKB method, but in the coming chapters we show that the instanton method can easily be generalized to Field Theory.

Dyson had conjectured that the behaviour of perturbation theory at large orders in QED was related to the vacuum instability for negative values of e^2 (e being the electric charge). Bender and Wu proved the analogous conjecture for the quartic anharmonic oscillator in quantum mechanics, showing that the behaviour of perturbation theory at large orders can indeed be obtained from a semi-classical calculation of barrier penetration effects occurring when the sign of the coupling constant is changed. We derive in Chapter 42 these results from instanton considerations, by a method that can be generalized to a large class of potentials (provided one considers also complex finite action solutions). Following Lipatov, we then extend this analysis to field theory, and characterize in Chapter 42 the large order behaviour of the perturbative expansion of several field theories. We finally show how this knowledge can be used to sum efficiently perturbation series. Remarkably enough one important application of these summation methods is a precise estimation of critical exponents. The results have been given in Chapter 29.

In this chapter we explain the role of instantons in unstable systems in quantum mechanics. We first discuss the quartic anharmonic oscillator with negative coupling and calculate the contributions of instantons at leading order. 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.

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 describe how analogous results can be obtained from calculations based on the WKB method, and comment about the average action in path integrals.

Let us finally mention that, although we shall deal only with euclidean theories, many aspects of the techniques we shall develop apply also to the calculation of effects coming from finite energy solutions of the real time field equations, called *soliton* solutions in the literature.

39.1 The Quartic Anharmonic Oscillator for Negative Coupling

We consider the hamiltonian H of the quartic anharmonic oscillator:

$$H = -\tfrac{1}{2}\left(\mathrm{d}/\mathrm{d}q\right)^2 + \tfrac{1}{2}q^2 + \tfrac{1}{4}gq^4. \tag{39.1}$$

The ground state energy $E_0(g)$ can be obtained from the large β limit of the partition function $\operatorname{tr}\mathrm{e}^{-\beta H}$:

$$E_0(g) = \lim_{\beta\to+\infty} -\frac{1}{\beta}\ln\operatorname{tr}\mathrm{e}^{-\beta H}.$$

Moreover, a systematic expansion of the partition function for β large also yields the energies of the excited states. We can thus derive the eigenvalues of H from the path integral representation of the partition function:

$$\operatorname{tr}\mathrm{e}^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q(t)]\exp\left[-\mathcal{S}\left(q(t)\right)\right], \tag{39.2}$$

where $\mathcal{S}(q)$ is the euclidean action:

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2}\left[\tfrac{1}{2}\dot{q}^2(t) + \tfrac{1}{2}q^2(t) + \tfrac{1}{4}gq^4(t)\right]\mathrm{d}t. \tag{39.3}$$

Generalization of arguments applicable to finite dimensional integrals indicates that the path integral (39.2) defines a function of g analytic in the half plane $\mathrm{Re}(g) > 0$. In this domain the integral is dominated for g small 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. This leads to the perturbative expansion of the partition function and thus of the ground state $E_0(g)$ by taking the large β limit.

Remarks.

(i) Note that we always expand in g before taking the large β limit. Since $E_N(g)$, the Nth eigenvalue of H, satisfies

$$E_N(g) = N + \tfrac{1}{2} + O(g),$$

the perturbative expansion can be written as

$$\operatorname{tr}\mathrm{e}^{-\beta H} = \sum_{N=0}\mathrm{e}^{-\beta E_N(g)} = \sum_{N=0}\mathrm{e}^{-(N+1/2)\beta}\sum_{k=0}\frac{1}{k!}(-\beta)^k\left(E_N - \tfrac{1}{2} - N\right)^k. \tag{39.4}$$

We observe that $E_N(g)$ can be deduced from the coefficient of $\mathrm{e}^{-(N+1/2)\beta}$, 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 rescaling $q(t)$:

$$q(t) \mapsto q(t)g^{-1/2},$$

we factorize the whole dependence in g in front of the action:

$$\mathcal{S}(q,g) = \frac{1}{g}\mathcal{S}(q\sqrt{g}). \tag{39.5}$$

The coupling constant g plays the same formal role as $\hbar$ in the semi-classical expansion.

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 $g = 0$ and the perturbation series is always divergent.

A wave function $\psi(t)$, localized at initial time $t = 0$ (t is here the *real physical time* of the Schrödinger equation) in the well of the potential near $q = 0$, decays due to barrier penetration. To calculate the decay rate of the wave function $\psi(t)$ we can use the following method: For g positive the time-dependent solution $\psi_0(t)$ of the Schrödinger equation associated with the ground state energy E_0 behaves like

$$\psi_0(t) \sim \mathrm{e}^{-iE_0 t}\,.$$

We then proceed by analytic continuation in the complex g plane to go to $g < 0$. After analytic continuation E_0 becomes complex and thus $\psi_0(t)$ decreases exponentially with time at a rate

$$|\psi_0(t)| \sim \mathrm{e}^{-|\operatorname{Im} E_0|t}\,.$$

$|\operatorname{Im} E_0|$ is the inverse lifetime of the wave function $\psi(t)$. Actually the decay of $\psi(t)$ 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. We shall, therefore, hereafter calculate $\operatorname{Im} E_0$ for g small and negative.

To now understand how we can define and evaluate $E_0(g)$ for g negative, we first study this problem on the example of a simple integral: the "zero-dimensional ϕ^4 field theory".

39.2 A Toy Model: A Simple Integral

We consider the integral which counts the number of Feynman diagrams contributing to the partition function in the ϕ^4 field theory:

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-(x^2/2 + gx^4/4)}\, \mathrm{d}x\,. \tag{39.6}$$

For g positive and small, the integral is dominated by the saddle point at the origin:

$$I(g) = 1 + O(g)\,. \tag{39.7}$$

The integral defines the function for $\operatorname{Re} g \geq 0$, but the function $I(g)$ is analytic in a cut plane. Its analytic continuation to $\operatorname{Re} g < 0$, can be obtained by rotating the contour of integration C in the complex plane as one changes the argument of g:

$$C:\ \arg x = -\tfrac{1}{4} \arg g \qquad (\mathrm{mod}\ \pi)\,,$$

so that $\operatorname{Re}(gx^4)$ remains positive. As a consequence, one obtains two different expressions for $I(g)$ depending on the direction in which one has rotated in the g-plane:

$$\text{for } g = -|g| + i0: \qquad I(g) = \int_{C_+} \mathrm{e}^{-(x^2/2 + gx^4/4)}\, \mathrm{d}x$$

$$\text{with } C_+: \qquad \arg x = -\frac{\pi}{4} \quad (\mathrm{mod}\ \pi)\,, \tag{39.8}$$

$$\text{for } g = -|g| - i0: \qquad I(g) = \int_{C_-} \mathrm{e}^{-(x^2/2 + gx^4/4)}\, \mathrm{d}x$$

$$\text{with } C_-: \qquad \arg x = \frac{\pi}{4} \quad (\mathrm{mod}\ \pi)\,. \tag{39.9}$$

For $g \to 0_-$, the two integrals are still dominated by the saddle point at the origin since the contribution of the other saddle points:

$$x + gx^3 = 0 \Rightarrow x^2 = -1/g\,, \tag{39.10}$$

are of the order

$$\mathrm{e}^{-(x^2/2+gx^4/4)} \sim \mathrm{e}^{1/4g} \ll 1\,. \tag{39.11}$$

However, the discontinuity of $I(g)$ on the cut is given by the difference between the two integrals:

$$I(g+i0) - I(g-i0) = 2i\,\mathrm{Im}\,I(g) = \frac{1}{\sqrt{\pi}} \int_{C_+ - C_-} \mathrm{e}^{-(x^2/2+gx^4/4)}\,\mathrm{d}x\,. \tag{39.12}$$

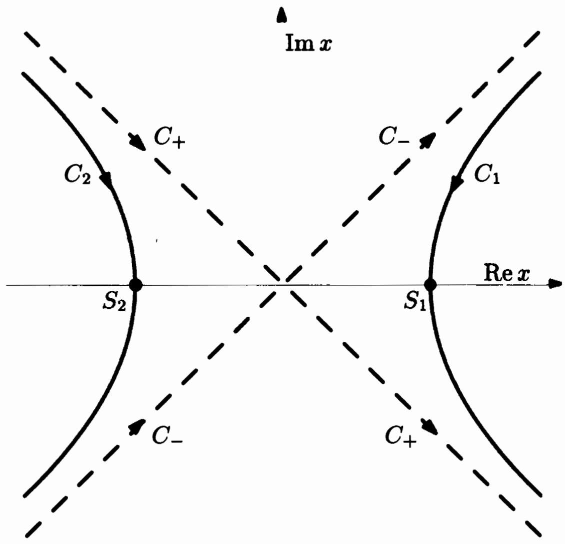


Fig. 39.1 The contours of integration C_+, C_-, C_1 and C_2.

It corresponds to the contour $C_+ - C_-$ which, as figure 39.1 shows, can be deformed into the sum of the contours C_1 and C_2 passing through the non-trivial saddle points S_1 and S_2: $x = \pm 1/\sqrt{-g}$. This means that the contribution of the saddle point at the origin cancels, and that the integral now is dominated by these saddle points:

$$\mathrm{Im}\,I(g) \sim \mathrm{e}^{1/4g}\,. \tag{39.13}$$

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.

39.3 Quantum Mechanics: Instantons

We now generalize the strategy followed in the case of the integral (39.6) to the path integral (39.2). We thus rotate the contour in the functional $q(t)$ space, as we change the argument g from g positive to g negative:

$$q(t) \mapsto q(t)\,\mathrm{e}^{-i\theta},$$

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), we verify that this is a reasonable procedure.

There is however one difference with the case of the simple integral: we have to stay in a domain in which $\mathrm{Re}\left[\dot{q}^2(t)\right] > 0$, since the kinetic part $\int \dot{q}^2(t)\mathrm{d}t$, as we have discussed in Chapter 2, favours the smooth paths and ensures, therefore, the existence of the continuum limit of the discretized path integral.

For g negative we thus integrate along the path:

$$\arg q(t) = -\theta, \quad \frac{1}{8}\pi < \theta < \frac{1}{4}\pi\,, \tag{39.14}$$

which satisfies the two conditions:

$$\mathrm{Re}\left[gq^4(t)\right] > 0\,, \qquad \mathrm{Re}\left[\dot{q}^2(t)\right] > 0\,. \tag{39.15}$$

For g small, the two path integrals corresponding to the two analytic continuations are here also 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:

$$-\ddot{q}(t) + q(t) + gq^3(t) = 0\,, \quad (g < 0)\,, \tag{39.16}$$

$$q(-\beta/2) = q(\beta/2). \tag{39.17}$$

The contribution of the constant saddle point:

$$q^2(t) = -1/g\,, \tag{39.18}$$

is of the order of $\mathrm{e}^{\beta/4g}$ and, therefore, negligible in the large β limit. We have to look for solutions which have an action which remains finite for $\beta \to +\infty$. These are called *instantons*.

The solutions of equations (39.16,39.17) correspond to a periodic motion in *real time* in the potential $-V(q)$

$$V(q) = \tfrac{1}{2}q^2 + \tfrac{1}{4}gq^4. \tag{39.19}$$

Solutions exist which correspond to oscillations around each of the minima of $-V$, $q = \pm\sqrt{-1/g}$. Integrating once equation (39.16), we find

$$\tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}q^2 - \tfrac{1}{4}gq^4 = \epsilon\,,$$

with $\epsilon < 0$. Calling q_0 and q_1 the points with $q > 0$ where the velocity $\dot{q}$ vanishes, we find for the period of such a solution:

$$\beta = 2 \int_{q_0}^{q_1} \frac{\mathrm{d}q}{\sqrt{q^2 + \frac{1}{2} g q^4 + 2\epsilon}}.$$

β can become large only if the constant ϵ and thus q_0 go to zero. With increasing β the classical trajectory comes closer to the origin. In the infinite β limit the classical solution becomes

$$q_c(t) = \pm \left(-\frac{2}{g}\right)^{1/2} \frac{1}{\cosh(t - t_0)}. \qquad (39.20)$$

The corresponding classical action is

$$\mathcal{S}(q_c) = -\frac{4}{3g} + O\left(\mathrm{e}^{-\beta}/g\right). \qquad (39.21)$$

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,

$$0 \le t_0 < \beta .$$

Therefore, in contrast to the simple integral, we do not find two degenerate saddle points but two one-parameter families.

Notice also that we could have 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 large β limit becomes

$$\mathcal{S}(q_c) = -n \frac{4}{3g}, \qquad (39.22)$$

and yields, therefore, a contribution proportional to $\mathrm{e}^{n4/3g}$. 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 $-n8/3g$.

39.4 Instanton Contribution at Leading Order

The gaussian approximation. To evaluate the contribution, at leading order, of the saddle points we expand the action around a saddle point, setting

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

and calculate for $g \to 0_-$, $\beta \to \infty$ the gaussian integration (one-loop order):

$$\begin{aligned} \operatorname{Im} \operatorname{tr} \mathrm{e}^{-\beta H} &= \frac{1}{i} \mathrm{e}^{4/3g} \int [\mathrm{d}r(t)] \exp\left[-\tfrac{1}{2} \int \mathrm{d}t \left(\dot{r}^2(t) + r^2(t) + 3g q_c^2(t) r^2(t)\right)\right], \\ &= \frac{1}{i} \mathrm{e}^{4/3g} \int [\mathrm{d}r(t)] \exp\left(-\tfrac{1}{2} \int \mathrm{d}t_1 \mathrm{d}t_2\, r(t_1) M(t_1, t_2) r\,(t_2)\right), \end{aligned}$$

where the operator M is given by

$$M(t_1, t_2) = \left.\frac{\delta^2 \mathcal{S}}{\delta q(t_1) \delta q(t_2)}\right|_{q = q_c} = \left[-\left(\mathrm{d}_{t_1}\right)^2 + 1 + 3g q_c^2(t_1)\right] \delta(t_1 - t_2). \qquad (39.23)$$

The path integral is normalized by comparing with the partition function of the harmonic oscillator.

The zero mode. Differentiating equation (39.16) with respect to t we obtain

$$-(\mathrm{d}_t)^2\dot{q}_c(t) + \dot{q}_c(t) + 3gq_c^2(t)\dot{q}_c(t) = 0\,. \tag{39.24}$$

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 zero:

$$M\dot{q}_c = 0\,. \tag{39.25}$$

Hence the naive gaussian approximation yields a result, proportional to $(\det M)^{-1/2}$, which is infinite!

We should not be too surprised by this result: as we have noted above, due to translation invariance in time, we have two one-parameter families of continuously connected degenerate saddle points. An infinitesimal variation of $q(t)$ which corresponds to a variation of the parameter t_0, that is, proportional to $\dot{q}_c$, leaves the action unchanged. The problem we face here is by no means special to path integrals as the following example shows.

Zero modes in simple integrals. We again consider a simple integral

$$I_2(g) = \int \mathrm{d}^\nu \mathbf{x}\; \mathrm{e}^{\mathbf{x}^2 - g(\mathbf{x}^2)^2}, \quad g > 0\,, \tag{39.26}$$

in which $\mathbf{x}$ is a ν-component vector, and the integrand is $O(\nu)$ invariant. For g small, this integral can be calculated by steepest descent. The saddle point is given by

$$\mathbf{x}_c\left(1 - 2g\mathbf{x}_c^2\right) = 0\,, \tag{39.27}$$

that is

$$|\mathbf{x}_c| = (2g)^{-1/2}. \tag{39.28}$$

We also find here a $\nu - 1$ parameter family of degenerate saddle points, since only the length of the vector $\mathbf{x}_c$ is determined by the saddle point equation. 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 $M_{\alpha\beta}$:

$$M_{\alpha\beta} = 8gx_\alpha x_\beta\,, \tag{39.29}$$

which is the projector on $\mathbf{x}$ and has, therefore, $(\nu - 1)$ zero eigenvalues.

The solution to this problem is obvious here: it is necessary to calculate the integral over angular variables exactly; only the integral over the radial variable can be evaluated by the steepest descent method.

In the case of the path integral it is also necessary to integrate exactly over the parameters which describe the saddle points, in the example above the time translation parameter. A suitable set of integration variables has to be found in which the time parameter explicitly appears. This is the method of so-called *collective coordinates.*

Remark. We have already studied theories possessing a symmetry group in which the classical minimum is not invariant under the group, for example the $(\phi^2)^2$ field theory in the ordered phase:

$$S(\phi) = \frac{1}{2}\int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi(x)\right)^2 + \frac{r}{2}\phi^2(x) + \frac{1}{4!}g\left(\phi^2(x)\right)^2\right]. \tag{39.30}$$

In this case we have generally chosen one classical minimum and made a systematic expansion around it. However, this is a procedure that is justified only when the symmetry is spontaneously broken. We have actually seen in Chapter 31 that the absence of symmetry breaking manifests itself in perturbation theory by the appearance of IR singularities. In the case of instanton solutions we have shown above that the propagator in an instanton background has an isolated pole at the origin. Therefore, all terms in perturbation theory are infinite. 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 23).

39.4.1 Collective coordinates and gaussian integration

In order to introduce an integration variable associated with time translation (a collective coordinate) a method inspired by the Faddeev–Popov quantization method of gauge theories can be used. We call t_0 the time collective coordinate and start from the identity:

$$1=\frac{1}{\sqrt{2\pi\xi}}\int \mathrm{d}t_0 \left[\int \mathrm{d}t\,\dot q_c(t)\dot q(t+t_0)\right]\exp\left\{-\frac{1}{2\xi}\left[\int \mathrm{d}t\,\dot q_c(t)\bigl(q(t+t_0)-q_c(t)\bigr)\right]^2\right\}, \tag{39.31}$$

which can be verified by changing variables $t_0\mapsto\lambda$

$$\lambda=\int \mathrm{d}t\,\dot q_c(t)\bigl(q(t+t_0)-q_c(t)\bigr).$$

The constant ξ has been introduced mainly for cosmetic reasons but is of order g.

We insert the identity (39.31) into the path integral. The new action

$$\mathcal{S}(q)+\frac{1}{2\xi}\left[\int \mathrm{d}t\,\dot q_c(t)\bigl(q(t+t_0)-q_c(t)\bigr)\right]^2$$

is no longer translation invariant. It leads to the saddle point equation

$$\frac{\delta\mathcal{S}}{\delta q(t)}+\frac{1}{\xi}\dot q_c(t-t_0)\int \mathrm{d}t'\,\dot q_c(t'-t_0)\bigl(q(t')-q_c(t'-t_0)\bigr)=0\,. \tag{39.32}$$

The equation is clearly satisfied for $q(t)=q_c(t-t_0)$. The determinant generated by the gaussian integration around the saddle point now is the determinant of a modified operator:

$$\det\frac{\delta^2\mathcal{S}}{\delta q_c(t_1)\delta q_c(t_2)}\mapsto\det\left\{\frac{\delta^2\mathcal{S}}{\delta q_c(t_1)\delta q_c(t_2)}+\frac{1}{\xi}\dot q_c(t_1-t_0)\dot q_c(t_2-t_0)\right\}.$$

The modified operator has all the same eigenvalues as the initial one but one: the eigenvalue that corresponds to the eigenvector $\dot q_c$ is now $\|\dot q_c\|^2/\xi$ 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 $\mathrm{e}^{-\beta/2}$. At $g=0$ the operator M becomes the operator M_0:

$$M_0(t_1,t_2)=\left[-\left(\mathrm{d}_{t_1}\right)^2+1\right]\delta(t_1-t_2). \tag{39.33}$$

As we indicate below, what 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 $\varepsilon \to 0$ it vanishes like ε and we thus set

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \det(M+\varepsilon)(M_0+\varepsilon)^{-1} \equiv \det{}' M M_0^{-1}. \tag{39.34}$$

Instead what we have to calculate has the form

$$\det(M+\mu\,|1\rangle\langle 1|)M_0^{-1} = \lim_{\varepsilon\to 0} \det(M+\varepsilon+\mu\,|1\rangle\langle 1|)(M_0+\varepsilon)^{-1},$$

where we have denoted by $|1\rangle$ the eigenvector proportional to $\dot{q}_c$ with unit norm, and μ is given by

$$\mu = \|\dot{q}_c\|^2/\xi\,.$$

Then,

$$\begin{aligned}\det(M+\varepsilon+\mu\,|1\rangle\langle 1|)(M_0+\varepsilon)^{-1} &= \det(M+\varepsilon)(M_0+\varepsilon)^{-1} \\ &\quad\times \det\bigl(1+\mu\,|1\rangle\langle 1|(M+\varepsilon)^{-1}\bigr) \\ &= \det(M+\varepsilon)(M_0+\varepsilon)^{-1}(1+\mu/\varepsilon).\end{aligned}$$

One thus finds

$$\det{}' M M_0^{-1}\, \|\dot{q}_c\|^2/\xi\,.$$

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 (39.31) at leading order we replace $q(t+t_0)$ by $q_c(t)$. The integral does not depend on t_0 anymore, we get a 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

$$\frac{\beta}{\sqrt{2\pi}}\mathcal{Z}_0(\beta)\|\dot{q}_c\|(\det{}' M M_0^{-1})^{-1/2}.$$

Taking into account the two families of saddle points, and dividing by $2i$ to get the imaginary part, we obtain for $\beta \to \infty$:

$$\operatorname{Im}\operatorname{tr} \mathrm{e}^{-\beta H} \sim \frac{2}{2i}\frac{\beta}{\sqrt{2\pi}}\,\mathrm{e}^{-\beta/2}\,\|\dot{q}_c\|\left[\det{}' M\,(\det M_0)^{-1}\right]^{-1/2}\mathrm{e}^{4/3g}\,. \tag{39.35}$$

39.4.2 The result at leading order

In Section 39.6 we calculate the determinant for a general potential in one-dimensional systems, and in Appendix A39.2 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 happens to be the hamiltonian corresponding to an exactly soluble Bargmann potential, and one finds

$$\det(M+\varepsilon)(M_0+\varepsilon)^{-1} = \frac{\sqrt{1+\varepsilon}-1}{\sqrt{1+\varepsilon}+1}\frac{\sqrt{1+\varepsilon}-2}{\sqrt{1+\varepsilon}+2}. \tag{39.36}$$

As we indicate below, what 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 $\varepsilon \to 0$ it vanishes like ε and we thus set

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\det(M+\varepsilon)(M_0+\varepsilon)^{-1} \equiv \det{}' M M_0^{-1}. \tag{39.34}$$

Instead what we have to calculate has the form

$$\det(M+\mu\,|1\rangle\langle 1|)M_0^{-1} = \lim_{\varepsilon\to 0}\det(M+\varepsilon+\mu\,|1\rangle\langle 1|)(M_0+\varepsilon)^{-1},$$

where we have denoted by $|1\rangle$ the eigenvector proportional to $\dot q_c$ with unit norm, and μ is given by

$$\mu = \|\dot q_c\|^2/\xi\,.$$

Then,

$$\begin{aligned}\det(M+\varepsilon+\mu\,|1\rangle\langle 1|)(M_0+\varepsilon)^{-1} &= \det(M+\varepsilon)(M_0+\varepsilon)^{-1}\\ &\quad\times\det\bigl(1+\mu\,|1\rangle\langle 1|(M+\varepsilon)^{-1}\bigr)\\ &= \det(M+\varepsilon)(M_0+\varepsilon)^{-1}(1+\mu/\varepsilon).\end{aligned}$$

One thus finds

$$\det{}' M M_0^{-1}\,\|\dot q_c\|^2/\xi\,.$$

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 (39.31) at leading order we replace $q(t+t_0)$ by $q_c(t)$. The integral does not depend on t_0 anymore, we get a 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

$$\frac{\beta}{\sqrt{2\pi}}\mathcal{Z}_0(\beta)\|\dot q_c\|(\det{}' M M_0^{-1})^{-1/2}.$$

Taking into account the two families of saddle points, and dividing by $2i$ to get the imaginary part, we obtain for $\beta\to\infty$:

$$\operatorname{Im}\operatorname{tr}\mathrm{e}^{-\beta H} \sim \frac{2}{2i}\frac{\beta}{\sqrt{2\pi}}\,\mathrm{e}^{-\beta/2}\,\|\dot q_c\|\left[\det{}' M\,(\det M_0)^{-1}\right]^{-1/2}\mathrm{e}^{4/3g}\,. \tag{39.35}$$

39.4.2 The result at leading order

In Section 39.6 we calculate the determinant for a general potential in one-dimensional systems, and in Appendix A39.2 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 happens to be the hamiltonian corresponding to an exactly soluble Bargmann potential, and one finds

$$\det(M+\varepsilon)(M_0+\varepsilon)^{-1} = \frac{\sqrt{1+\varepsilon}-1}{\sqrt{1+\varepsilon}+1}\,\frac{\sqrt{1+\varepsilon}-2}{\sqrt{1+\varepsilon}+2}. \tag{39.36}$$

Simple arguments show that the ground state wave function has no node, the wave function of the first excited state one node, etc. The wave function $\dot{q}_c(t)$ vanishes once at the turning point, therefore, it does not correspond to the ground state. There exists one other state which is the ground state, and M has, therefore, one negative eigenvalue, as is apparent on expression (39.36).

Then,

$$\det{}' M \, (\det M_0)^{-1} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \det (M + \varepsilon) \, (M_0 + \varepsilon)^{-1} = -\frac{1}{12} \, . \tag{39.37}$$

We still have an ambiguity of sign since we have to take the square root of expression (39.36). This ambiguity can only be resolved by following the analytic continuation from g positive to g negative. However, in both cases the square root of the determinant is imaginary, so that the final result is real, as expected.

The norm $\|\dot{q}_c\|$ is easily calculable, but the important point is that q_c is proportional to $1/\sqrt{g}$. As we shall see later, this is the first example of a general situation: each time the instanton solution breaks some 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

$$\|\dot{q}_c\| = \frac{2}{\sqrt{3}} \frac{1}{\sqrt{-g}} . \tag{39.38}$$

The expression then becomes

$$\operatorname{Im} \operatorname{tr} \mathrm{e}^{-\beta H} = -\beta \, \mathrm{e}^{-\beta/2} \frac{4}{\sqrt{2\pi}} \frac{1}{\sqrt{-g}} \mathrm{e}^{4/3g} \left[1 + O\left(g, \mathrm{e}^{-\beta}\right) \right] , \quad \text{for } g \to 0_- \, , \ \beta \to \infty \, . \tag{39.39}$$

For β large, the l.h.s. has the form:

$$\operatorname{Im} \operatorname{tr} \mathrm{e}^{-\beta H} \sim \operatorname{Im} \mathrm{e}^{-\beta E_0(g)} \equiv \operatorname{Im} \mathrm{e}^{-\beta (\operatorname{Re} E_0(g) + i \operatorname{Im} E_0(g))} , \quad \text{for } g \to 0_- \, , \ \beta \to \infty \, . \tag{39.40}$$

We have learned that 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

$$\operatorname{Im} \operatorname{tr} \mathrm{e}^{-\beta H} \sim -\beta \operatorname{Im} (E_0(g)) \, \mathrm{e}^{-\beta \operatorname{Re} E_0(g)} \sim -\beta \, \mathrm{e}^{-\beta/2} \operatorname{Im} E_0 \, . \tag{39.41}$$

Equation (39.39) then leads to

$$\operatorname{Im} E_0(g) = \frac{4}{\sqrt{2\pi}} \frac{\mathrm{e}^{4/3g}}{\sqrt{-g}} \left[1 + O(g) \right] , \quad g \to 0_- \, . \tag{39.42}$$

A systematic expansion around the saddle point then generates an expansion in powers of g.

Remark. We have obtained the behaviour, for g small and negative, of the imaginary part of the ground state energy and, therefore, the decay rate of a state localized in the unbounded potential corresponding to the anharmonic oscillator with negative coupling. In Chapter 42 we show that this result also leads to an evaluation of the large order behaviour of perturbation series for the anharmonic oscillator.

39.5 General Potentials: Instanton Contributions

We now calculate, still in one-dimensional systems but for a general class of analytic potentials, the decay rate of a wave packet located at initial time at a relative minimum of a potential and decaying through barrier penetration, generalizing the method of preceding sections.

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 some analytic continuation becomes a relative minimum of the potential. As we have discussed in Section 39.1, 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. We calculate the contribution of instantons at leading order.

The instanton solution. We consider hamiltonians of the form:

$$H = -\tfrac{1}{2}\left(\mathrm{d}/\mathrm{d}q\right)^2 + g^{-1}V(q\sqrt{g}). \tag{39.43}$$

The function $V(q)$ is an analytic function of q which, for q small, behaves like

$$V(q) = \tfrac{1}{2}q^2 + O\left(q^3\right). \tag{39.44}$$

Again in the hamiltonian (39.43) the potential has been parametrized in such a way that g plays the formal role of $\hbar$.

The path integral representation of the partition function $\operatorname{tr} \mathrm{e}^{-\beta H}$ is

$$\operatorname{tr} \mathrm{e}^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} \left[\mathrm{d}q(t)\right] \exp\left[-\mathcal{S}(q)\right], \tag{39.45}$$

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}^2(t) + g^{-1}V(q(t)\sqrt{g})\right] \mathrm{d}t\,. \tag{39.46}$$

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 regular and thus continuous, has at least another zero. For β infinite, one instanton solution $q_c(t)$ starts from the origin at time $-\infty$, is reflected on the zero of the potential, and comes back to the origin at time $+\infty$.

The euclidean equation of motion is

$$\ddot{q}_c(t) = \frac{1}{\sqrt{g}} V'\left(q_c(t)\sqrt{g}\right). \tag{39.47}$$

Integrating once, we obtain for a finite action solution (for $\beta = \infty$):

$$\tfrac{1}{2}\dot{q}_c^2(t) - g^{-1}V\left(q_c(t)\sqrt{g}\right) = 0\,. \tag{39.48}$$

Calling x_0 the relevant zero of $V(x)$, we can write the corresponding action:

$$\mathcal{S}(q_c) = \int_{-\infty}^{+\infty} \dot{q}_c^2(t)\mathrm{d}t = \frac{a}{g}, \qquad a = 2\int_0^{x_0} \sqrt{2V(x)}\mathrm{d}x\,. \tag{39.49}$$

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 around the saddle point $q_c(t)$ in the gaussian approximation. However, as we have shown in previous section, we must first separate a collective coordinate corresponding to time translation. This yields a factor β and a jacobian J:

$$J = \left[\int_{-\infty}^{+\infty} \dot{q}_c^2(t)\mathrm{d}t\right]^{1/2} = \left(\frac{a}{g}\right)^{1/2}. \tag{39.50}$$

Since the function $\dot{q}_c(t)$, which is an eigenfunction of $\delta^2\mathcal{S}/(\delta q_c)^2$ with eigenvalue zero, has a node at the turning point $x_0/\sqrt{g}$, there exists an eigenfunction associated with a negative eigenvalue. Consequently, the determinant which appears in the gaussian integration around the saddle point is negative.

Collecting all factors, we obtain

$$\begin{aligned} \operatorname{Im} \operatorname{tr} \mathrm{e}^{-\beta H} &\sim \frac{1}{2}\frac{\beta}{\sqrt{2\pi}}\,\mathrm{e}^{-\beta/2}\sqrt{\frac{a}{g}}\left[\det M_0 \left(-\det{}' M\right)^{-1}\right]^{1/2}\mathrm{e}^{-a/g} \\ &\text{for } g \to 0\,, \quad \beta \to \infty\,. \end{aligned} \tag{39.51}$$

This yields for the imaginary part of the "ground state" energy E_0:

$$\operatorname{Im} E_0 = \frac{1}{2}\sqrt{\frac{a}{2\pi g}}\left[\det M_0 \left(-\det{}' M\right)^{-1}\right]^{1/2}\mathrm{e}^{-a/g}. \tag{39.52}$$

The definitions of M_0, M, $\det'$ are analogous to those of Section 39.4:

$$M(t_1,t_2) = \delta^2\mathcal{S}/\delta q_c(t_1)\delta q_c(t_2), \qquad M_0(t_1,t_2) = \left[-\left(\mathrm{d}_{t_1}\right)^2 + 1\right]\delta(t_1 - t_2), \tag{39.53}$$

and $\det'$ means determinant in the subspace orthogonal to $\dot{q}_c$.

To complete the calculation we evaluate $\det' M$ explicitly, a calculation which can always be done in one-dimensional systems, and more generally in classically integrable systems. In Appendix A39.2, for a comparison, we recall how the calculation can be done with the WKB method.

39.6 Gaussian Integration: The Shifting Method

For reasons which will become apparent later, we first calculate the general matrix element:

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle = \int_{q(-\beta/2)=x'}^{q(\beta/2)=x} [\mathrm{d}q(t)] \exp\left[-\mathcal{S}(q)\right]. \tag{39.54}$$

We call $\mathcal{S}_c(x',x;\beta)$ the classical action corresponding to the classical solution with the prescribed boundary conditions:

$$\begin{cases} \mathcal{S}_c(x',x;\beta) = \displaystyle\int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}_c^2(t) + g^{-1}V\left(q_c(t)\sqrt{g}\right)\right]\mathrm{d}t\,, \\ q_c(-\beta/2) = x'\,, \quad q_c(\beta/2) = x\,. \end{cases} \tag{39.55}$$

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 around the saddle point $q_c(t)$ in the gaussian approximation. However, as we have shown in previous section, we must first separate a collective coordinate corresponding to time translation. This yields a factor β and a jacobian J:

$$J = \left[\int_{-\infty}^{+\infty} \dot{q}_c^2(t)\mathrm{d}t\right]^{1/2} = \left(\frac{a}{g}\right)^{1/2}. \tag{39.50}$$

Since the function $\dot{q}_c(t)$, which is an eigenfunction of $\delta^2\mathcal{S}/(\delta q_c)^2$ with eigenvalue zero, has a node at the turning point $x_0/\sqrt{g}$, there exists an eigenfunction associated with a negative eigenvalue. Consequently, the determinant which appears in the gaussian integration around the saddle point is negative.

Collecting all factors, we obtain

$$\begin{aligned}\operatorname{Im}\operatorname{tr}\mathrm{e}^{-\beta H} &\sim \frac{1}{2}\frac{\beta}{\sqrt{2\pi}}\mathrm{e}^{-\beta/2}\sqrt{\frac{a}{g}}\left[\det M_0\left(-\det{}'M\right)^{-1}\right]^{1/2}\mathrm{e}^{-a/g}\\ &\text{for } g\to 0\,,\quad \beta\to\infty\,.\end{aligned} \tag{39.51}$$

This yields for the imaginary part of the "ground state" energy E_0:

$$\operatorname{Im} E_0 = \frac{1}{2}\sqrt{\frac{a}{2\pi g}}\left[\det M_0\left(-\det{}'M\right)^{-1}\right]^{1/2}\mathrm{e}^{-a/g}. \tag{39.52}$$

The definitions of M_0, M, $\det'$ are analogous to those of Section 39.4:

$$M(t_1,t_2) = \delta^2\mathcal{S}\,/\delta q_c(t_1)\delta q_c(t_2), \qquad M_0(t_1,t_2) = \left[-\left(\mathrm{d}_{t_1}\right)^2 + 1\right]\delta(t_1-t_2), \tag{39.53}$$

and $\det'$ means determinant in the subspace orthogonal to $\dot{q}_c$.

To complete the calculation we evaluate $\det'M$ explicitly, a calculation which can always be done in one-dimensional systems, and more generally in classically integrable systems. In Appendix A39.2, for a comparison, we recall how the calculation can be done with the WKB method.

39.6 Gaussian Integration: The Shifting Method

For reasons which will become apparent later, we first calculate the general matrix element:

$$\left\langle x'\left|\mathrm{e}^{-\beta H}\right|x\right\rangle = \int_{q(-\beta/2)=x'}^{q(\beta/2)=x}[\mathrm{d}q(t)]\exp\left[-\mathcal{S}(q)\right]. \tag{39.54}$$

We call $\mathcal{S}_c(x',x;\beta)$ the classical action corresponding to the classical solution with the prescribed boundary conditions:

$$\begin{cases}\mathcal{S}_c(x',x;\beta) = \displaystyle\int_{-\beta/2}^{\beta/2}\left[\tfrac{1}{2}\dot{q}_c^2(t) + g^{-1}V\left(q_c(t)\sqrt{g}\right)\right]\mathrm{d}t\,,\\ q_c(-\beta/2) = x'\,,\quad q_c(\beta/2) = x\,.\end{cases} \tag{39.55}$$

Setting

$$q(t) = q_c(t) + r(t)\,, \tag{39.56}$$

we obtain, at leading order, the path integral:

$$\begin{aligned} \left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle &= \mathrm{e}^{-\mathcal{S}_c} \int_{r(-\beta/2)=0}^{r(\beta/2)=0} [\mathrm{d}r(t)] \exp\left[-\Sigma(r)\right], \\ \Sigma(r) &= \int_{-\beta/2}^{\beta/2} \tfrac{1}{2} \left[\dot r^2(t) + V''(q_c\sqrt{g}) r^2(t)\right] \mathrm{d}t\,. \end{aligned} \tag{39.57}$$

We now calculate the gaussian integral over $r(t)$ using the so-called *shifting method.* The main drawback of this method is that it involves a dangerous change of variables and the final result is at first sight ill-defined. On the other hand it allows a rather straightforward evaluation of the determinant. 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 which maps any hamiltonian system onto a standard one (here we choose a free hamiltonian). For details see Appendix A39.1.

39.6.1 The shifting method

We first set:

$$V''\left(\sqrt{g} q_c(t)\right) = \ddot\kappa(t)/\kappa(t). \tag{39.58}$$

We know at least one possible choice for $\kappa(t)$. If we differentiate the equation of motion (39.47), we find

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 \dot q_c(t) = V''(q_c\sqrt{g})\dot q_c(t)\,. \tag{39.59}$$

If $\dot q_c(t)$ does not vanish on the classical trajectory, we choose

$$\kappa(t) = \dot q_c(t)\,. \tag{39.60}$$

Otherwise, we look for a linear combination of the two independent solutions of equation (39.58), $\dot q_c(t)$ and

$$\dot q_c(t) \int^t \frac{\mathrm{d}\tau}{\left[\dot q_c(\tau)\right]^2},$$

which does not vanish on the classical trajectory.

The action in expression (39.57) then becomes

$$\int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot r^2(t) + \frac{\ddot\kappa(t)}{\kappa(t)} r^2(t)\right] \mathrm{d}t = \int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot r^2(t) - \dot\kappa(t) \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{r^2(t)}{\kappa(t)}\right)\right] \mathrm{d}t\,. \tag{39.61}$$

In the integration by parts, the integrated terms vanish due to the boundary conditions. The evaluation of the r.h.s. of equation (39.61) leads to the remarkable identity:

$$\int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot r^2(t) + \frac{\ddot\kappa(t)}{\kappa(t)} r^2(t)\right] \mathrm{d}t = \int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot r(t) - \frac{\dot\kappa}{\kappa} r(t)\right]^2 \mathrm{d}t\,. \tag{39.62}$$

This suggests an obvious linear change of variable, $r(t) \mapsto \sigma(t)$:

$$\dot r(t) - \frac{\dot\kappa(t)}{\kappa(t)} r(t) = \dot\sigma(t)\,, \qquad \sigma(-\beta/2) = 0\,, \tag{39.63}$$

which transforms the hamiltonian of the time-dependent harmonic oscillator, into the free hamiltonian:

$$\int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot{r}^2(t) + \frac{\ddot{\kappa}(t)}{\kappa(t)} r^2(t)\right] \mathrm{d}t = \int_{-\beta/2}^{\beta/2} \frac{1}{2}\left[\dot{\sigma}^2(t)\right] \mathrm{d}t . \tag{39.64}$$

The change of variables (39.63) reminds us of the Langevin equation we discussed in Chapter 4, $r(t)$ playing the role of the dynamical variable and $\dot{\sigma}$ the role of the noise. We are, therefore, not too surprised to encounter the same difficulty as in naive continuum derivations of the Fokker–Planck equation. Integrating equation (39.63), we obtain

$$r(t) = \kappa(t) \int_{-\beta/2}^{t} \mathrm{d}\tau \, \frac{\dot{\sigma}(\tau)}{\kappa(\tau)} = \sigma(t) + \kappa(t) \int_{-\beta/2}^{t} \mathrm{d}\tau \, \sigma(\tau) \frac{\dot{\kappa}(\tau)}{\kappa^2(\tau)} . \tag{39.65}$$

The jacobian J of this transformation is formally the determinant of the kernel (see Section 4.6)

$$J = \det \frac{\delta r(t_2}{\delta \sigma(t_1)} = \det \left[\delta(t_1 - t_2) + \theta(t_2 - t_1)\kappa(t_2) \frac{\dot{\kappa}(t_1)}{\kappa^2(t_1)}\right] , \tag{39.66}$$

where $\theta(t)$ is the step function. Using,

$$\ln \det(1 + M) = \operatorname{tr} \ln(1 + M) = \operatorname{tr} M - \tfrac{1}{2} \operatorname{tr} M^2 + \cdots , \tag{39.67}$$

we get,

$$\ln J = \theta(0) \int_{-\beta/2}^{\beta/2} \mathrm{d}t \frac{\dot{\kappa}(t)}{\kappa(t)} . \tag{39.68}$$

For reasons we have discussed in Section 3.2.1 (commutation of derivative and expectation value that is required to justify the identity (39.62) within the path integral), the proper prescription is

$$\theta(0) = \tfrac{1}{2} . \tag{39.69}$$

Integrating over t we then obtain the jacobian

$$J = \sqrt{\frac{\kappa(\beta/2)}{\kappa(-\beta/2)}} . \tag{39.70}$$

In the path integral we still have to impose the boundary condition:

$$0 = (\beta/2) = \kappa(\beta/2) \int_{-\beta/2}^{\beta/2} \mathrm{d}t \, \frac{\dot{\sigma}(t)}{\kappa(t)} . \tag{39.71}$$

This can be achieved by introducing a δ-function for which, as usual, we use a Fourier representation

$$\delta\bigl(r(\beta/2)\bigr) = \frac{1}{\kappa(\beta/2)} \int \frac{\mathrm{d}\lambda}{2\pi} \exp\left(i\lambda \int_{-\beta/2}^{\beta/2} \mathrm{d}t \, \frac{\dot{\sigma}(t)}{\kappa(t)}\right) .$$

The complete expression then reads:

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \mathrm{e}^{-\mathcal{S}_c} \int_{\sigma(-\beta/2)=0} [\mathrm{d}\sigma(t)] \frac{\mathrm{d}\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} \mathrm{e}^{-\mathcal{S}(\sigma,\lambda)}, \tag{39.72}$$

with

$$\mathcal{S}(\sigma,\lambda) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \left(\frac{1}{2}\dot{\sigma}^2(t) - i\lambda \frac{\dot{\sigma}(t)}{\kappa(t)}\right). \tag{39.73}$$

To eliminate the term linear in $\dot{\sigma}(t)$ in equation (39.73), we shift $\dot{\sigma}(t)$,

$$\dot{\sigma}(t) = i\frac{\lambda}{\kappa(t)} + \dot{\sigma}'(t). \tag{39.74}$$

After this shift the path integral becomes

$$\begin{aligned}\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \mathrm{e}^{-\mathcal{S}_c} &\int_{\sigma(-\beta/2)=0} [\mathrm{d}\sigma(t)] \frac{\mathrm{d}\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} \\ &\times \exp\left[-\frac{1}{2}\lambda^2 \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} - \frac{1}{2}\int_{-\beta/2}^{\beta/2} \dot{\sigma}^2(t)\mathrm{d}t\right].\end{aligned} \tag{39.75}$$

We integrate over λ:

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \mathrm{e}^{-\mathcal{S}_c(x',x;\beta)} \left[\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)}\right]^{-1/2} \mathcal{N}(\beta). \tag{39.76}$$

The constant $\mathcal{N}(\beta)$ does not depend on x and x', and is proportional to the matrix element $\langle 0 | \mathrm{e}^{-\beta H_0} | 0 \rangle$ in which H_0 is the free hamiltonian:

$$\left\langle x' \left| \mathrm{e}^{-\beta H_0} \right| x \right\rangle = (2\pi\beta)^{-1/2} \mathrm{e}^{-(x'-x)^2/2\beta}. \tag{39.77}$$

To determine $\mathcal{N}(\beta)$ we set $H = H_0$ in equation (39.76) and note that in this case $\kappa(t)$ is a constant.

The final result is

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \mathrm{e}^{-\mathcal{S}_c(x',x;\beta)} \left[2\pi\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)}\right]^{-1/2}. \tag{39.78}$$

We leave as an exercise to show that the result (39.78) is formally independent of the particular linear combination of the two solutions of equation (39.58) one has chosen. To obtain a more explicit expression we then substitute for example $\kappa(t) = \dot{q}_c(t)$. We integrate the equation of motion (39.47), taking into account the boundary conditions:

$$\tfrac{1}{2}\dot{q}_c^2 = g^{-1}\left[V\left(q_c(t)\sqrt{g}\right) + E\right], \tag{39.79}$$

and, therefore,

$$\beta = \int_{x'\sqrt{g}}^{x\sqrt{g}} \frac{\mathrm{d}q}{\left[2\left(E + V(q)\right)\right]^{1/2}}. \tag{39.80}$$

The complete expression then reads:

$$\left\langle x' \left| e^{-\beta H} \right| x \right\rangle \sim e^{-\mathcal{S}_c} \int_{\sigma(-\beta/2)=0} [d\sigma(t)] \frac{d\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} e^{-\mathcal{S}(\sigma,\lambda)}, \tag{39.72}$$

with

$$\mathcal{S}(\sigma,\lambda) = \int_{-\beta/2}^{\beta/2} dt \left(\frac{1}{2}\dot{\sigma}^2(t) - i\lambda \frac{\dot{\sigma}(t)}{\kappa(t)} \right). \tag{39.73}$$

To eliminate the term linear in $\dot{\sigma}(t)$ in equation (39.73), we shift $\dot{\sigma}(t)$,

$$\dot{\sigma}(t) = i\frac{\lambda}{\kappa(t)} + \dot{\sigma}'(t). \tag{39.74}$$

After this shift the path integral becomes

$$\begin{aligned} \left\langle x' \left| e^{-\beta H} \right| x \right\rangle \sim e^{-\mathcal{S}_c} & \int_{\sigma(-\beta/2)=0} [d\sigma(t)] \frac{d\lambda}{2\pi} \frac{1}{\sqrt{\kappa(\beta/2)\kappa(-\beta/2)}} \\ & \times \exp\left[-\frac{1}{2}\lambda^2 \int_{-\beta/2}^{\beta/2} \frac{dt}{\kappa^2(t)} - \frac{1}{2}\int_{-\beta/2}^{\beta/2} \dot{\sigma}^2(t) dt \right]. \end{aligned} \tag{39.75}$$

We integrate over λ:

$$\left\langle x' \left| e^{-\beta H} \right| x \right\rangle \sim e^{-\mathcal{S}_c(x',x;\beta)} \left[\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{dt}{\kappa^2(t)} \right]^{-1/2} \mathcal{N}(\beta). \tag{39.76}$$

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:

$$\left\langle x' \left| e^{-\beta H_0} \right| x \right\rangle = (2\pi\beta)^{-1/2} e^{-(x'-x)^2/2\beta}. \tag{39.77}$$

To determine $\mathcal{N}(\beta)$ we set $H = H_0$ in equation (39.76) and note that in this case $\kappa(t)$ is a constant.

The final result is

$$\left\langle x' \left| e^{-\beta H} \right| x \right\rangle \sim e^{-\mathcal{S}_c(x',x;\beta)} \left[2\pi\kappa(\beta/2)\kappa(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{dt}{\kappa^2(t)} \right]^{-1/2}. \tag{39.78}$$

We leave as an exercise to show that the result (39.78) is formally independent of the particular linear combination of the two solutions of equation (39.58) one has chosen. To obtain a more explicit expression we then substitute for example $\kappa(t) = \dot{q}_c(t)$. We integrate the equation of motion (39.47), taking into account the boundary conditions:

$$\tfrac{1}{2}\dot{q}_c^2 = g^{-1}\left[V\left(q_c(t)\sqrt{g}\right) + E \right], \tag{39.79}$$

and, therefore,

$$\beta = \int_{x'\sqrt{g}}^{x\sqrt{g}} \frac{dq}{\left[2(E+V(q))\right]^{1/2}}. \tag{39.80}$$

Differentiating equation (39.80) with respect to β, we obtain

$$1 = -\int_{x'\sqrt{g}}^{x\sqrt{g}} \frac{\mathrm{d}q}{\left[2\left(E + V(q)\right)\right]^{3/2}} \frac{\partial E}{\partial \beta}, \tag{39.81}$$

which can be written as

$$\frac{\partial E}{\partial \beta} = -\left[\int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)}\right]^{-1}. \tag{39.82}$$

The result (39.78) can then also be written as

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \mathrm{e}^{-\mathcal{S}_c(x',x;\beta)} \frac{1}{\sqrt{2\pi \dot{q}_c(\beta/2)\dot{q}_c(-\beta/2)}} \left(-\frac{\partial E}{\partial \beta}\right)^{1/2}. \tag{39.83}$$

We leave as an exercise to verify the identity

$$\kappa(\beta/2)\kappa\,(-\beta/2) \int_{-\beta/2}^{\beta/2} \frac{\mathrm{d}t}{\kappa^2(t)} = \left(-\frac{\partial^2 \mathcal{S}_c}{\partial x \partial x'}\right)^{-1}. \tag{39.84}$$

Substituting equation (39.84) into equation (39.78) one then obtains Van Vleck's formula:

$$\left\langle x' \left| \mathrm{e}^{-\beta H} \right| x \right\rangle \sim \left(-\frac{1}{2\pi} \frac{\partial^2 \mathcal{S}_c}{\partial x \partial x'}\right)^{1/2} \exp\left[-\mathcal{S}_c(x',x;\beta)\right]. \tag{39.85}$$

Several degrees of freedom. The calculation of the instanton contribution by the shifting method can be generalized to $d > 1$ degrees of freedom provided one can find a non-singular matrix K solution of the equation:

$$\ddot{K}_{ij} = \frac{\partial V\left(\mathbf{q}_c(t)\right)}{\partial q_i \partial q_k} K_{kj}\,. \tag{39.86}$$

The change of variables (39.63) then takes the form:

$$\dot{\mathbf{r}} - \dot{K}K^{-1}\mathbf{r} = \dot{\boldsymbol{\sigma}}\,. \tag{39.87}$$

The matrix K can be chosen in such a way that $\dot{K}K^{-1}$ is symmetric. It is then easy to verify that all arguments can be repeated and one finally obtains an expression similar to (39.83):

$$\begin{aligned}\left\langle \mathbf{x}' \left| \mathrm{e}^{-\beta H} \right| \mathbf{x} \right\rangle \sim{}& \left\{(2\pi)^d \det\left[K(\beta/2)K(-\beta/2) \int_{-\beta/2}^{\beta/2} \mathrm{d}t\ \left({}^T\!K\right)^{-1} K^{-1}\right]\right\}^{-1/2} \\ &\times \exp\left[-\mathcal{S}_c(\mathbf{x}',\mathbf{x};\beta)\right].\end{aligned} \tag{39.88}$$

This expression is again equivalent to Van Vleck's formula (see Appendix A39.2) 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 corresponds instead to the special class of integrable hamiltonians. A simple example is provided by $O(N)$ symmetric systems.

39.6.2 The partition function

In order to calculate $\operatorname{tr} e^{-\beta H}$ we now impose periodic boundary conditions. Then,

$$[\dot{q}_c(\beta/2)\dot{q}_c(-\beta/2)]^{-1/2} = \left\{\frac{2}{g}\left[V(x\sqrt{g}) + E\right]\right\}^{-1/2}. \tag{39.89}$$

Integrating over x we obtain the trace. Using equation (39.80), we find

$$\int \mathrm{d}x \left[\frac{2}{g}\left(V(x\sqrt{g}) + E\right)\right]^{-1/2} = \beta\,. \tag{39.90}$$

We now collect all factors and obtain a more explicit expression:

$$\operatorname{Im} \operatorname{tr} e^{-\beta H} \sim \frac{\beta}{2i}\left(-\frac{\partial E}{\partial \beta}\frac{1}{2\pi g}\right)^{1/2} e^{-A(\beta)/g}, \tag{39.91}$$

where $E(\beta)$ and $A(\beta)$ are defined by

$$\beta = 2\int_{x_-}^{x_+} \frac{\mathrm{d}x}{\left[2\left(E(\beta) + V(x)\right)\right]^{1/2}}, \tag{39.92}$$

$$A(\beta) = 2\int_{x_-}^{x_+} \mathrm{d}x \left[2\left(E(\beta) + V(x)\right)\right]^{1/2} - \beta E(\beta). \tag{39.93}$$

The quantities x_+ and x_- are the zeros of $E(\beta) + V(x)$. Notice the useful relation

$$\partial A/\partial \beta = -E(\beta). \tag{39.94}$$

It is clear, at least for β large enough, that $E(\beta)$ is a negative increasing function of β. Therefore, $-\partial E/\partial \beta$ is negative and the result is real as expected:

$$\operatorname{Im} \operatorname{tr} e^{-\beta H} \sim -\frac{\beta}{2}\left(\frac{\partial E}{\partial \beta}\frac{1}{2\pi g}\right)^{1/2} e^{-A(\beta)/g} \quad \text{for } g \to 0\,. \tag{39.95}$$

This completes the calculation at 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 $V(x)$:

$$\begin{cases} V(x) \sim V_0 - \frac{1}{2}\omega^2 (x - x_0)^2 + O\left[(x - x_0^3\right], \\ V_0 > 0\,. \end{cases} \tag{39.96}$$

Let us parametrize E, x_+ and x_-,

$$E = -V_0 + \tfrac{1}{2}\omega^2\varepsilon^2, \qquad x_\pm = x_0 \pm \varepsilon\,, \tag{39.97}$$

then,

$$\beta = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{\mathrm{d}x}{\left[\omega^2\varepsilon^2 - \omega^2\left(x - x_0\right)^2\right]^{1/2}}, \ \Rightarrow\ \lim_{\varepsilon\to 0} \beta = \beta_c = 2\pi/\omega\,. \tag{39.98}$$

For $\beta \le \beta_c$, no instanton solution can be found and it is on the contrary the perturbative expansion around the classical extremum $x = x_0$ of the potential which is relevant.

39.7 Low Temperature Evaluation

Writing the equation (39.92) as

$$\beta = 2\int_{x_-}^{x_+} \left\{[2\,(V(x)+E)]^{-1/2} - \left(x^2+2E\right)^{-1/2} + \left(x^2+2E\right)^{-1/2}\right\} \mathrm{d}x\,, \qquad (39.99)$$

we can explicitly evaluate the last term and neglect E in the difference between the first two terms. This leads to

$$E(\beta) \sim -2C\,\mathrm{e}^{-\beta}, \qquad C = x_+^2 \exp\left[2\int_0^{x_+}\left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x}\right)\mathrm{d}x\right], \qquad (39.100)$$

where x_+ is now the zero of the potential. In the same way the equation (39.93) becomes

$$A(\beta) = a - 2C\,\mathrm{e}^{-\beta} + O\left(\mathrm{e}^{-2\beta}\right), \qquad a = 2\int_0^{x_+}\sqrt{2V(x)}\,. \qquad (39.101)$$

Substituting into equation (39.95), we obtain at leading order

$$\mathrm{Im}\,\mathrm{e}^{-\beta E_0(g)} \underset{g\to 0}{\sim} \frac{\beta}{2}\,\mathrm{e}^{-\beta/2}\left(\frac{C}{\pi g}\right)^{1/2}\mathrm{e}^{-a/g}, \qquad (39.102)$$

and thus,

$$\mathrm{Im}\,E_0(g) \underset{g\to 0}{\sim} -\frac{1}{2}\left(\frac{C}{\pi g}\right)^{1/2}\mathrm{e}^{-a/g}. \qquad (39.103)$$

We have calculated here only the imaginary part of the would-be ground state energy. To obtain the imaginary part of the excited levels we have to keep the correction of order $\mathrm{e}^{-\beta}$ in $A(\beta)$ for β large. We then expand $\exp[-g^{-1}A(\beta)]$ in powers of $\mathrm{e}^{-\beta}$. The coefficient of $\mathrm{e}^{-N\beta}$ in the expansion yields the imaginary part of the Nth 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 several instanton solutions corresponding to different 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 39.1 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 unstable 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 is more puzzling. First at the energy of the initial state, which is large compared to the true ground state, corresponds in the semi-classical limit an almost continuous spectrum outside the well. Moreover, in the semi-classical limit, the lifetime of the metastable state is very long. For times which 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 thè decay law will be modified.

Bibliographical Notes

Common references to Chapters 39–43 are

J. Zinn-Justin, contribution to *Recent Advances in Field Theory and Statistical Mechanics*, Les Houches 1982, R. Stora and J.-B. Zuber eds. (Elsevier, Amsterdam 1984),

which we follow closely and which also contains additional references and

S. Coleman, *The Whys of Subnuclear Physics, Erice 1977*, A. Zichichi ed. (Plenum, New York 1979).

For an early investigation of the analytic properties of the anharmonic oscillator, see

C.M. Bender and T.T. Wu, *Phys. Rev.* 184 (1969) 1231.

The content of Chapter 39 is based on

E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D15 (1977) 1544, 1558.

The method of collective coordinates was first described in this context by

J. Zittartz and J.S. Langer, *Phys. Rev.* 148 (1966) 741; J.S. Langer, *Ann. Phys.* (NY) 41 (1967) 108.

The analytic continuation in the g-plane is discussed in

G. Parisi, *Phys. Lett.* 66B (1977) 167; E.B. Bogolmony, *Phys. Lett.* 67B (1977) 193.

For the calculation of the determinant arising in the gaussian integration around the instanton see, for example,

R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York 1966, 2nd edn. Springer-Verlag, New York 1982) p. 433.

Rigorous results concerning the spectrum of the quartic anharmonic oscillator can be found in

J.J. Loeffel, A. Martin, B. Simon and A.S. Wightman, *Phys. Lett.* 30B (1968) 1656; S. Graffi, V. Grecchi and B. Simon, *Phys. Lett.* 32B (1970) 631; B. Simon, *Functional Integration and Quantum Physics* (Academic Press, New York 1979).

The shifting method is exposed in

I.M. Gel'fand and A.M. Yaglom, *J. Math. Phys.* 1 (1960) 48; R.F. Dashen, B. Hasslacher and A. Neveu, *Phys. Rev.* D10 (1974) 4114.

The appendix is inspired by the classical paper

J.H. Van Vleck, *Proc. Natl. Acad. Sci. USA* 14 (1928) 178.

APPENDIX A39

In the appendix we discuss analogous semi-classical calculations in the framework of the WKB method. For this purpose we first recall a few properties of the classical equations of motion. We finally comment about average action in the path integral.

A39.1 Classical Equations of Motion

We consider general hamiltonian systems for which the classical equations of motion can be solved with arbitrary boundary conditions. This means in particular that we can explicitly calculate the classical action on the trajectory as a function of initial and final positions $\mathbf{x}'$ and $\mathbf{x}$, and times T' and T. The classical action $\mathcal{A}$ corresponding to a hamiltonian $H(p,q,t)$ is

$$\mathcal{A}(\mathbf{p},\mathbf{q}) = \int_{T'}^{T} \mathrm{d}t \left[\mathbf{p}(t)\cdot\dot{\mathbf{q}}(t) - H\left(\mathbf{p}(t),\mathbf{q}(t);t\right)\right], \qquad (A39.1)$$

with $q_i(T') = x'_i$ and $q_i(T) = x_i$.

The action A_c corresponding to a trajectory $\{\mathbf{p}_c(t),\mathbf{q}_c(t)\}$ in phase space solution of the classical of motions is a function of the initial and final positions, and time:

$$A_c(\mathbf{x}',\mathbf{x};T) = \mathcal{A}(\mathbf{p}_c,\mathbf{q}_c). \qquad (A39.2)$$

In this appendix, in order to obtain the usual expressions of classical mechanics, we shall work in *real time*. The analytic continuation of all expressions to imaginary time is straightforward.

As is well known and will be verified below, the classical action A_c satisfies the Hamilton–Jacobi equations.

Preliminary remarks. Let us first recall a few classical results. Let us call $S(\mathbf{Q},\mathbf{q};t)$ a function satisfying the Hamilton–Jacobi equations:

$$\frac{\partial S}{\partial t} = -H\left(\frac{\partial S}{\partial \mathbf{q}},\mathbf{q};t\right), \qquad (A39.3)$$

with $S(\mathbf{Q},\mathbf{Q};T') = 0$ and with additional implicit boundary conditions at $t = T'$ which will be explained below. We then use S to generate a time-dependent canonical transformation in phase space, transforming the set $(\mathbf{p},\mathbf{q})$ into $(\mathbf{P},\mathbf{Q})$:

$$p_i = \partial S/\partial q_i\,, \quad P_i = -\partial S/\partial Q_i\,. \qquad (A39.4)$$

The implicit boundary conditions come from the conditions:

$$t = T' \to P_i = p_i \quad \text{and} \quad Q_i = q_i\,. \qquad (A39.5)$$

We make the transformation $(A39.4)$ in the action $(A39.1)$:

$$\mathcal{A}(\mathbf{p},\mathbf{q}) = \int_{T'}^{T} \left(\frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t}\right) \mathrm{d}t\,. \qquad (A39.6)$$

The quantities $\mathbf{q}$ and $\mathbf{Q}$ are now considered as time-dependent. Expression ($A39.6$) can be rewritten as

$$\mathcal{A}(\mathbf{p},\mathbf{q}) = \int_{T'}^{T} \frac{\mathrm{d}}{\mathrm{d}t} S\bigl(\mathbf{q}(t),\mathbf{Q}(t);t\bigr)\mathrm{d}t - \int_{T'}^{T} \frac{\partial S}{\partial Q_i}\dot{Q}_i \mathrm{d}t\,. \qquad (A39.7)$$

Using again relations ($A39.4$), we finally obtain

$$\mathcal{A}(\mathbf{p},\mathbf{q}) = S(\mathbf{x},\mathbf{Q}(T);T) + \int_{T'}^{T} P_i\dot{Q}_i \mathrm{d}t\,. \qquad (A39.8)$$

The equations of motion are now trivial,

$$\dot{Q}_i = 0 \Longrightarrow Q_i(t) = Q_i(T'). \qquad (A39.9)$$

The conditions ($A39.5$) determine the solution

$$Q_i(t) = x'_i\,. \qquad (A39.10)$$

We have, therefore, shown that

$$A_c(\mathbf{x}',\mathbf{x};T) \equiv S(\mathbf{x},\mathbf{x}';T), \qquad (A39.11)$$

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

$$H = \tfrac{1}{2}\sum p_i^2\,, \qquad (A39.12)$$

or a harmonic oscillator

$$H = \tfrac{1}{2}\sum_i \left(p_i^2 + q_i^2\right), \qquad (A39.13)$$

we can map the original hamiltonian system onto any convenient hamiltonian.

Let us also verify that transformation ($A39.4$) leaves the measure in phase space invariant. We perform the transformation in two steps. First we go from p_i to Q_i:

$$\prod_i \mathrm{d}p_i\mathrm{d}q_i = \prod_i \mathrm{d}q_i\mathrm{d}Q_i \det \frac{\partial^2 S}{\partial q_i \partial Q_j}. \qquad (A39.14)$$

We now eliminate q_i in favour of P_i:

$$\frac{\partial q_i}{\partial P_j} = \left[\frac{\partial P}{\partial q}\right]^{(-1)}_{ij} = -\left[\frac{\partial^2 S}{\partial q \partial Q}\right]^{(-1)}_{ij}. \qquad (A39.15)$$

Therefore, the second jacobian cancels the first one (note that a similar argument directly shows the invariance already proven above of the symplectic form $\mathrm{d}p_i \wedge \mathrm{d}q_i$). This analysis suggests that we can perform two transformations ($A39.4$) 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:

$$\langle \mathbf{x}' | U(T', T) | \mathbf{x} \rangle = \int \prod_i [\mathrm{d}q_i(t)\mathrm{d}p_i(t)] \exp [i\mathcal{A}(\mathbf{p}, \mathbf{q})] , \qquad (A39.16)$$
$$\text{with } q_i(T') = x'_i , \quad q_i(T) = x_i .$$

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. This would in particular apply to systems with one degree of freedom, with a hamiltonian H of the form

$$H = \tfrac{1}{2}p^2 + V(q) . \qquad (A39.17)$$

Unfortunately it is easy to verify that the result is wrong at least in Quantum Mechanics. Actually the whole procedure is somewhat ill-defined. This comes from the fact that 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 about 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 on the discretized form of the phase space path integral. It should be possible in this way to derive the semi-classical result which we shall now establish, but, to our knowledge, no one has yet done so. On the other hand, as often, the situation seems to be more favourable in field theory. The problem we have described above comes mainly from commutation of quantum operators. We have emphasized in Chapter 7 that the commutators are infinite in field theory and disappear in the renormalization. Therefore one expects, as this has been verified on examples, that the semi-classical approximations of classically integrable field theories reproduce features of the exact solution.

A39.2 The WKB Method

We now explicitly write the Schrödinger equation for the evolution operator

$$H\left(\frac{\hbar}{i}\frac{\partial}{\partial \mathbf{x}}, \mathbf{x}; t\right) U(\mathbf{x}', \mathbf{x}; t) = i\hbar \frac{\partial U}{\partial t}(\mathbf{x}', \mathbf{x}; t), \qquad (A39.18)$$

with the definition

$$U(\mathbf{x}', \mathbf{x}; t) = \langle \mathbf{x}' | \mathbf{U}(t) | \mathbf{x} \rangle , \qquad (A39.19)$$

and the boundary condition

$$\mathbf{U}(t = T') = \mathbf{1} . \qquad (A39.20)$$

In order to write equation ($A39.18$), it has been necessary to associate with the classical function $H(\mathbf{p}, \mathbf{q}, ; t)$ a quantum operator.

We show below how much the semi-classical result depends upon the quantization. We have reintroduced the quantity $\hbar$ which we usually set equal to one, to make the expansion parameter explicit. The ansatz will be to set

$$U(\mathbf{x}', \mathbf{x}; t) = G(\mathbf{x}', \mathbf{x}; t)\, \mathrm{e}^{iA(\mathbf{x}', \mathbf{x}; t)/\hbar} [1 + O(\hbar)] . \qquad (A39.21)$$

Introducing the ansatz into equation ($A39.18$) and keeping the two first terms in $\hbar$, we obtain two equations. The first equation is

$$H\left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t\right) = -\frac{\partial A}{\partial t}. \tag{A39.22}$$

This equation involves only the classical hamiltonian and is the Hamilton–Jacobi equation for the classical action on the classical trajectory. Together with the boundary conditions implied by the condition ($A39.20$), it determines A completely. The derivation of the second equation involves some more work. Let us first note that

$$H\left(\frac{\hbar}{i}\frac{\partial}{\partial \mathbf{x}}, \mathbf{x}; t\right) G = GH - i\hbar \frac{\partial H}{\partial p_i}\frac{\partial G}{\partial x_i} + O\left(\hbar^2\right). \tag{A39.23}$$

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:

$$\begin{aligned} \mathrm{e}^{-iA/\hbar} H\, \mathrm{e}^{iA/\hbar} = H\left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t\right) &- \frac{i\hbar}{2}\frac{\partial^2 H}{\partial p_j \partial q_j}\left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t\right) \\ &- \frac{i\hbar}{2}\frac{\partial^2 H}{\partial p_j \partial p_k}\left(\frac{\partial A}{\partial \mathbf{x}}, \mathbf{x}; t\right)\frac{\partial^2 A}{\partial x_j \partial x_k} + O\left(\hbar^2\right). \end{aligned} \tag{A39.24}$$

The second term in the r.h.s. comes from commuting all the derivatives completely on the right. It relies on the assumption that the classical hamiltonian H is real and can be quantized to generate a hermitian operator. Indeed let us first assume that we have symmetrized all monomials:

$$p^n q^m \to \tfrac{1}{2}\left(p^n q^m + q^m p^n\right). \tag{A39.25}$$

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 ($A39.25$). 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 ($A39.24$) arises from two derivatives acting on the action. The factor 1/2 is a counting factor. We then obtain an equation for G:

$$\frac{\partial H}{\partial p_i}\frac{\partial G}{\partial x_i} + \frac{1}{2}\left(\frac{\partial^2 H}{\partial p_j \partial q_j} + \frac{\partial^2 H}{\partial p_j \partial p_k}\frac{\partial^2 A}{\partial x_j \partial x_k}\right) G = -\frac{\partial G}{\partial t}. \tag{A39.26}$$

Let us now differentiate equation ($A39.22$) with respect to x'_i and x_j successively:

$$\frac{\partial^2 A}{\partial x'_i \partial x_k}\frac{\partial H}{\partial p_k} = -\frac{\partial^2 A}{\partial t \partial x'_i}, \tag{A39.27}$$

$$\frac{\partial^3 A}{\partial x'_i \partial x_j \partial x_k}\frac{\partial H}{\partial p_k} + \frac{\partial^2 A}{\partial x'_i \partial x_k}\frac{\partial^2 A}{\partial x_j \partial x_l}\frac{\partial^2 H}{\partial p_k \partial p_l} + \frac{\partial^2 A}{\partial x'_i \partial x_k}\frac{\partial^2 H}{\partial p_k \partial q_j} = -\frac{\partial^3 A}{\partial t \partial x'_i \partial x_j}. \tag{A39.28}$$

Let us introduce a matrix notation

$$M_{ij} = \frac{\partial^2 A}{\partial x_i' \partial x_j}, \quad H_{ij} = \frac{\partial^2 H}{\partial p_i \partial p_j}, \quad \widetilde{H}_{ij} = \frac{\partial^2 H}{\partial p_i \partial q_j}. \tag{A39.29}$$

The equation ($A39.28$) can then be rewritten as

$$\frac{\partial \mathbf{M}}{\partial x_k}\frac{\partial H}{\partial p_k} + \mathbf{MHM} + \mathbf{M}\widetilde{\mathbf{H}} = -\frac{\partial \mathbf{M}}{\partial t}. \tag{A39.30}$$

All multiplications are meant in a matrix sense. We now multiply equation ($A39.30$) by $\mathbf{M}^{-1}$ on the left and take the trace:

$$\frac{\partial H}{\partial p_k} \operatorname{tr} \mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial x_k} + \operatorname{tr}\left(\mathbf{HM} + \widetilde{\mathbf{H}}\right) = -\operatorname{tr} \mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial t}. \tag{A39.31}$$

Let us note that

$$\frac{\partial}{\partial z} \ln \det \mathbf{M}(z) = \frac{\partial}{\partial z} \operatorname{tr} \ln \mathbf{M}(z) = \operatorname{tr} \frac{\partial \mathbf{M}}{\partial z}\mathbf{M}^{-1}. \tag{A39.32}$$

The equation ($A39.31$) can then be rewritten as

$$\frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} \ln \det \mathbf{M} + \frac{\partial^2 H}{\partial p_j \partial q_j} + \frac{\partial^2 H}{\partial p_j \partial p_k}\frac{\partial^2 A}{\partial x_j \partial x_k} = -\frac{\partial}{\partial t} \ln \det \mathbf{M}, \tag{A39.33}$$

while equation ($A39.26$) can be written as

$$\frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} \ln G + \frac{1}{2}\frac{\partial^2 H}{\partial p_j \partial q_j} + \frac{1}{2}\frac{\partial^2 H}{\partial p_j \partial p_k}\frac{\partial^2 A}{\partial x_j \partial x_k} = -\frac{\partial}{\partial t} \ln G\,. \tag{A39.34}$$

A comparison between these two equations shows that a solution to equation ($A39.34$) is

$$\ln G = \tfrac{1}{2} \ln \det \mathbf{M} + \text{ const.}\,. \tag{A39.35}$$

Taking into account the boundary conditions we finally obtain Van Vleck's formula:

$$\langle \mathbf{x}' | U(T', T) | \mathbf{x} \rangle \sim \frac{1}{(2\pi i \hbar)^{n/2}} \left(-\det \frac{\partial^2 A}{\partial \mathbf{x} \partial \mathbf{x}'}\right)^{1/2} \mathrm{e}^{(i/\hbar)A}\,. \tag{A39.36}$$

It is straightforward to derive from this equation the corresponding expression for imaginary time.

Let us introduce a matrix notation

$$M_{ij} = \frac{\partial^2 A}{\partial x'_i \partial x_j}, \quad H_{ij} = \frac{\partial^2 H}{\partial p_i \partial p_j}, \quad \widetilde{H}_{ij} = \frac{\partial^2 H}{\partial p_i \partial q_j}. \tag{A39.29}$$

The equation ($A39.28$) can then be rewritten as

$$\frac{\partial \mathbf{M}}{\partial x_k}\frac{\partial H}{\partial p_k} + \mathbf{MHM} + \mathbf{M}\widetilde{\mathbf{H}} = -\frac{\partial \mathbf{M}}{\partial t}. \tag{A39.30}$$

All multiplications are meant in a matrix sense. We now multiply equation ($A39.30$) by $\mathbf{M}^{-1}$ on the left and take the trace:

$$\frac{\partial H}{\partial p_k} \operatorname{tr} \mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial x_k} + \operatorname{tr}\left(\mathbf{HM} + \widetilde{\mathbf{H}}\right) = -\operatorname{tr} \mathbf{M}^{-1}\frac{\partial \mathbf{M}}{\partial t}. \tag{A39.31}$$

Let us note that

$$\frac{\partial}{\partial z} \ln \det \mathbf{M}(z) = \frac{\partial}{\partial z} \operatorname{tr} \ln \mathbf{M}(z) = \operatorname{tr} \frac{\partial \mathbf{M}}{\partial z}\mathbf{M}^{-1}. \tag{A39.32}$$

The equation ($A39.31$) can then be rewritten as

$$\frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} \ln \det \mathbf{M} + \frac{\partial^2 H}{\partial p_j \partial q_j} + \frac{\partial^2 H}{\partial p_j \partial p_k}\frac{\partial^2 A}{\partial x_j \partial x_k} = -\frac{\partial}{\partial t} \ln \det \mathbf{M}, \tag{A39.33}$$

while equation ($A39.26$) can be written as

$$\frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} \ln G + \frac{1}{2}\frac{\partial^2 H}{\partial p_j \partial q_j} + \frac{1}{2}\frac{\partial^2 H}{\partial p_j \partial p_k}\frac{\partial^2 A}{\partial x_j \partial x_k} = -\frac{\partial}{\partial t} \ln G\,. \tag{A39.34}$$

A comparison between these two equations shows that a solution to equation ($A39.34$) is

$$\ln G = \tfrac{1}{2} \ln \det \mathbf{M} + \text{ const.}\,. \tag{A39.35}$$

Taking into account the boundary conditions we finally obtain Van Vleck's formula:

$$\langle \mathbf{x}' \,| U(T',T) |\, \mathbf{x} \rangle \sim \frac{1}{(2\pi i \hbar)^{n/2}} \left(-\det \frac{\partial^2 A}{\partial \mathbf{x} \partial \mathbf{x}'} \right)^{1/2} \mathrm{e}^{(i/\hbar)A}\,. \tag{A39.36}$$

It is straightforward to derive from this equation the corresponding expression for imaginary time.

A39.3 The Average Action in Path Integrals

Since we have emphasized finite action solutions, one could have the impression that the result of the path integral is dominated by finite action paths. As the two examples below show, this is never the case: the average action is infinite, though the path integral is dominated by paths which are in the neighbourhood of finite action solutions.

(i) ***The average action in a path integral is infinite.*** We give first a very simple example to illustrate this point:

$$I = \lim_{N\to\infty} I_N, \quad I_N = (2\pi)^{-N/2} \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2\right). \tag{A39.37}$$

Obviously $I = 1$. Setting

$$\sum_{i=1}^{N} x_i^2 = R^2, \tag{A39.38}$$

we obtain

$$I_N = (2\pi)^{-N/2} \frac{2\pi^{N/2}}{\Gamma(N/2)} \int_0^\infty R^{N-1}\mathrm{d}R\, \mathrm{e}^{-R^2/2}. \tag{A39.39}$$

Finite action means that R is bounded:

$$I_N(R_0) = \frac{2^{1-N/2}}{\Gamma(N/2)} \int_0^{R_0} R^{N-1}\mathrm{d}R\, \mathrm{e}^{-R^2/2} < \left(\frac{R_0^2}{2}\right)^{N/2} \frac{1}{\Gamma(1+N/2)}. \tag{A39.40}$$

Therefore,

$$\lim_{N\to\infty} I_N(R_0) = 0. \tag{A39.41}$$

More precisely the integral ($A39.39$) can be calculated by steepest descent. The saddle point R_s is

$$R_s = N^{1/2}, \tag{A39.42}$$

and the classical actions A which are relevant to the path integral satisfy

$$A = \tfrac{1}{2}N + O\left(N^{1/2}\right). \tag{A39.43}$$

(ii) *Average action in Quantum Mechanics.* We consider the path integral:

$$\operatorname{tr} \mathrm{e}^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q(t)] \exp\left\{-\int_{-\beta/2}^{\beta/2} \mathrm{d}t \left[\frac{1}{2}\dot{q}^2(t) + V\left(q(t)\right)\right]\right\}. \tag{A39.44}$$

The average action $\langle S(q)\rangle$ contains two contributions, the average potential, which is finite in general, and the average kinetic term. We have shown in Section A2.1 quite generally

$$\left\langle \left(q(t+\varepsilon) - q(t)\right)^2 \right\rangle \underset{\varepsilon\to 0}{\sim} |\varepsilon|.$$

Therefore,

$$\left\langle \frac{1}{2}\int_{-\beta/2}^{\beta/2} \mathrm{d}t\, \dot{q}^2(t) \right\rangle = \lim_{\varepsilon\to 0} \frac{1}{2}\int_{-\beta/2}^{\beta/2} \mathrm{d}t \left\langle \left(q(t+\varepsilon) - q(t)\right)^2 \right\rangle / \varepsilon^2 \sim \frac{\beta}{2\varepsilon} \to \infty.$$

A39.3 The Average Action in Path Integrals

Since we have emphasized finite action solutions, one could have the impression that the result of the path integral is dominated by finite action paths. As the two examples below show, this is never the case: the average action is infinite, though the path integral is dominated by paths which are in the neighbourhood of finite action solutions.

(i) *The average action in a path integral is infinite.* We give first a very simple example to illustrate this point:

$$I = \lim_{N\to\infty} I_N, \quad I_N = (2\pi)^{-N/2} \int \mathrm{d}x_1 \cdots \mathrm{d}x_N \exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2\right). \tag{A39.37}$$

Obviously $I = 1$. Setting

$$\sum_{i=1}^{N} x_i^2 = R^2, \tag{A39.38}$$

we obtain

$$I_N = (2\pi)^{-N/2} \frac{2\pi^{N/2}}{\Gamma(N/2)} \int_0^\infty R^{N-1} \mathrm{d}R\, \mathrm{e}^{-R^2/2}. \tag{A39.39}$$

Finite action means that R is bounded:

$$I_N(R_0) = \frac{2^{1-N/2}}{\Gamma(N/2)} \int_0^{R_0} R^{N-1}\mathrm{d}R\, \mathrm{e}^{-R^2/2} < \left(\frac{R_0^2}{2}\right)^{N/2} \frac{1}{\Gamma(1+N/2)}. \tag{A39.40}$$

Therefore,

$$\lim_{N\to\infty} I_N(R_0) = 0. \tag{A39.41}$$

More precisely the integral (A39.39) can be calculated by steepest descent. The saddle point R_s is

$$R_s = N^{1/2}, \tag{A39.42}$$

and the classical actions A which are relevant to the path integral satisfy

$$A = \tfrac{1}{2}N + O\left(N^{1/2}\right). \tag{A39.43}$$

(ii) *Average action in Quantum Mechanics.* We consider the path integral:

$$\operatorname{tr} \mathrm{e}^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q(t)] \exp\left\{-\int_{-\beta/2}^{\beta/2} \mathrm{d}t \left[\frac{1}{2}\dot{q}^2(t) + V\big(q(t)\big)\right]\right\}. \tag{A39.44}$$

The average action $\langle S(q)\rangle$ contains two contributions, the average potential, which is finite in general, and the average kinetic term. We have shown in Section A2.1 quite generally

$$\left\langle \big(q(t+\varepsilon) - q(t)\big)^2 \right\rangle \underset{\varepsilon\to 0}{\sim} |\varepsilon|.$$

Therefore,

$$\left\langle \frac{1}{2}\int_{-\beta/2}^{\beta/2} \mathrm{d}t\, \dot{q}^2(t) \right\rangle = \lim_{\varepsilon\to 0} \frac{1}{2} \int_{-\beta/2}^{\beta/2} \mathrm{d}t \left\langle \big(q(t+\varepsilon) - q(t)\big)^2 \right\rangle / \varepsilon^2 \sim \frac{\beta}{2\varepsilon} \to \infty.$$

40 UNSTABLE VACUA IN QUANTUM FIELD THEORY

We begin with this chapter a semi-classical study of barrier penetration in quantum field theory, generalizing the methods explained in quantum mechanics. We have shown that in quantum mechanics barrier penetration is associated with classical motion in imaginary time; it is thus natural here to consider quantum field theory in its euclidean formulation.

In the functional representation of field theory barrier penetration in the semi-classical limit is also related to finite action solutions (instantons) of the classical field equations. We first try to characterize such solutions. We then explain how to evaluate the instanton contributions at leading order, the main new problem arising from UV divergences.

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. Dividing by the volume we, therefore, obtain the probability per unit time and unit volume of a metastable pseudo-vacuum to decay.

We first discuss a scalar field theory with a ϕ^4 interaction, generalization of the quartic anharmonic oscillator considered in Chapter 39, in the dimensions in which it is super-renormalizable, that is, two and three dimensions. We then consider more general scalar field theories, of a form analogous to the quantum mechanical models discussed in Section 39.5.

We finally calculate instanton contributions at leading order explicitly in the ϕ^4 in four dimensions, the dimension in which the theory is renormalizable. Several new problems arise. As we show in Appendix A40.2, using Sobolev inequalities, in dimension 4 the massive field equation has no instanton solution and the relevant instanton is a solution of the massless field equation instead. We, therefore, first study the massless ϕ^4 theory 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 properties of the theory.

The last section 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 theory, and RG properties and conformal invariance relevant to the instanton calculations in ϕ^4_4.

40.1 The ϕ^4 Field Theory for Negative Coupling

We consider the d-dimensional field theory for a scalar field ϕ, corresponding in the tree approximation to the action

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left(\frac{1}{2} \left(\partial_\mu \phi(x)\right)^2 + \frac{1}{2} m^2 \phi^2(x) + \frac{1}{4!} g m^{4-d} \phi^4(x) \right), \tag{40.1}$$

m being the mass and g the dimensionless coupling constant (the power of m which appears in front of the interaction term ϕ^4 takes care of the dimension).

The complete n-point correlation function has the functional representation:

$$Z^{(n)}(x_1, \dots, x_n) = \int [\mathrm{d}\phi(x)]\, \phi(x_1)\phi(x_2)\dots\phi(x_n) \exp[-\mathcal{S}(\phi)]. \tag{40.2}$$

We normalize the functional 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 used in Chapter 39, we assume that we start from positive values of g and proceed by analytic continuation to define the functional 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-trivial saddle points contribute to the imaginary part. Therefore, we look for non-trivial finite action solutions of the euclidean field equations, that is, instanton configurations, and then calculate the corresponding contributions.

40.1.1 *Instantons: classical solutions and classical action*

The instanton solutions. The field equation corresponding to the action (40.1) is

$$\left(-\nabla^2 + m^2\right) \phi(x) + \tfrac{1}{6} g m^{4-d} \phi^3(x) = 0\,. \tag{40.3}$$

We set (g is negative),

$$\phi(x) = (-6/g)^{1/2} m^{d/2-1} f(mx)\,. \tag{40.4}$$

In terms of f the classical action (40.1) reads

$$\mathcal{S}(f) = -\frac{6}{g} \int \mathrm{d}^d x \left[\tfrac{1}{2} \left(\partial_\mu f\right)^2 + \tfrac{1}{2} f^2 - \tfrac{1}{4} f^4 \right]. \tag{40.5}$$

The function $f(x)$ then satisfies a parameter-free equation,

$$\left(-\nabla^2 + 1\right) f(x) - f^3(x) = 0\,. \tag{40.6}$$

It can be shown (for details see Appendix A40.2) that the solution with the smallest action is spherically symmetric. We therefore choose an arbitrary origin x_0 and set

$$r = |x - x_0|\,. \tag{40.7}$$

A function $f(x)$ which depends only on the radial variable r satisfies the differential equation,

$$\left[-\left(\frac{\mathrm{d}}{\mathrm{d}r}\right)^2 - \frac{d-1}{r}\frac{\mathrm{d}}{\mathrm{d}r} + 1 \right] f(r) - f^3(r) = 0\,. \tag{40.8}$$

Interpreting r as a time, we note that this equation describes the motion of a particle in a potential $-V(f)$:

$$V(f) = \tfrac{1}{2}f^2 - \tfrac{1}{4}f^4, \tag{40.9}$$

submitted in addition to a viscous damping force.

Since we look for finite action solutions we have to impose the boundary condition

$$f(r) \to 0 \quad \text{for} \quad r \to \infty\,. \tag{40.10}$$

Equation (40.8) shows that if $f(r)$ goes to zero at infinity 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 tends at infinity towards a minimum of the potential $f = \pm 1$. The condition (40.10) is satisfied only for a discrete set of initial values of $f(0)$. It can moreover be shown that the minimal action solution corresponds to the function for which $|f(0)|$ is minimal in the set, and which vanishes only at infinity.

The classical action. We denote by f_c a solution. Since g is dimensionless the corresponding classical action has the form

$$\mathcal{S}(\phi_c) \equiv \mathcal{S}(f_c) = -A/g\,, \tag{40.11}$$

in which A is a pure number. The considerations of Section A40.1 lead to the relations

$$A = \frac{6}{d}\int [\partial_\mu f(x)]^2\,\mathrm{d}^d x = \frac{3}{2}\int f^4(x)\mathrm{d}^d x = \frac{6}{4-d}\int f^2(x)\mathrm{d}^d x\,, \tag{40.12}$$

which show that A is positive. We also note that these relations can only be true for $d < 4$ and thus the dimension 4 is singular (see Section A40.2).

We give here the numerical results for $d = 2$ and $d = 3$:

$$d = 2: \quad f_c(0) = 2.20620086465 \quad A = 35.102689573 \tag{40.13}$$

$$d = 3: \quad f_c(0) = 4.33738767997 \quad A = 113.383507815 \tag{40.14}$$

40.1.2 The gaussian integration

We want to perform the gaussian integration in the neighbourhood of the saddle point. This involves studying the eigenvalues of the operator $\mathbf{M}$:

$$\langle x\,|\mathbf{M}|\,x'\rangle \equiv M(x,x') = \frac{\delta^2\mathcal{S}}{\delta\phi_c(x)\delta\phi_c(x')}, \tag{40.15}$$

$$= \left[\left(-\nabla_x^2 + m^2\right) + \tfrac{1}{2}gm^{4-d}\phi_c^2(x)\right]\delta(x-x'), \tag{40.16}$$

or, introducing the solution $f(x)$,

$$M(x,x') = \left[\left(-\nabla_x^2 + m^2\right) - 3m^2 f^2(mx)\right]\delta(x-x'). \tag{40.17}$$

Differentiating the equation of motion (40.3) with respect to x_μ, we learn that the functions $\partial_\mu\phi_c$ are d eigenvectors of $\mathbf{M}$ with vanishing eigenvalue:

$$\left(-\nabla^2 + m^2\right)\partial_\mu\phi_c(x) + \tfrac{1}{2}gm^{4-d}\phi_c^2(x)\partial_\mu\phi_c(x) = 0 \quad \Longleftrightarrow \quad \mathbf{M}\,\partial_\mu\phi_c = 0\,. \tag{40.18}$$

This property had to be expected. Due to translation symmetry we have found a family of degenerate saddle points $\phi_c(x)$ depending on d parameters $x_{0\mu}$ (equation (40.7)). As in quantum mechanics, we have to sum over all saddle points, and thus to take the collective coordinates $x_{0\mu}$ as d of our integration variables, over which we eventually integrate exactly. To change variables, we can for example apply the identity (39.31) to each variable. As the result (39.35) shows, the change of variables leads to the determinant of $\mathbf{M}$ in the subspace orthogonal to the zero eigenvalue sector, and a jacobian J, which, at leading order, is

$$J = \prod_{\mu=1}^{d} \|\partial_\mu \phi_c(x)\| , \tag{40.19}$$

and a factor $(2\pi)^{-1/2}$ from the gaussian integration, for each variable.

Since the solution is spherically symmetric, we can rewrite J as

$$J = \left[\frac{1}{d}\int \mathrm{d}^d x \sum_\mu (\partial_\mu \phi_c)^2\right]^{d/2} . \tag{40.20}$$

Using the relation (40.12) we can express the jacobian in terms of the classical action

$$J = (-A/g)^{d/2} . \tag{40.21}$$

Note one important feature of this expression: each component of x_0 has generated a factor $(-g)^{-1/2}$. Finally, in expression (40.2) we 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 we find,

$$\operatorname{Im} Z^{(n)}(x_1,\ldots,x_n) = \frac{1}{2i}\left(\frac{A}{2\pi}\right)^{d/2} \Omega \frac{\mathrm{e}^{A/g}}{(-g)^{(d+n)/2}} F_n(x_1,\ldots,x_n), \tag{40.22}$$

with

$$F_n(x_1,\ldots,x_n) = m^{d+n(d-2)/2} 6^{n/2} \int \mathrm{d}^d x_0 \prod_{i=1}^{n} f\big(m(x_i - x_0)\big), \tag{40.23}$$

and

$$\langle x\,|\mathbf{M}_0|\,x'\rangle = \left(-\nabla_x^2 + m^2\right)\delta(x-x'), \tag{40.24a}$$

$$\Omega = \left(\det{}' \mathbf{M}\mathbf{M}_0^{-1}\right)^{-1/2}\Big|_{m=1} = \lim_{\epsilon\to 0} \epsilon^{-d} \det\left[(\mathbf{M}+\epsilon)\,\mathbf{M}_0^{-1}\right]\big|_{m=1} . \tag{40.24b}$$

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.

Wave function arguments of the kind used for the Schrödinger equation show that $\partial_\mu \phi_c$ is not the ground state of $\mathbf{M}$. There is one state with a negative eigenvalue so that the final result is real as expected. In Appendix A40.2 we give a proof of this property using Sobolev inequalities.

Discussion. A few comments concerning expression (40.22) are here in order. We have obtained a result for the complete correlation functions, improperly normalized, for convenience, with respect to the free field theory. We notice, however, that, 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 (40.1) have the same behaviour as those renormalized with respect to the free field theory.

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 a power of g. If we call $W^{(n)}(x_1,\ldots,x_n)$ the connected n-point function, we thus find at leading order:

$$\operatorname{Im} W^{(n)} \sim \operatorname{Im} Z^{(n)},$$

a result that is consistent with the observation that the explicit expression (40.22) is indeed connected. To pass from connected correlation functions to 1PI functions, we have first to subtract the reducible contributions which involve functions with a smaller number of arguments and which are, therefore, negligible at leading order, and then to amputate the remaining part. Again for the same reason only the real part of the propagator matters; to amputate expression (40.22) we, therefore, have to simply multiply it by the product of the inverse free propagators corresponding to each external line. Introducing $\tilde{f}(p/m)$, the Fourier transform of $f(mx)$, and writing the n-point 1PI function $\tilde{\Gamma}^{(n)}$ in momentum space representation, we obtain

$$\operatorname{Im} \tilde{\Gamma}^{(n)}(p_1,\ldots,p_n) \sim -\frac{1}{2i}\left(\frac{A}{2\pi}\right)^{d/2} \Omega \frac{e^{A/g}}{(-g)^{(d+n)/2}} m^{d-n(d/2+1)} \prod_{i=1}^{n} \sqrt{6}\tilde{f}\left(\frac{p_i}{m}\right)(p_i^2+m^2). \tag{40.25}$$

The structure, at leading order, of the imaginary part of the n-point function is particularly simple in momentum space, in particular it depends only on the square of the momenta and not of their scalar products.

Up to this point the discussions of the ϕ^4 field theory and of the anharmonic oscillator have been remarkably similar. Now comes one important difference: as we shall see the determinant of the operator $\mathbf{M}$ is actually UV divergent and we shall have to deal with this new problem.

40.1.3 Renormalization

To define properly the ϕ^4 theory in two and three dimensions, we have first to introduce a UV cut-off and then to add to the classical action a mass counterterm before taking the infinite cut-off limit: the cut-off dependent action $\mathcal{S}_\Lambda$ has the form:

$$\mathcal{S}_\Lambda(\phi) = \int d^d x \left[\tfrac{1}{2}\phi(x)\left(-\nabla^2 + \nabla^4/\Lambda^2 + m^2\right)\phi(x) + \tfrac{1}{4!}g\phi^4(x) + \tfrac{1}{2}\delta m^2(\Lambda)\,\phi^2(x)\right]. \tag{40.26}$$

Let us examine, at cut-off Λ large but fixed, the effect of these modifications. The additional term

$$\frac{1}{\Lambda^2}\int \phi(x)\nabla^4\phi(x)d^d x\,,$$

modifies the equation of motion but when Λ becomes large this modification vanishes like $1/\Lambda^2$.

The counterterm on the other hand 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 counterterm does not contribute to the classical equation of motion.

$\phi_c(x)$ 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 (40.1) have the same behaviour as those renormalized with respect to the free field theory.

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 a power of g. If we call $W^{(n)}(x_1,\ldots,x_n)$ the connected n-point function, we thus find at leading order:

$$\operatorname{Im} W^{(n)} \sim \operatorname{Im} Z^{(n)},$$

a result that is consistent with the observation that the explicit expression (40.22) is indeed connected. To pass from connected correlation functions to 1PI functions, we have first to subtract the reducible contributions which involve functions with a smaller number of arguments and which are, therefore, negligible at leading order, and then to amputate the remaining part. Again for the same reason only the real part of the propagator matters; to amputate expression (40.22) we, therefore, have to simply multiply it by the product of the inverse free propagators corresponding to each external line. Introducing $\tilde{f}(p/m)$, the Fourier transform of $f(mx)$, and writing the n-point 1PI function $\tilde{\Gamma}^{(n)}$ in momentum space representation, we obtain

$$\operatorname{Im} \tilde{\Gamma}^{(n)}(p_1,\ldots,p_n) \sim -\frac{1}{2i}\left(\frac{A}{2\pi}\right)^{d/2} \Omega \frac{\mathrm{e}^{A/g}}{(-g)^{(d+n)/2}} m^{d-n(d/2+1)} \prod_{i=1}^{n} \sqrt{6}\tilde{f}\left(\frac{p_i}{m}\right)(p_i^2+m^2). \tag{40.25}$$

The structure, at leading order, of the imaginary part of the n-point function is particularly simple in momentum space, in particular it depends only on the square of the momenta and not of their scalar products.

Up to this point the discussions of the ϕ^4 field theory and of the anharmonic oscillator have been remarkably similar. Now comes one important difference: as we shall see the determinant of the operator $\mathbf{M}$ is actually UV divergent and we shall have to deal with this new problem.

40.1.3 Renormalization

To define properly the ϕ^4 theory in two and three dimensions, we have first to introduce a UV cut-off and then to add to the classical action a mass counterterm before taking the infinite cut-off limit: the cut-off dependent action $\mathcal{S}_\Lambda$ has the form:

$$\mathcal{S}_\Lambda(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\phi(x)\left(-\nabla^2 + \nabla^4/\Lambda^2 + m^2\right)\phi(x) + \tfrac{1}{4!}g\phi^4(x) + \tfrac{1}{2}\delta m^2(\Lambda)\,\phi^2(x)\right]. \tag{40.26}$$

Let us examine, at cut-off Λ large but fixed, the effect of these modifications. The additional term

$$\frac{1}{\Lambda^2}\int \phi(x)\nabla^4\phi(x)\mathrm{d}^d x\,,$$

modifies the equation of motion but when Λ becomes large this modification vanishes like $1/\Lambda^2$.

The counterterm on the other hand 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 counterterm does not contribute to the classical equation of motion.

On the other hand, if we calculate the contribution of the counterterm to the classical action we find that the one-loop counterterm, which is proportional to g, gives a contribution of order 1 in g because $\phi_c(x)$ is proportional to $1/\sqrt{-g}$. It, therefore, generates an additional multiplicative factor. We now consider the determinant of $\mathbf{M}$ in the regularized theory:

$$M(x,x') = \left[\left(-\nabla^2 + \nabla^4/\Lambda^2 + m^2\right) + \tfrac{1}{2} g m^{4-d} \phi_c^2(x)\right] \delta(x-x'), \tag{40.27}$$

We can expand it in powers of $\phi_c^2(x)$ using the identity $\operatorname{tr}\ln = \ln\det$:

$$\begin{aligned}\ln\det \mathbf{M} &= \ln\det\left(-\nabla^2 + \nabla^4/\Lambda^2 + m^2\right) \\ &\quad - \sum_{k=1}^{\infty} \frac{1}{k} \operatorname{tr}\left[\tfrac{1}{2} g m^{4-d} \phi_c^2(x) \left(\nabla^2 - \nabla^4/\Lambda^2 - m^2\right)^{-1}\right]^k .\end{aligned} \tag{40.28}$$

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. Let us rewrite more explicitly the $k=1$ term:

$$\tfrac{1}{2} \operatorname{tr} g m^{4-d} \phi_c^2(x) \left(-\nabla^2 + \nabla^4/\Lambda^2 + m\right)^{-1} = \tfrac{1}{2} g m^{4-d} D \int \mathrm{d}^d x\, \phi_c^2(x)\,, \tag{40.29}$$

in which D is the free regularized propagator at coinciding arguments,

$$D = \frac{1}{(2\pi)^d} \int \frac{\mathrm{d}^d p}{p^2 + p^4/\Lambda^2 + m^2}. \tag{40.30}$$

The determinant of the operator $\mathbf{M}$, therefore, contains a factor which diverges at large cut-off:

$$\exp\left[-\frac{1}{4} g m^{4-d} D \int \mathrm{d}^d x\, \phi_c^2(x)\right]. \tag{40.31}$$

This factor exactly cancels the infinite factor coming from the counterterm, so that the final expression for the imaginary part is finite. The fact that really we have to calculate $\det' \mathbf{M}$ does not change this argument, because UV divergences are insensitive to the omission of a finite number of eigenvalues of $\mathbf{M}$, as the second form (40.24*b*) explicitly shows.

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 function concentrated around small fields. However, because we have expanded around a relative minimum of the potential, this state actually is unstable. We have calculated its decay rate due to barrier penetration. We note that the integral over $x_{0\mu}$ in equation (40.23) 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 would-be ground state ("false vacuum") of the theory to decay. Some implications of such a result will be discussed in a slightly more general context in Section 40.8.

40.2 General Potentials: Instanton Contributions

We now extend the analysis of Section 39.5 to the analogous scalar field theories, using the techniques developed in Section 40.1. We consider an euclidean action of the form:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left(\tfrac{1}{2} (\partial_\mu \phi)^2 + g^{-1} V(\phi\sqrt{g}) \right), \tag{40.32}$$

in which the potential $V(\phi)$ has a stable and an unstable minimum, and is of the type discussed in Section 39.5. Assuming that at some initial time the quantum mechanical state corresponds to a field concentrated around the unstable minimum of the potential, the "false" vacuum, we want to evaluate semi-classically 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 and a gaussian integration around the instanton.

40.2.1 Calculation of the instanton contribution

We define the field in such a way that the unstable minimum corresponds to $\phi = 0$. The discussion of the existence of an instanton solution is similar to the one given in Section 40.1. A theorem establishes, under mild assumptions, that spherically symmetric solutions give the minimal action. We, therefore, 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

$$\frac{\mathrm{d}^2 f}{\mathrm{d}r^2} + \frac{d-1}{r}\frac{\mathrm{d}f}{\mathrm{d}r} = V'(f). \tag{40.33}$$

This again is the equation governing the motion of a particle in a potential $V(f)$ and submitted to a viscous damping force. The solution depends on its value at the origin $f(0)$.

We call f_+ the absolute minimum of the potential. If we choose $f(0)$ too close to f_+, $f(r) - f(0)$ will remain very small until r becomes very large. When r is large, the damping force is small so that energy is almost conserved and the particle will overshoot. If $f(0)$ is too close to zero, the particle will lose too much energy and, therefore, undershoot, the asymptotic value $f(r)$ then corresponding to the maximum of $V(f)$. Thus somewhere in between we expect to find a value $f(0)$, which corresponds to a solution which goes to zero at infinity and, therefore, has a finite action.

The virial theorem, derived in Section A40.1, implies that the corresponding action is positive:

$$\mathcal{S}(\phi_c) = A/g, \tag{40.34}$$

with

$$A = \frac{1}{d} \int (\partial_\mu f)^2 \,\mathrm{d}^d x > 0. \tag{40.35}$$

Moreover, we also derive in Section A40.1 that $\delta^2 \mathcal{S}/(\delta\phi_c)^2$ has one and only one negative eigenvalue.

40.2 General Potentials: Instanton Contributions

We now extend the analysis of Section 39.5 to the analogous scalar field theories, using the techniques developed in Section 40.1. We consider an euclidean action of the form:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left(\tfrac{1}{2} (\partial_\mu \phi)^2 + g^{-1} V(\phi\sqrt{g}) \right), \tag{40.32}$$

in which the potential $V(\phi)$ has a stable and an unstable minimum, and is of the type discussed in Section 39.5. Assuming that at some initial time the quantum mechanical state corresponds to a field concentrated around the unstable minimum of the potential, the "false" vacuum, we want to evaluate semi-classically 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 and a gaussian integration around the instanton.

40.2.1 Calculation of the instanton contribution

We define the field in such a way that the unstable minimum corresponds to $\phi = 0$. The discussion of the existence of an instanton solution is similar to the one given in Section 40.1. A theorem establishes, under mild assumptions, that spherically symmetric solutions give the minimal action. We, therefore, 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

$$\frac{\mathrm{d}^2 f}{\mathrm{d}r^2} + \frac{d-1}{r}\frac{\mathrm{d}f}{\mathrm{d}r} = V'(f). \tag{40.33}$$

This again is the equation governing the motion of a particle in a potential $V(f)$ and submitted to a viscous damping force. The solution depends on its value at the origin $f(0)$.

We call f_+ the absolute minimum of the potential. If we choose $f(0)$ too close to f_+, $f(r) - f(0)$ will remain very small until r becomes very large. When r is large, the damping force is small so that energy is almost conserved and the particle will overshoot. If $f(0)$ is too close to zero, the particle will lose too much energy and, therefore, undershoot, the asymptotic value $f(r)$ then corresponding to the maximum of $V(f)$. Thus somewhere in between we expect to find a value $f(0)$, which corresponds to a solution which goes to zero at infinity and, therefore, has a finite action.

The virial theorem, derived in Section A40.1, implies that the corresponding action is positive:

$$\mathcal{S}(\phi_c) = A/g, \tag{40.34}$$

with

$$A = \frac{1}{d}\int (\partial_\mu f)^2 \, \mathrm{d}^d x > 0. \tag{40.35}$$

Moreover, we also derive in Section A40.1 that $\delta^2 \mathcal{S}/(\delta\phi_c)^2$ has one and only one negative eigenvalue.

We now have to take x_0 as an integration variable. This leads to a jacobian J which, as we have seen in Section 40.1, can be written as

$$J = \prod_{\mu=1}^{d} \|\partial_\mu \phi_c\| = \left[\frac{1}{d}\int \mathrm{d}^d x \sum_\mu (\partial_\mu \phi_c)^2\right]^{d/2} = \left(\frac{A}{g}\right)^{d/2}, \tag{40.36}$$

where equations (40.34,40.35) have been used.

All other details of the calculation can be borrowed from the ϕ^4 example, and we finally obtain an explicit expression for the imaginary part of the n-point correlation function.

Let us just make a few remarks concerning the renormalization. (If we suppose the theory super-renormalizable or renormalizable).

Renormalization. To construct the renormalized theory we can proceed in an inductive way: we regularize the theory, we then add the counterterms which make the theory finite order by order in a loopwise expansion, that is, here an expansion in powers of g. The renormalized action $\mathcal{S}_\mathrm{r}(\phi)$ has the form:

$$\mathcal{S}_\mathrm{r}(\phi) = \frac{1}{g}\mathcal{S}_0(\phi\sqrt{g}) + \mathcal{S}_1(\phi\sqrt{g}) + \cdots + g^{L-1}\mathcal{S}_L(\phi\sqrt{g}) + \cdots .$$

At leading order only the one-loop counterterms matter in the instanton calculation. To evaluate them we expand the effective potential $\Gamma(\phi)$ (the generating functional of 1PI correlation functions) at one-loop order using the regularized action, and calculate its divergent part. In Chapter 7 we have derived:

$$\Gamma_{1\,\mathrm{loop}}(\phi) = \mathcal{S}(\phi) + \frac{1}{2}\operatorname{tr}\ln\frac{\delta^2\mathcal{S}}{\delta\phi(x)\delta\phi(x')}. \tag{40.37}$$

To render correlation functions finite we have to subtract to the regularized action the divergent part of the one loop term of $\Gamma(\phi)$:

$$\mathcal{S}_1(\phi\sqrt{g}) = -\frac{1}{2}\left(\operatorname{tr}\ln\frac{\delta^2\mathcal{S}}{\delta\phi\delta\phi}\right)_{\mathrm{div.}} .$$

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.

40.3 The ϕ^4 Field Theory in Dimension 4

The ϕ^4 field theory in four dimensions is just renormalizable, and, as we show in Appendix A40.2, 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 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.

The euclidean action of the massless theory ϕ^4 theory can be written as

$$\mathcal{S}(\phi) = \int \mathrm{d}^4 x \left[\tfrac{1}{2} \left(\partial_\mu \phi\right)^2 + \tfrac{1}{4} g \phi^4 \right], \tag{40.38}$$

and the corresponding field equation reads:

$$-\nabla^2 \phi(x) + g\phi^3(x) = 0\,. \tag{40.39}$$

Note the unconventional normalization of the coupling constant. To return to the usual convention one has to substitute $g \mapsto g/6$.

We know that the solution of minimal action is spherically symmetric, thus we set:

$$\phi(x) = \frac{1}{\sqrt{-g}} f(r)\,, \tag{40.40}$$

with

$$r = |x - x_0|\,. \tag{40.41}$$

We then obtain a differential equation

$$-\left[\left(\frac{\mathrm{d}}{\mathrm{d}r}\right)^2 + \frac{3}{r}\frac{\mathrm{d}}{\mathrm{d}r}\right] f(r) = f^3(r). \tag{40.42}$$

We now use the scale invariance of the classical theory (the theory is actually conformal invariant, see Appendix A40.4). If $\phi(x)$ is the solution to the equation, then $\psi(x)$ is also a solution with

$$\phi(x) = \lambda \psi(\lambda x). \tag{40.43}$$

This suggests the following change:

$$f(r) = \mathrm{e}^{-t}\, h(t), \qquad r = \mathrm{e}^t\,, \tag{40.44}$$

which transforms equation (40.42) into

$$\ddot{h}(t) = h(t) - h^3(t). \tag{40.45}$$

We recognize the equation of motion of the anharmonic oscillator that we have solved in Chapter 39:

$$h_c(t) = \pm \frac{\sqrt{2}}{\cosh(t - t_0)}\,. \tag{40.46}$$

The solution $\phi_c(x)$ of equation (40.39) then is

$$f(r) = \pm \frac{2\sqrt{2}\lambda}{1 + \lambda^2 r^2}\,, \tag{40.47a}$$

$$\Rightarrow \phi_c(x) = \pm \frac{1}{\sqrt{-g}} \frac{2\sqrt{2}\lambda}{1 + \lambda^2 \left(x - x_0\right)^2}\,, \tag{40.47b}$$

where we have defined $\lambda = \mathrm{e}^{-t_0}$. The corresponding classical action $\mathcal{S}(\phi_c)$ is

$$\mathcal{S}(\phi_c) = -A/g\,, \qquad A = 8\pi^2/3\,. \tag{40.48}$$

With the standard normalization of g one finds $A = 16\pi^2$.

Because the classical theory is scale invariant, the instanton solution now depends on a scale parameter λ, in addition to the four translation parameters $x_{0\mu}$. We have, therefore, to introduce five collective coordinates to calculate the instanton contribution.

40.4 Instanton Contributions at Leading Order

The general strategy. The second derivative $M(x, x')$ of the action at the saddle point is

$$M(x, x') = \frac{\delta^2 \mathcal{S}}{\delta\phi_c(x)\delta\phi_c(x')} = \left[-\nabla^2 - \frac{24\lambda^2}{(1+\lambda^2 x^2)^2}\right] \delta^{(4)}(x - x'). \tag{40.49}$$

To find the eigenvalues the operator of $\mathbf{M}$, one has to solve a 4-dimensional Schrödinger equation with a spherically symmetric potential. We immediately note at this stage two serious problems. The operator $\mathbf{M}$ has, as expected, five eigenvectors, $\partial_\mu \phi_c(x)$ and $(\mathrm{d}/\mathrm{d}\lambda)\phi_c(x)$, with eigenvalue zero, but the last of these eigenvectors is not normalizable with the natural measure of the problem,

$$\int \left[\frac{\mathrm{d}}{\mathrm{d}\lambda}\phi_c(x)\right]^2 \mathrm{d}^4 x = \infty\,. \tag{40.50}$$

This is an IR problem which arises because the theory is massless.

Another difficulty comes from the mass counterterm which has to be added to the action. It has the form:

$$\tfrac{1}{2}\delta m_0^2 \int \mathrm{d}^4 x\, \phi_c^2(x) = \infty\,. \tag{40.51}$$

The integral of $\phi_c^2(x)$ is also IR divergent, and this IR divergence is expected to cancel with an IR divergence of $\det \mathbf{M}$. Thus we need in general some kind of IR regularization. In the particular case of the dimensional regularization, this problem is postponed to 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 (40.24)). 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 indicate here only the various steps, without giving all details.

The transformation. We extend the transformation (40.44) to arbitrary fields, setting:

$$\phi(x) = \mathrm{e}^{-t}\, h(t, \hat{n}) \qquad \text{with} \quad t = \ln|x|, \qquad \hat{n}_n = \frac{x^\mu}{|x|}\,. \tag{40.52}$$

The classical action then becomes

$$\mathcal{S}(\phi) = \tilde{\mathcal{S}}(h) = \int \mathrm{d}t\, \mathrm{d}\Omega \left[\frac{1}{2}\left(\dot{h} - h\right)^2 + h\mathbf{L}^2 h + \frac{1}{4} g h^4\right]. \tag{40.53}$$

The symbol $\int \mathrm{d}\Omega$ 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 (40.53) can be rewritten

$$\tilde{\mathcal{S}}(h) = \int \mathrm{d}t\, \mathrm{d}\Omega \left\{\frac{1}{2}\left[\dot{h}^2 + h\left(\mathbf{L}^2 + 1\right) h\right] + \frac{1}{4} g h^4\right\}. \tag{40.54}$$

The integral of $\dot{h}h$ vanishes due to boundary conditions.

With the parametrization

$$\lambda = e^{-t_0}, \qquad \mathbf{x}_0 = e^{t_0}\,\mathbf{v}\,,$$

the classical solution (40.47*b*) transforms into $h_c(t)$:

$$\sqrt{-g}h_c(t) = \frac{\pm 2\sqrt{2}}{e^{(t-t_0)} - 2\mathbf{v}\cdot\mathbf{n} + e^{-(t-t_0)}(\mathbf{v}^2+1)}. \tag{40.55}$$

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 $t_0 = x_{0\mu} = 0$)

$$\mathbf{M} = \frac{\delta^2 \mathcal{S}}{\delta h_c \delta h_c} = -\left(\frac{d}{dt}\right)^2 + \mathbf{L}^2 + 1 - \frac{6}{\cosh^2 t}. \tag{40.56}$$

The natural measure associated to this hamiltonian problem is

$$\int dt\, d\Omega\,,$$

which in the original language means

$$\int \frac{d^4 x}{\mathbf{x}^2}.$$

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 $x_{0\mu} = 0$.

40.4.1 The jacobian

With the new measure $d\phi_c/d\lambda$ is normalizable:

$$J_1 = \left[\int \frac{d^4 x}{\mathbf{x}^2}\left(\frac{d}{d\lambda}\phi_c(x)\right)^2\right]^{1/2}, \tag{40.57}$$

$$= \left[\frac{16\pi^2}{(-g)}\int_0^\infty r dr \frac{\left(1-\lambda^2 r^2\right)^2}{\left(1+\lambda^2 r^2\right)^4}\right]^{1/2}. \tag{40.58}$$

This leads to a first factor:

$$J_1 = \frac{1}{\lambda}\sqrt{\frac{8}{3}}\frac{\pi}{\sqrt{-g}}\,. \tag{40.59}$$

The second jacobian J_2 comes from the collective coordinates $x_{0\mu}$:

$$J_2 = \left[\frac{1}{4}\int \frac{d^4 x}{\mathbf{x}^2}\sum_{\mu=1}^{4}\left(\partial_\mu \phi_c\right)^2\right]^2, \tag{40.60}$$

$$= \frac{1}{g^2}\left[16\pi^2 \int_0^\infty \frac{r^3 dr \lambda^6}{\left(1+\lambda^2 r^2\right)^4}\right]^2 = \frac{\lambda^4}{g^2}\times\frac{16}{9}\pi^4. \tag{40.61}$$

The complete jacobian J is thus

$$J = J_1 J_2 = \frac{\lambda^3}{(-g)^{5/2}}\pi^5 \times \frac{32\sqrt{2}}{9\sqrt{3}}\,. \tag{40.62}$$

40.4.2 The determinant

Using the result of equation (39.36) it is possible to calculate the determinant of $\mathbf{M}$ for each value l of the angular momentum:

$$M_l = -\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 + (1+l)^2 - \frac{6}{\cosh^2 t}. \tag{40.63}$$

The determinant thus is

$$\det\left(M_l+\varepsilon\right)\left(M_{0l}+\varepsilon\right)^{-1} = \frac{\sqrt{\varepsilon+(l+1)^2}-1}{\sqrt{\varepsilon+(l+1)^2}+2}\,\frac{\sqrt{\varepsilon+(l+1)^2}-2}{\sqrt{\varepsilon+(l+1)^2}+1]}, \tag{40.64}$$

in which M_{0l} is the operator of the corresponding free theory. As we know, this determinant is UV divergent and we have to renormalize it. However, let us first calculate formally the unrenormalized determinant:

$$l\ge 2: \qquad \det M_l M_{0l}^{-1} = \frac{l(l-1)}{(l+2)(l+3)}, \tag{40.65}$$

$$l=1: \qquad \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\det\left(M_1+\varepsilon\right)\left(M_{01}+\varepsilon\right)^{-1} = \frac{1}{48}, \tag{40.66}$$

$$l=0: \qquad \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\det\left(M_{l=0}+\varepsilon\right)\left(M_{0l=0}+\varepsilon\right)^{-1} = -\frac{1}{12}. \tag{40.67}$$

As expected the determinant is negative and we obtain the formal expression

$$\det{}'\,\mathbf{M}\mathbf{M}_0^{-1} = -\frac{1}{12}\times\left(\frac{1}{48}\right)^4\times\prod_{l=2}^{\infty}\left[\frac{l(l-1)}{(l+2)(l+3)}\right]^{(l+1)^2}. \tag{40.68}$$

Renormalization. In these variables, the UV divergences appear as divergences of the infinite product on l. We thus use in an intermediate step 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. We, therefore, proceed in the following way: the determinant of the operator $\mathbf{M}(s)$,

$$\mathbf{M}(s) = -\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 - \frac{s(s+1)}{\cosh^2 t}, \tag{40.69}$$

is exactly known

$$\det\left[\mathbf{M}(s)+z\right]\left[\mathbf{M}_0+z\right]^{-1} = \frac{\Gamma(1+\sqrt{z})\Gamma(\sqrt{z})}{\Gamma(1+s+\sqrt{z})\Gamma(\sqrt{z}-s)}. \tag{40.70}$$

Setting:

$$s(s+1) = 6\gamma, \tag{40.71}$$

one expands $\ln\det\mathbf{M}(s)$ in powers of γ. One deduces 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 (40.64). One then subtracts these two terms from $\ln\det\mathbf{M}$ as obtained from the representation (40.68). One then verifies that indeed the large L limit of the subtracted quantity:

$$\left\{\det{}' \mathbf{M}\mathbf{M}_0^{-1}\right\}_{\text{ren}}^{-1/2} = \lim_{L\to+\infty} i2\sqrt{3} \times (48)^2 \prod_{l=2}^{L} \left[\frac{(l+2)(l+3)}{(l-1)}\right]^{(l+1)^2/2} \prod_{l=0}^{L} \mathrm{e}^{-3(l+1)}$$
$$\times \prod_{l=0}^{L} \mathrm{e}^{-18(l+1)^2} \left[\sum_{k=l+1}^{\infty} \frac{1}{k^2} - \frac{1}{l+1} - \frac{1}{2(l+1)^2}\right], \tag{40.72}$$

is finite. We set:

$$\left\{\det{}' \mathbf{M}\mathbf{M}_0^{-1}\right\}_{\text{ren}}^{-1/2} = iC_1 \,. \tag{40.73}$$

Taking into account the jacobians, the factor $(2\pi)^{-1/2}$ for each collective mode, the factor $(2i)^{-1}$ and a factor two for the two saddle points, we get a first factor C_2 of the form

$$C_2 = \frac{\lambda^3}{(-g)^{5/2}} \times \pi^5 \times \frac{32\sqrt{2}}{9\sqrt{3}} \times \frac{C_1}{(2\pi)^{5/2}}, \tag{40.74}$$

which we write as

$$C_2 = C_3 \lambda^3/(-g)^{5/2} \,. \tag{40.75}$$

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 counterterms. 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 the next section we calculate explicitly the finite difference between this term and the coupling constant counter-term which cancels the divergence.

40.5 Coupling Constant Renormalization

The terms we want to calculate involve the renormalized four-point function. We have to choose a renormalization scheme: we assume, therefore, that we have renormalized the field theory by minimal subtraction after dimensional regularization. The renormalization constants have been calculated in Section 11.5. Notice the different normalization of the coupling constant. The contribution δS_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

$$\delta S_2 = \frac{9}{4}\frac{N_d}{\varepsilon} g^2 \int \phi_c^4(x) \mathrm{d}^4 x - \frac{9}{4} g^2 \operatorname{tr}\left[\phi_c^2 \left(-\nabla^2\right)^{-1} \phi_c^2 \left(-\nabla^2\right)^{-1}\right], \tag{40.76}$$

in which N_d is the usual loop factor:

$$N_d = 2(4\pi)^{-d/2}/\Gamma(d/2) \,. \tag{40.77}$$

and $d = 4-\varepsilon$. The expression can be rewritten as

$$\delta S_2 = -\frac{9}{4} g^2 \int \mathrm{d}^4 x \, \mathrm{d}^4 y \frac{\mathrm{d}^4 p}{(2\pi)^4} \mathrm{e}^{ip(x-y)} \phi_c^2(x)\phi_c^2(y) \lim_{d\to 4}\left(\int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{\mu^\varepsilon}{q^2(p-q)^2} - \frac{N_d}{\varepsilon}\right), \tag{40.78}$$

in which μ is the renormalization scale. The integral over $\mathbf{q}$ has been performed in Section 9.6 (see equation (9.66)):

$$\int \frac{\mathrm{d}^d q}{(2\pi)^d} \frac{1}{q^2 (p-q)^2} - \frac{N_d}{\varepsilon} = \frac{1}{8\pi^2}\left(\frac{1}{2} - \ln p\right) + O(\varepsilon) . \tag{40.79}$$

We also introduce the Fourier transform of the function $f^2(r)$ ($f(r)$ is given by equation (40.47a)):

$$v(p) = \frac{1}{(2\pi)^4} \int \mathrm{d}^4 x \frac{8\, \mathrm{e}^{ipx}}{(1+x^2)^2} . \tag{40.80}$$

The solution $\phi_c(x)$ depends on the scale λ. Rescaling the variables x, y, and p, we can then write the total expression more explicitly:

$$\delta S_2 = -\frac{9\pi^2}{2} \int \mathrm{d}^4 p\, v^2(p) \left[\tfrac{1}{2} - \ln(\lambda p/\mu)\right] . \tag{40.81}$$

From the definition of $v(p)$ we deduce after a short calculation:

$$\int \mathrm{d}^4 p\, v^2(p) = \frac{2}{(3\pi^2)} , \tag{40.82}$$

$$\int \mathrm{d}^4 p \ln p\, v^2(p) = \frac{2}{3\pi^2}\left(\ln 2 + \gamma + \frac{1}{6}\right) , \tag{40.83}$$

in which γ is Euler's constant: $\gamma \doteq -\psi(1) = 0.577215\ldots$. We then obtain

$$\delta S_2 = 3 \ln \lambda/\mu - \ln C_4 \tag{40.84}$$

with

$$\ln C_4 = 1 - 3\ln 2 - 3\gamma . \tag{40.85}$$

We note that the r.h.s. of equation (40.84) 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

$$\frac{8\pi^2}{3g} - 3 \ln \lambda/\mu = \frac{8\pi^2}{3g(\lambda)} + O(g) , \tag{40.86}$$

in which $g(\lambda)$ is the effective coupling at the scale λ, solution of the renormalization group equation,

$$\frac{\mathrm{d}g(\lambda)}{\mathrm{d}\ln\lambda} = \beta\left[g(\lambda)\right] , \tag{40.87}$$

with

$$\beta(g) = \frac{9}{8\pi^2} g^2 + O\left(g^3\right) . \tag{40.88}$$

This property is expected. The renormalization of the perturbative expansion renders the instanton contribution, before integration over dilatation, finite. Consequently this contribution should satisfy a renormalization group equation, and the coupling constant g can be present only in the combination $g(\lambda)$, since λ fixes the scale in the calculation (for more details see Appendix A40.3).

40.6 The Imaginary Part of the n-Point Function

We can now write the complete contribution to the imaginary part of the n-point function,

$$\operatorname{Im} Z^{(n)}(x_1,\ldots,x_n) \underset{g\to 0_-}{\sim} C_5 \int \mathrm{d}^4 x_0 \int_0^\infty \frac{\mathrm{d}\lambda}{\lambda}\lambda^4 \prod_{i=1}^{n} \frac{2\sqrt{2}\lambda}{1+\lambda^2 (x_i - x_0)^2} \frac{\mathrm{e}^{8\pi^2/3g(\lambda)}}{(-g)^{(n+5)/2}}, \tag{40.89}$$

where we have set:

$$C_5 = C_3 C_4 \,.$$

To calculate the Fourier transform of the expression (40.89), we introduce

$$u(p) = 2\sqrt{2}\int \mathrm{e}^{ipx}\,\frac{\mathrm{d}^4 x}{1+x^2}. \tag{40.90}$$

Then, after factorizing the δ-function of momentum conservation,

$$\operatorname{Im} \tilde{Z}^{(n)}(p_1,\ldots,p_n) \sim \frac{C_5}{(-g)^{(n+5)/2}} \int_0^\infty \mathrm{d}\lambda\, \lambda^{3-3n}\, \mathrm{e}^{8\pi^2/3g(\lambda)} \prod_{i=1}^{n} u(p_i/\lambda). \tag{40.91}$$

We can express this result in terms of 1PI correlation functions $\tilde{\Gamma}^{(n)}t(p_1,\ldots,p_n)$:

$$\operatorname{Im} \tilde{\Gamma}^{(n)}(p_1,\ldots,p_n) \sim \frac{C_5}{(-g)^{(n+5)/2}} \int_0^\infty \frac{\mathrm{d}\lambda}{\lambda}\lambda^{4-n}\, \mathrm{e}^{8\pi^2/3g(\lambda)} \prod_{i=1}^{n} (p_i^2/\lambda^2)\, u\,(p_i/\lambda). \tag{40.92}$$

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, but involves instead 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

$$\int^\infty \mathrm{d}\lambda\, \lambda^{3-n}\, \mathrm{e}^{8\pi^2/3g(\lambda)}, \tag{40.93}$$

and, therefore, we have to examine the behaviour of $g(\lambda)$ for λ large. From equation (40.88) we see that the theory is UV asymptotically free because for g negative, that is, $g(\lambda)$ goes to zero for λ large. Thus perturbation theory is applicable and we can use the approximation (40.86). The argument remains true even if we take g slightly complex. Thus the integral has the form

$$\int^\infty \mathrm{d}\lambda\, \lambda^{-n}. \tag{40.94}$$

We see 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 (40.94) 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

$\tilde{\Gamma}^{(n)}(p_1, \ldots, p_n)$ for $n \le 4$. In the ϕ^6 field theory in dimension 3, for example, these contributions cancel the divergences coming from the integral over λ. Here instead the integrals over λ are finite at this order. This implies in particular that in the minimal subtraction scheme the imaginary parts of the renormalization constants vanish at leading order. In another renormalization scheme (fixed momentum subtraction for example) 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 totally 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 $g(\lambda)$ is not too dramatic, the integrals will converge and it will be justified to replace $g(\lambda)$ by the expansion (40.86). 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 complicated.

40.7 The Massive Theory

We show in Section A40.2 that the massive field equations have instanton solution, 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, we notice a difficulty: as explained in Section 40.4 the integral of ϕ_c^2 is IR divergent. We have thus to modify the field configuration at large distances, by connecting it smoothly to the solution of the massive free equation with mass m. Qualitatively speaking we consider 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 will be met in Chapter 43 in the case of multi-instanton configuration. 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 with respect to m. The classical action has the form:

$$\mathcal{S}_m(\phi_c) = -\frac{1}{g}\left(\frac{8\pi^2}{3} + 8\pi^2 \frac{m^2}{\lambda^2} \ln \frac{\lambda}{m}\right) \quad \text{for} \quad \lambda \gg m\,, \tag{40.95}$$

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.

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 solved because the λ integral is cut at a scale $m/\sqrt{-g}$. For correlation functions the integral will be cut by the largest between momenta and $m/\sqrt{-g}$. This implies that the limits $m \to 0$ and $g \to 0$ do not commute.

$\tilde{\Gamma}^{(n)}(p_1, \ldots, p_n)$ for $n \leq 4$. In the ϕ^6 field theory in dimension 3, for example, these contributions cancel the divergences coming from the integral over λ. Here instead the integrals over λ are finite at this order. This implies in particular that in the minimal subtraction scheme the imaginary parts of the renormalization constants vanish at leading order. In another renormalization scheme (fixed momentum subtraction for example) 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 totally 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 $g(\lambda)$ is not too dramatic, the integrals will converge and it will be justified to replace $g(\lambda)$ by the expansion (40.86). 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 complicated.

40.7 The Massive Theory

We show in Section A40.2 that the massive field equations have instanton solution, 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, we notice a difficulty: as explained in Section 40.4 the integral of ϕ_c^2 is IR divergent. We have thus to modify the field configuration at large distances, by connecting it smoothly to the solution of the massive free equation with mass m. Qualitatively speaking we consider 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 will be met in Chapter 43 in the case of multi-instanton configuration. 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 with respect to m. The classical action has the form:

$$\mathcal{S}_m(\phi_c) = -\frac{1}{g}\left(\frac{8\pi^2}{3} + 8\pi^2\frac{m^2}{\lambda^2}\ln\frac{\lambda}{m}\right) \quad \text{for} \quad \lambda \gg m\,, \tag{40.95}$$

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.

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 solved because the λ integral is cut at a scale $m/\sqrt{-g}$. For correlation functions the integral will be cut by the largest between momenta and $m/\sqrt{-g}$. This implies that the limits $m \to 0$ and $g \to 0$ do not commute.

40.8 Cosmology: The Decay of the False Vacuum

In preceding sections we have in particular determined the probability for a "false vacuum" of a Field Theory to decay through barrier penetration. It has been speculated 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. 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 is the kind of physical speculation that we here have in mind.

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, absorbing eventually the whole space. To discuss what happens once a bubble has been created, it is useful to consider first the analogous problem in ordinary quantum mechanics.

Quantum mechanics. In the language of particle physics, a semi-classical description of the decay process would be the following: a particle is sitting in the well of the potential corresponding to the unstable 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.

Field theory. We apply the same ideas to the field theoretical model we discuss here. At time zero the system makes a quantum jump. According to the previous discussion, the value of the field at time zero is then (with the choice $x_{0\mu} = 0$)

$$\phi(t = 0, \mathbf{x}) = \phi_c(x_d = 0, \mathbf{x}), \qquad (\mathbf{x} = x_1, \ldots, x_{d-1}), \tag{40.96}$$

and its time derivative vanishes,

$$\partial_t \phi(t = 0, \mathbf{x}) = 0 . \tag{40.97}$$

At a later time $\phi(t, \mathbf{x})$ then obeys the *real-time* field equation,

$$\left[\nabla_i^2 - \partial_t^2\right] \phi(t, \mathbf{x}) = \frac{1}{\sqrt{g}} V'(\sqrt{g}\phi). \tag{40.98}$$

The first equation (40.96) 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 (40.96,40.97) and obviously obeys the field equation (40.98). It is, therefore, 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,

$$0 \le \mathbf{x}^2 - t^2 = r^2 \ll 1 .$$

Therefore, after a short time the bubble starts expanding at almost the speed of light.

Bibliographical Notes

The first calculations of instanton contributions in Field Theory are due to
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.
We have followed here the method presented in
E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D15 (1977) 1558.
See also
A.J. McKane and D.J. Wallace, *J. Phys. A: Math. Gen.* 11 (1978) 2285,
and the review
J. Zinn-Justin, *Phys. Rep.* 70 (1981) 109.
To simplify expressions we have used scaling arguments
G.H. Derrick, *J. Math. Phys.* 5 (1964) 1252.
The field equations for the ϕ^4 field theory have been solved numerically by Zinn-Justin unpublished, and the determinant calculated in
E. Brézin and G. Parisi, *J. Stat. Phys.* 19 (1978) 269.
The relativistic generalization of the method of collective coordinates is described in
J.-L. Gervais and B. Sakita, *Phys. Rev.* D11 (1975) 2943; V.E. Korepin, P.P. Kulish and L.D. Faddeev, *JETP Lett.* 21 (1975) 138; N.H. Christ and T.D. Lee, *Phys. Rev.* D12 (1975) 1606.
The standard early reference on vacuum decay is
M.B. Voloshin, I.Yu. Kobzarev and L.B. Okun, *Yad. Fiz.* 20 (1974) 1229 (*Sov. J. Nucl. Phys.* 20 (1975) 644).
The instanton approach to this problem has been developed in
S. Coleman, *Phys. Rev.* D15 (1977) 2929; C.G. Callan and S. Coleman, *Phys. Rev.* D16 (1977) 1762; M. Stone, *Phys. Rev.* D14 (1976) 3568, *Phys. Lett.* 67B (1977) 186; S. Coleman and F. De Luccia, *Phys. Rev.* D21 (1980) 3305; P.J. Steinhardt, *Phys. Rev.* D24 (1981) 842.
Properties of the minimal action are proven in
S. Coleman, V. Glaser and A. Martin, *Commun. Math. Phys.* 58 (1978) 211.
For a discussion of Sobolev inequalities see, for example,
E.M. Stein, *Singular Integrals and Differentiability Properties of Functions* (Princeton University Press, Princeton 1970); J.R. Klauder, *Acta Phys. Austr. Suppl.* XI (1973) 341.
The relevance of Sobolev inequalities for the instanton problem in the ϕ^4 field theory has been emphasized by
G. Parisi, *Phys. Lett.* 66B (1977) 167; 68B (1977) 361.
For a discussion of the massive theory see
Y. Frishman and S. Yankielowicz, *Phys. Rev.* D19 (1979) 540; I. Affleck, *Nucl. Phys.* B191 (1981) 429.
Renormalons have been discussed in
G. Parisi in *Cargèse Lectures 1977*, vol. B39 (Plenum, New York 1979) and *Phys. Lett.* 76B (1978) 65; B. Lautrup, *Phys. Lett.* 69B (1977) 109;
G. 't Hooft, *Erice Lectures 1977*, A. Zichichi ed. (Plenum, New York 1979); P. Olesen, *Phys. Lett.* 73B (1978) 327.
Renormalons have IR analogues in massless field theories
G. Parisi, *Nucl. Phys.* B150 (1979) 163; F. David, *Nucl. Phys.* B209 (1982) 433; B234 (1984) 237; B263 (1986) 637.

APPENDIX A40
INSTANTONS: ADDITIONAL REMARKS

We prove here a few simple equalities and inequalities concerning classical solutions that we have used in several places in the chapter, discuss the RG properties and conformal invariance of the ϕ_4^4 massless field theory.

A40.1 Virial Theorem

We consider the general action:

$$\mathcal{S}(\phi) = \int \mathrm{d}^d x \left[\tfrac{1}{2}\left(\partial_\mu \phi(x)\right)^2 + \mathcal{V}(\phi(x))\right], \qquad (A40.1)$$

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 find

$$\mathcal{S}(\phi_c, \lambda) = \lambda^{2-d} \int \tfrac{1}{2}\left(\partial_\mu \phi_c(x)\right)^2 \mathrm{d}^d x + \lambda^{-d} \int \mathcal{V}(\phi_c(x))\,\mathrm{d}^d x \,. \qquad (A40.2)$$

Since $\phi_c(x)$ satisfies the field equation the variation of the action vanishes for $\lambda = 1$:

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathcal{S}(\phi_c, \lambda)\right|_{\lambda=1} = 0 \;\Rightarrow\; (d-2)\int \tfrac{1}{2}\left(\partial_\mu \phi_c(x)\right)^2 \mathrm{d}^d x + d \int \mathcal{V}(\phi_c(x))\,\mathrm{d}^d x = 0\,. \qquad (A40.3)$$

The classical action $\mathcal{S}(\phi_c)$ can thus be expressed in terms of the kinetic term only:

$$\mathcal{S}(\phi_c) = \frac{1}{d}\int \left(\partial_\mu \phi_c(x)\right)^2 \mathrm{d}^d x > 0\,, \qquad (A40.4)$$

a form that shows that $\mathcal{S}(\phi_c)$ is always positive.

It is also interesting to calculate the second derivative of $\mathcal{S}(\phi_c, \lambda)$

$$\left.\frac{\mathrm{d}^2}{(\mathrm{d}\lambda)^2}\mathcal{S}(\phi_c, \lambda)\right|_{\lambda=1} = (2-d)\int \left[\partial_\mu \phi_c(x)\right]^2 \mathrm{d}^d x \,. \qquad (A40.5)$$

For $d \geq 2$ the solution is not a local minimum of the action, and the operator M

$$M(x, x') = \frac{\delta^2 \mathcal{S}}{\delta\phi_c(x)\delta\phi_c(x')},$$

has at least one negative eigenvalue.

Moreover, a general theorem established by Coleman *et al* states that $\phi_c(x)$ corresponds to an absolute minimum of $\mathcal{S}(\phi)$ at fixed integral of the potential $\int \mathrm{d}^d x\, \mathcal{V}(\phi_c(x))$. If 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 M has at most one negative eigenvalue. Since the equation ($A40.5$) shows that it has at least one, M has one and only one negative eigenvalue.

Special potentials. If we now consider potentials that have the special form:

$$\mathcal{V}(\phi) = \tfrac{1}{2}m^2\phi^2 + g\phi^N, \tag{A40.6}$$

we can derive an additional relation. If the action $\mathcal{S}(\phi_c)$ is finite, so is the action $\mathcal{S}(\lambda\phi_c)$:

$$\mathcal{S}(\lambda\phi_c) = \lambda^2 \int \frac{1}{2}\left[(\partial_\mu\phi_c(x))^2 + m^2\phi_c^2(x)\right] \mathrm{d}^d x + \lambda^N \int \phi_c^N(x)\mathrm{d}^d x \,.$$

Again the derivative for $\lambda = 1$ vanishes if ϕ_c is a solution. We now have two relations and we can express the action in terms of the mass term. After some simple algebra one obtains

$$\mathcal{S}(\phi_c) = \frac{N-2}{2d - N(d-2)} m^2 \int \phi_c^2(x)\mathrm{d}^d x \,. \tag{A40.7}$$

In particular these results are consistent only if the expression is positive, which implies

$$N \le 2d/(d-2),$$

and, therefore, the field theory must be renormalizable.

A40.2 Sobolev Inequalities

We consider the following functional $R(\varphi)$,

$$R(\varphi) = \frac{\left\{\int \mathrm{d}^d x \left[(\partial_\mu\varphi(x))^2 + \varphi^2(x)\right]\right\}^2}{\int \varphi^4(x)\mathrm{d}^d x}. \tag{A40.8}$$

If the dimension d is not larger than 4, Sobolev inequalities tell us that

$$R(\varphi) \ge R > 0 \,. \tag{A40.9}$$

In addition, for $d < 4$, there exists a spherically symmetric, zero-free function $\varphi_c(x)$ which saturates the bound

$$R(\varphi_c) = R \,, \tag{A40.10}$$

and is a solution of the variational equation

$$\delta R / \delta\varphi_c(x) = 0 \,. \tag{A40.11}$$

Dimensional smaller than four. The equation (A40.11) has the explicit form,

$$\left(-\nabla^2 + 1\right)\varphi(x) - \varphi^3(x) K = 0 \,, \tag{A40.12}$$

in which we have defined

$$K = \int \mathrm{d}^d x \left[(\partial_\mu\varphi_c)^2 + \varphi_c^2\right] \Big/ \int \varphi_c^4(x)\mathrm{d}^d x \,. \tag{A40.13}$$

This equation is, up to a rescaling of $\varphi_c(x)$, the equation of motion (40.6). Both equations become identical if we choose the scale of $\varphi_c(x)$ which is otherwise arbitrary such that

$$K = 1 \ \Rightarrow f(x) = \varphi_c(x) \,. \tag{A40.14}$$

For each instanton solution we have derived the identities (40.12). Combining them with $K=1$, we obtain

$$A=\frac{3}{2}\int \mathrm{d}^d x\, f^4(x)=\frac{3}{2}R(\varphi_c)\,. \tag{A40.15}$$

The smallest action solution thus corresponds to the minimum of $R(\varphi)$:

$$A=3R/2\,, \tag{A40.16}$$

and the solution $f(x)$ we are looking for is given by

$$f(x)=\varphi_c(x) \quad \text{for} \quad K=1\,. \tag{A40.17}$$

The introduction of the functional $R(\varphi)$ has the following advantage: the action (40.1) is obviously not bounded from below. But if we restrict ourselves to fields $\phi(x)$ 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. In the operator sense:

$$\frac{\delta^2 R}{\delta\varphi_c(x)\delta\varphi_c(x')}\geq 0\,. \tag{A40.18}$$

Calculating the second derivative of R explicitly, we obtain

$$\frac{\delta^2 R}{\delta\varphi_c(x)\delta\varphi_c(x')}=4\left\{\left[-\nabla^2+1-3\varphi_c^2(x)\right]\delta(x-x')+\frac{2\varphi_c^3(x)\varphi_c^3(x')}{\int\varphi_c^4(y)\mathrm{d}^d y}\right\}. \tag{A40.19}$$

We have again set $K=1$. Let us write below $M(x,x')$ in terms of $f(x)$ or $\varphi_c(x)$ for $m=1$,

$$M(x,x')=\left[-\nabla^2+1-3\varphi_c^2(x)\right]\delta(x-x'). \tag{A40.20}$$

We, therefore, derive the relation:

$$M(x',x)=\frac{1}{4}\frac{\delta^2 R}{\delta\varphi_c(x)\delta\varphi_c(x')}-2\left(\frac{\varphi_c^3(x')\varphi_c^3(x)}{\int\varphi_c^4(y)\mathrm{d}^d y}\right). \tag{A40.21}$$

(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 zero, thus

$$\int\varphi_c(x')\varphi_c(x)M(x',x)\mathrm{d}^d x\,\mathrm{d}^d x'=-2\int\varphi_c^4(x)<0\,. \tag{A40.22}$$

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, we 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 $\delta^2 R/(\delta\varphi_c)^2$ is not a positive operator.

We conclude that $\mathbf{M}$ has *one and only one* negative eigenvalue.

Dimension four. Let us calculate $R[\varphi(\lambda x)]$ for $d \leq 4$. Then, changing λx in x' in the various integrals, we obtain

$$R\left[\varphi(\lambda x)\right] = \frac{\left\{\int \mathrm{d}^d x \left[\lambda^{2-d}\left(\partial_\mu\varphi\right)^2 + \lambda^{-d}\varphi^2\right]\right\}^2}{\lambda^{-d}\int \varphi^4 \mathrm{d}^d x}. \tag{A40.23}$$

We can now write

$$R = \min_{\{\varphi(x)\}} \min_{\lambda} R\left[\varphi(\lambda x)\right]. \tag{A40.24}$$

The minimum in λ of expression (A40.23) is obtained for

$$\lambda = \left[\frac{d}{(4-d)}\frac{\int \varphi^2(x)\mathrm{d}^d x}{\int (\partial_\mu\varphi)^2 \mathrm{d}^d x}\right]^{1/2}, \tag{A40.25}$$

and equation (A40.24) becomes

$$R = \min_{\{\varphi(x)\}} \frac{16}{d^{d/2}(4-d)^{(4-d)/2}} \frac{\left(\int \left(\partial_\mu\varphi\right)^2 \mathrm{d}^d x\right)^{d/2}\left(\int \varphi^2(x)\mathrm{d}^d x\right)^{2-d/2}}{\int \varphi^4 \mathrm{d}^d x}. \tag{A40.26}$$

For $d = 4$, we see that the solution is $\lambda = \infty$ and expression (A40.26) is just the equivalent of expression (A40.8) for the massless ϕ^4 field theory. 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 will apply to the ϕ^4 action.

The minimum of $R(\varphi)$ will be 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_4^4 field theory in four dimensions that we have discussed in Section 40.3.

A40.3 Instantons and RG Equations

We now briefly describe a few RG properties of the instanton contributions in the ϕ_4^4 field theory discussed starting with Section 40.3.

The instanton contribution to the 1PI n-point function can be written as

$$\operatorname{Im}\Gamma^{(n)}(p_i;\mu,g) = \int_0^\infty \frac{\mathrm{d}\lambda}{\lambda} F^{(n)}(p_i;\mu,g,\lambda), \tag{A40.27}$$

in which μ represents the subtraction scale, and λ the dilatation parameter. The counterterms which renormalize the perturbative expansion, also render $F^{(n)}$ finite for reasons we have already explained. Therefore, $F^{(n)}$ satisfies an RG equation:

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right] F^{(n)}(p_i;\mu,g,\lambda) = 0\,, \tag{A40.28}$$

Dimension four. Let us calculate $R[\varphi(\lambda x)]$ for $d \le 4$. Then, changing λx in x' in the various integrals, we obtain

$$R\left[\varphi(\lambda x)\right] = \frac{\left\{\int \mathrm{d}^d x \left[\lambda^{2-d}\left(\partial_\mu \varphi\right)^2 + \lambda^{-d}\varphi^2\right]\right\}^2}{\lambda^{-d}\int \varphi^4 \mathrm{d}^d x}. \tag{A40.23}$$

We can now write

$$R = \min_{\{\varphi(x)\}} \min_{\lambda} R\left[\varphi(\lambda x)\right]. \tag{A40.24}$$

The minimum in λ of expression ($A40.23$) is obtained for

$$\lambda = \left[\frac{d}{(4-d)}\frac{\int \varphi^2(x)\mathrm{d}^d x}{\int(\partial_\mu\varphi)^2\mathrm{d}^d x}\right]^{1/2}, \tag{A40.25}$$

and equation ($A40.24$) becomes

$$R = \min_{\{\varphi(x)\}} \frac{16}{d^{d/2}(4-d)^{(4-d)/2}} \frac{\left(\int\left(\partial_\mu\varphi\right)^2 \mathrm{d}^d x\right)^{d/2} \left(\int \varphi^2(x)\mathrm{d}^d x\right)^{2-d/2}}{\int \varphi^4 \mathrm{d}^d x}. \tag{A40.26}$$

For $d = 4$, we see that the solution is $\lambda = \infty$ and expression ($A40.26$) is just the equivalent of expression ($A40.8$) for the massless ϕ^4 field theory. 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 will apply to the ϕ^4 action.

The minimum of $R(\varphi)$ will be 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_4 field theory in four dimensions that we have discussed in Section 40.3.

A40.3 Instantons and RG Equations

We now briefly describe a few RG properties of the instanton contributions in the ϕ^4_4 field theory discussed starting with Section 40.3.

The instanton contribution to the 1PI n-point function can be written as

$$\operatorname{Im}\Gamma^{(n)}(p_i;\mu,g) = \int_0^\infty \frac{\mathrm{d}\lambda}{\lambda} F^{(n)}(p_i;\mu,g,\lambda), \tag{A40.27}$$

in which μ represents the subtraction scale, and λ the dilatation parameter. The counterterms which renormalize the perturbative expansion, also render $F^{(n)}$ finite for reasons we have already explained. Therefore, $F^{(n)}$ satisfies an RG equation:

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right] F^{(n)}(p_i;\mu,g,\lambda) = 0, \tag{A40.28}$$

which, integrated by the method of characteristics, yields

$$F^{(n)}(p_i;\mu;g;\lambda) = Z^{-n/2}(\tau)F^{(n)}(p_i;\mu\tau,g(\tau),\lambda), \tag{A40.29}$$

with the definitions:

$$\begin{aligned} \ln\tau &= \int_g^{g(\tau)} \frac{dg'}{\beta(g')}, \\ \ln Z &= \int_g^{g(\tau)} \frac{\eta(g')}{\beta(g')}. \end{aligned} \tag{A40.30}$$

The coupling constant $g(\tau)$ is the effective coupling constant at the scale τ.
Ordinary dimensional considerations now tell us that

$$F^{(n)}(p_i;\mu;g;\lambda) = \lambda^{4-n}F^{(n)}(p_i/\lambda,\mu/\lambda,g,1). \tag{A40.31}$$

Applied to the r.h.s. of equation (A40.29), this identity yields

$$F^{(n)}(p_i;\mu;g;\lambda) = Z^{-n/2}(\tau)\lambda^{4-n}F^{(n)}(\frac{p_i}{\lambda},\frac{\mu\tau}{\lambda},g(\tau)). \tag{A40.32}$$

The choice

$$\tau = \lambda/\mu\,,$$

finally leads to the relation:

$$F^{(n)}(p_i;\mu;g;\lambda) = [Z(\lambda/\mu)]^{-n/2}\,\lambda^{4-n}F^{(n)}\,[p_i/\lambda;g(\lambda/\mu)]\,. \tag{A40.33}$$

A40.4 Conformal Invariance

The scale invariance of the classical ϕ_4^4 field theory has allowed us to obtain in Section 40.3 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 39, and calculate explicitly the instanton contribution at leading order. Actually the scale invariant classical ϕ_4^4 theory is also conformal invariant (see Appendix A13.5). This property, which also holds for other scale invariant field theories like gauge theories, can be used more directly to calculate the instanton contribution. The conformal group is isomorphic to $SO(5,1)$. It is expected that the minimal action solution will be invariant under a maximal compact subgroup of $SO(5,1)$, that is, $SO(5)$. It is then convenient to perform a stereographic mapping of $\mathbb{R}^4$ onto the sphere S_4 to simplify the $SO(5)$ transformations. One sets:

$$\xi^\mu = \frac{2x^\mu}{1+\mathbf{x}^2}, \qquad \xi^5 = \frac{1-\mathbf{x}^2}{1+\mathbf{x}^2} \quad \Rightarrow \quad \sum_{a=1}^{5}\xi^a\xi^a = 1\,. \tag{A40.34}$$

It is also useful to introduce a field which has simple transformation properties under $SO(5)$. In the ϕ_4^4 theory, the conformal transformation properties of the ϕ-field lead to set:

$$\phi = \frac{1}{1+\mathbf{x}^2}\psi\,. \tag{A40.35}$$

To express the classical action (40.38) in terms of these new variables we perform the transformations in two steps: first we keep the variables x^μ, but now considered as coordinates on S_4, and perform only the substitution ($A40.35$). The metric $g_{\mu\nu}$ on S_4 in the coordinates x^μ is

$$g_{\mu\nu} = 4\frac{\delta_{\mu\nu}}{(1+\mathbf{x}^2)^2}. \qquad (A40.36)$$

The invariant measure on the sphere involves the square root of the determinant of the metric $\mathbf{g}$ (see Section 22.5):

$$\sqrt{\det \mathbf{g}} = \frac{16}{(1+\mathbf{x}^2)^4}. \qquad (A40.37)$$

Finally, after an integration by parts the kinetic term can be rewritten as

$$\int \mathrm{d}^4x\, (\partial_\mu \phi)^2 = \int \mathrm{d}^4x \left[\frac{(\partial_\mu \psi)^2}{(1+\mathbf{x}^2)^2} + \frac{8\psi^2}{(1+\mathbf{x}^2)^4}\right]. \qquad (A40.38)$$

The classical action then reads:

$$\mathcal{S}(\psi) = \int \mathrm{d}^4x \sqrt{\det \mathbf{g}} \left(\frac{1}{8} g^{\mu\nu}\partial_\mu\psi\partial_\nu\psi + \frac{1}{4}\psi^2 + \frac{g}{64}\psi^4\right). \qquad (A40.39)$$

In this covariant form the change of coordinates ($A40.34$) is straightforward and hardly necessary. One solution of minimal action is a constant:

$$\psi^2 = -1/8g. \qquad (A40.40)$$

The classical action is proportional to the surface of S_4 which is $8\pi^2/3$. The operator $\mathbf{M}$ second derivative of the action is given by (see Section 3.4)

$$\mathbf{M} = \tfrac{1}{4}\mathbf{L}^2 - 1, \qquad (A40.41)$$

in which $\mathbf{L}$ is the angular momentum in 5 space dimensions. The eigenvalues of $\mathbf{L}^2$ are $l(l+3)$ with a degeneracy δ_l:

$$\delta_l = \frac{1}{6}\frac{(2l+3)\Gamma(l+3)}{\Gamma(l+1)}. \qquad (A40.42)$$

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 40.4. We leave up to the reader, as an exercise, to verify other results.

41 DEGENERATE CLASSICAL MINIMA AND INSTANTONS

We consider in this chapter a somewhat different situation in which instantons play a role: classical actions with degenerate isolated minima. The simplest examples are provided by simple quantum mechanics when potentials have degenerate minima. Classically the state of minimum energy corresponds to a particle sitting at any of the minima of the potential. The position of the particle breaks (spontaneously) the symmetry of the system. In quantum mechanics on the contrary, we expect the modulus of the ground state wave function to be large near all minima of the potential, as a consequence of barrier penetration effects. We discuss in this chapter this phenomenon on two typical examples: the double-well potential and the periodic cosine potential.

In Section 41.3 we relate instantons to Arrhenius law in the context of stochastic dynamics. The proof of existence of instantons relies on an inequality related to supersymmetric structures, and which generalizes to the following field theory examples.

In field theory the problem is more subtle as the study of phase transitions has shown. However, the presence of instantons again indicates that the classical minima are connected and that the symmetry between them is not spontaneously broken. Examples of instantons of this type are provided in two dimensions by the $CP(N-1)$ models and in four dimensions by $SU(2)$ gauge theories.

41.1 The Double-Well Potential

We first discuss the simple example of the double-well potential:

$$H = -\tfrac{1}{2}\left(\mathrm{d}/\mathrm{d}x\right)^2 + \tfrac{1}{2}x^2\left(1 - x\sqrt{g}\right)^2 . \tag{41.1}$$

The potential has two degenerate minima located at the origin and at $1\big/\sqrt{g}$. In addition the hamiltonian commutes with the reflection operator P which exchanges x and $g^{-1/2} - x$. This last property is not essential for the existence of instanton solutions. It is just a simplifying feature, which moreover is present in several examples of physical interest.

The structure of the ground state. Due to the symmetry of the potential, we can expand around each of the minima of the potential, and we find the same perturbative expansion to all orders. It would, therefore, seem that the quantum hamiltonian also has a doubly degenerate ground state, corresponding to two eigenfunctions respectively concentrated around each of the classical minima of the potential. Actually we know that, due to barrier penetration, this is not the case and the true eigenstates are also eigenstates of the reflection operator P, the ground state being an even state. The reflection symmetry is not spontaneously broken in quantum mechanics in the case of regular potentials: as we have extensively discussed, 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 1D Ising model (see Section 23.1). Nevertheless, if we use the quantity $\operatorname{tr} \mathrm{e}^{-\beta H}$ to calculate the ground state energy, since we expand in g small first, we find it difficult to separate the two first eigenstates in the large β limit:

$$\begin{aligned} &\operatorname{tr} \mathrm{e}^{-\beta H} \sim \mathrm{e}^{-\beta H_+} + \mathrm{e}^{-\beta H_-} \sim 2\,\mathrm{e}^{-\beta(E_+ + E_-)/2} \cosh[\beta(E_+ - E_-)/2] \\ &\quad \text{for } g \to 0\,, \quad \beta \to \infty\,. \end{aligned} \tag{41.2}$$

The partition function is dominated by the perturbative expansion of 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 $(E_+ - E_-)^2$:

$$-\frac{1}{\beta}\ln \operatorname{tr} \mathrm{e}^{-\beta H} = \frac{1}{2}(E_+ + E_-) - \frac{1}{\beta}\ln 2 + O\left[\mathrm{e}^{-\beta}, \beta(E_+ - E_-)^2\right]$$
$$\text{for } g \to 0, \quad \beta \to \infty . \tag{41.3}$$

To calculate the difference $(E_+ - E_-)$, it is, therefore, much easier to evaluate the quantity $\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H}$ (see Section 23.2). In the same limits we now find:

$$\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H} \sim \mathrm{e}^{-\beta E_+} - \mathrm{e}^{-\beta E_-} \sim -2\sinh[\beta(E_+ - E_-)/2]\, \mathrm{e}^{-\beta(E_+ + E_-)/2} \tag{41.4}$$
$$g \to 0, \quad \beta \to \infty .$$

Since $E_+ - E_-$ vanishes in perturbation theory, the r.h.s. is dominated by

$$\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H} \sim -\beta\, \mathrm{e}^{-\beta/2}(E_+ - E_-)\left[1 + O\left(g, \mathrm{e}^{-\beta}\right)\right]. \tag{41.5}$$

It is actually more convenient to consider the ratio between the quantities (41.2) and (41.4):

$$\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H} / \operatorname{tr} \mathrm{e}^{-\beta H} \sim -\tfrac{1}{2}\beta(E_+ - E_-)\left[1 + O\left(\mathrm{e}^{-\beta}, (E_+ - E_-)^2\right)\right]. \tag{41.6}$$

By evaluating this ratio, we can 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 in perturbation theory, we have to look for non-perturbative effects: they are here due to instantons.

Instanton contributions. The path integral representation of $\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H}$ differs from the representation of the partition function only by boundary conditions:

$$\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H} = \int_{q(-\beta/2)+q(\beta/2)=g^{-1/2}} [\mathrm{d}q(t)] \exp\left[-\mathcal{S}(q)\right], \tag{41.7}$$

with

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}^2(t) + \tfrac{1}{2}q^2(t)\left(1 - \sqrt{g}\,q(t)\right)^2\right] \mathrm{d}t . \tag{41.8}$$

The path integral representation of $\operatorname{tr} \mathrm{e}^{-\beta H}$ is dominated by the trivial saddle points:

$$q(t) = 0 \quad \text{or} \quad q(t) = g^{-1/2},$$

and this leads to the usual perturbative expansion. However, these paths do not contribute to the path integral (41.7) because they do not satisfy the boundary conditions. This is not too surprising since we have already seen that the difference $E_+ - E_-$ vanishes to all orders in an expansion in powers of g. We, therefore, have to look for non-trivial solutions of the equation of motion which have a finite action in the infinite β limit. The boundary conditions then impose

$$q(\mp\infty) = 0 \qquad \text{and} \qquad q(\pm\infty) = g^{-1/2} .$$

The non-degeneracy of the ground state thus depends on the existence of an instanton solution connecting the two minima of the potential.

The euclidean equation of motion is

$$-\ddot{q}(t) + q(t)\left(1 - q(t)\sqrt{g}\right)\left(1 - 2q(t)\sqrt{g}\right) = 0 . \tag{41.9}$$

In the infinite β limit, this equation indeed has two solutions with finite classical action which we call generically instantons or instanton and anti-instanton when it is necessary to distinguish between them:

$$q_c^{\pm}(t) = \frac{1}{\sqrt{g}} \frac{1}{\left(1 + \mathrm{e}^{\mp(t-t_0)}\right)} . \tag{41.10}$$

The methods of Section 39.6 can then be adapted to the present problem. From equations analogous to (39.92,39.93), we obtain the expansions of the classical energy and action for β large:

$$E(\beta) = -2\,\mathrm{e}^{-\beta} + O\left(\mathrm{e}^{-2\beta}\right), \tag{41.11}$$

$$\mathcal{S}(q_c) = \frac{1}{g}\left[\frac{1}{6} - 2\,\mathrm{e}^{-\beta} + O\left(\mathrm{e}^{-2\beta}\right)\right]. \tag{41.12}$$

The determinant resulting from the integration around the saddle point can also be evaluated by the method explained in Section 39.5. The only noticeable modification comes from the fact that $\dot{q}_c(t)$ has no zero. It thus corresponds to the ground state of the hamiltonian $\partial^2 \mathcal{S}/(\partial q_c)^2$, which is, therefore, a positive operator. The final result is real, as expected. We can use the expression (39.91) to obtain it, except that no $1/2i$ factor is needed here, but instead one has to multiply by a factor 2 since the two solutions q_c^+ and q_c^- give identical contributions:

$$\operatorname{tr} \mathrm{P}\, \mathrm{e}^{-\beta H} \sim \frac{2}{\sqrt{\pi g}} \beta\, \mathrm{e}^{-\beta/2}\, \mathrm{e}^{-1/6g} \left(1 + O\left(g\right)\right), \quad g \to 0, \quad \beta \to \infty . \tag{41.13}$$

Using then equation (41.6) we find the asymptotic behaviour of $E_+ - E_-$ for g small,

$$E_+ - E_- \underset{g \to 0}{=} -\frac{2}{\sqrt{\pi g}}\, \mathrm{e}^{-1/6g} \left(1 + O\left(g\right)\right). \tag{41.14}$$

The difference is exponentially small in $1/g$ and this is consistent with the property that it vanishes to all orders in g.

41.2 The Periodic Cosine Potential

We now consider a slightly more complicated problem:

$$H = -\tfrac{1}{2}\left(\mathrm{d}/\mathrm{d}x\right)^2 + g^{-1}\left(1 - \cos x\sqrt{g}\right). \tag{41.15}$$

The potential is periodic and has, therefore, an infinite number of degenerate minima. We can expand perturbation series around each of these minima and, therefore, to all orders in perturbation theory, the quantum hamiltonian also has 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.

41.2.1 The structure of the ground state

More precisely, let us introduce the translation operator T, corresponding to an elementary translation of one period $2\pi/\sqrt{g}$. Since it commutes with the hamiltonian,

$$[T, H] = 0\,, \tag{41.16}$$

both operators can be diagonalized simultaneously. Because the eigenfunctions of H must be bounded at infinity, the eigenvalues of T must be pure phases. Each eigenfunction of H is thus characterized by an angle φ (pseudo-momentum) eigenvalue of T:

$$T\left|\varphi\right\rangle = \mathrm{e}^{i\varphi}\left|\varphi\right\rangle. \tag{41.17}$$

In the limit $g \to 0$, H has the spectrum of the harmonic oscillator, each state being infinitely degenerate. In the basis in which T is diagonal, all the eigenstates corresponding to the same eigenvalue are indexed by the angle φ. When $g \neq 0$, each point of the spectrum becomes a band. In a band the energy eigenvalue is a periodic function of φ which can be expanded in a Fourier series:

$$E(\varphi) = \sum_{-\infty}^{+\infty} E_l\, \mathrm{e}^{il\varphi}, \quad E_l = E_{-l}\,. \tag{41.18}$$

All coefficients E_l except E_0 vanish to all orders in a perturbative expansion in g.

We now consider the partition function which is here $\mathrm{tr}'\, T\,\mathrm{e}^{-\beta H}$. The notation tr' has the following meaning: since the diagonal matrix elements of $\mathrm{e}^{-\beta H}$ in configuration space are periodic functions, we integrate only over one period.

The large β limit selects the lowest band and we obtain (see Appendix A41.1):

$$\mathrm{tr}'\,\mathrm{e}^{-\beta H} \underset{\beta\to\infty}{\sim} \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\varphi\, \mathrm{e}^{-\beta E(\varphi)}\,. \tag{41.19}$$

Like in the case of the double-well potential, we observe that it is difficult to determine the dependence on φ of the energy levels from the partition function. Let us instead consider:

$$\mathrm{tr}'\,T\,\mathrm{e}^{-\beta H} \underset{\beta\to\infty}{\sim} \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\varphi\, \mathrm{e}^{i\varphi}\,\mathrm{e}^{-\beta E(\varphi)}\,. \tag{41.20}$$

For g small $E(g) - E_0(g)$ vanishes faster than any power of g. We can therefore expand equation (41.20):

$$\mathrm{tr}'\; T\;\mathrm{e}^{-\beta H} \sim \mathrm{e}^{-\beta E_0}\int \frac{\mathrm{d}\varphi}{2\pi}\,\mathrm{e}^{i\varphi}\left[1 - \beta(E - E_0) + \cdots\right] \quad \text{for } g \to 0\,, \quad \beta \to \infty\,. \tag{41.21}$$

The integration over φ selects E_1:

$$\mathrm{tr}'\,T\,\mathrm{e}^{-\beta H} \sim -\beta\,\mathrm{e}^{-\beta E_0}\,E_1(g), \quad g \to 0\,, \quad \beta \to \infty\,. \tag{41.22}$$

This equation can be more conveniently rewritten as

$$\mathrm{tr}'\,T\,\mathrm{e}^{-\beta H} \,/\, \mathrm{tr}'\,\mathrm{e}^{-\beta H} \sim -\beta E_1(g). \tag{41.23}$$

As indicated above, if E_1 does not vanish this implies that the translation symmetry is not spontaneously broken.

Remark. To evaluate the other Fourier series coefficients $E_2, E_3, \ldots$, for g small, the most convenient method is to consider $\mathrm{tr}'\,T^n\,\mathrm{e}^{-\beta H}$ for $n = 2, 3, \ldots$. This evaluation leads to additional problems that we solve in Chapter 43.

41.2.2 The instanton contributions

The path integral representations of the partition function $\mathrm{tr}'\,\mathrm{e}^{-\beta H}$ and of $\mathrm{tr}'\,T\,\mathrm{e}^{-\beta H}$ again differ only by the boundary conditions. The operator T has the effect of translating the argument q in the matrix element $\langle q'|\,\mathrm{tr}'\,\mathrm{e}^{-\beta H}\,|q\rangle$ before taking the trace:

$$\mathrm{tr}'\,T\,\mathrm{e}^{-\beta H} = \int_{q(\beta/2)=q(-\beta/2)+2\pi/\sqrt{g}} [\mathrm{d}q(t)]\exp\left[-\mathcal{S}(q)\right], \tag{41.24}$$

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}^2(t) + g^{-1}\left(1-\cos\sqrt{g}q(t)\right)\right]\mathrm{d}t\,. \tag{41.25}$$

We recall that $q(-\beta/2)$ varies over only 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. Solving the equation of motion explicitly, we find

$$q_c(t) = \frac{4}{\sqrt{g}}\tan^{-1}\mathrm{e}^{(t-t_0)}, \tag{41.26}$$

and the corresponding classical action, in the infinite β limit, is

$$\mathcal{S}(q_c) = 8/g\,. \tag{41.27}$$

For all potentials for which the minima can be exchanged by a reflection, the analogue of expression (39.100) is

$$E(\beta) \sim -\,\mathrm{e}^{-\beta}\,\frac{x_0^2}{2}\exp\left[2\int_0^{x_0/2}\left(\frac{1}{\sqrt{2V(x)}}-\frac{1}{x}\right)\mathrm{d}x\right], \tag{41.28}$$

in which x_0 is the location of the other minimum. Applying equation (41.28) to the analogue of equation (39.91), we obtain

$$-\beta\,\mathrm{e}^{-\beta/2}\,E_1(g) \underset{g\to 0}{\sim} \frac{4\beta\,\mathrm{e}^{-\beta/2}}{\sqrt{\pi g}}\,\mathrm{e}^{-8/g}, \tag{41.29}$$

or

$$E_1(g) \underset{g\to 0}{\sim} -\frac{4}{\sqrt{\pi g}}\,\mathrm{e}^{-8/g}\,. \tag{41.30}$$

Without evaluating E_n for $n \ge 2$ explicitly, one verifies that the corresponding boundary conditions for $\mathrm{tr}\,T^n\,\mathrm{e}^{-\beta H}$ which are

$$q(\beta/2) = q(-\beta/2) + n\frac{2\pi}{\sqrt{g}},$$

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

$$E_\varphi(g) = E_0(g) - \frac{8}{\sqrt{\pi g}}\,\mathrm{e}^{-8/g}\left[1+O(g)\right]\cos\varphi + O\left(\mathrm{e}^{-16/g}\right). \tag{41.31}$$

General discussion. We have illustrated with two examples 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. An analogous conclusion has been reached on the lattice in Sections 23.1,23.2 where we have also shown that, in contrast, spontaneous symmetry breaking of discrete symmetries is possible in higher dimensions.

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 Chapter 31).

While in theories in which the dynamical variables live in flat euclidean space, instantons are always 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 x 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 3.3) or a one-dimensional classical spin chain in a magnetic field. The classical minimum is no longer degenerate because all minima are identified to one point on the circle. The quantum ground state is equally unique since the Hilbert space consists in strictly periodic eigenfunctions ($\varphi = 0$). Still instanton solutions exist but they 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 to the perturbative expansion that we have described before.

Finally, further insight into the problem can be gained by generalizing the hamiltonian to the $O(N)$ rotator of Section 3.4 in a potential $1 - x_1$. 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 42).

41.3 Instantons and Stochastic Dynamics

We now discuss 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 have been considered already in Chapter 40, 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.

We recall (see Chapter 4) that to the Langevin equation (4.38),

$$\dot{\mathbf{q}} = -\tfrac{1}{2}\nabla E(\mathbf{q}) + \boldsymbol{\nu}(t)\,, \tag{41.32}$$

with gaussian white noise (4.2)

$$[\mathrm{d}\rho(\boldsymbol{\nu})] = [\mathrm{d}\boldsymbol{\nu}] \exp\left[-\frac{1}{2}\int \mathrm{d}t\, \boldsymbol{\nu}^2(t)/\Omega\right],$$

is associated the *hermitian* hamiltonian H (4.41):

$$H = \tfrac{1}{2}\left[-\Omega\nabla^2 + \tfrac{1}{4}\Omega^{-1}\left(\nabla E(\hat{\mathbf{q}})\right)^2 - \tfrac{1}{2}\nabla^2 E(\hat{\mathbf{q}})\right], \tag{41.33}$$

and the dynamic action:

$$\mathcal{A}(\mathbf{q}) = \tfrac{1}{2}\Omega^{-1}\int \left[\dot{\mathbf{q}}^2 + \tfrac{1}{4}\left(\nabla E\right)^2 - \tfrac{1}{2}\Omega\nabla^2 E\right]\mathrm{d}t\,. \tag{41.34}$$

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 $\Omega\nabla^2 E$ can be omitted. Therefore, the classical minima of the action (41.34) 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 $\mathbf{q} = 0$:

$$E(\mathbf{q}) = \tfrac{1}{2}\omega^2\mathbf{q}^2 + O\left(|\mathbf{q}|^3\right).$$

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 the thermal fluctuations described by the Langevin equation. Our 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 $\mathrm{e}^{-E(\mathbf{q})/\Omega}$ is not normalizable, like in one dimension

$$E(q) = q^2 - 2q^3/3\,. \tag{41.35}$$

We then know that τ is the inverse of the smallest eigenvalue of the hamiltonian (41.33) (see Section 4.4) 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 $\mathrm{e}^{-E(\mathbf{q})/(2\Omega)}$ is the formal ground state eigenvector associated with the eigenvalue 0. It follows that the calculation of the eigenvalue is non-perturbative. We now show that the instantons connecting the critical points of the function $E(\mathbf{q})$ provide a solution to the problem.

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

$$\int_{-\infty}^{+\infty}\mathrm{d}t\left(\dot{\mathbf{q}}_c \pm \tfrac{1}{2}\nabla E(\mathbf{q}_c)\right)^2 \geq 0\,, \tag{41.36}$$

and therefore,

$$2\Omega\mathcal{A}(\mathbf{q}_c) \geq |Q(\mathbf{q}_c)|\,, \tag{41.37}$$

with

$$Q(\mathbf{q}_c) = \int_{-\infty}^{+\infty}\mathrm{d}t\,\dot{\mathbf{q}}_c\cdot\nabla E(\mathbf{q}_c) = E(\mathbf{q}_0) - E(0). \tag{41.38}$$

We conclude that the action satisfies

$$\mathcal{A}(\mathbf{q}_c) \geq \tfrac{1}{2}\Omega^{-1}\left|E(\mathbf{q}_0) - E(0)\right|. \tag{41.39}$$

The equality corresponds to a local minimum of the action and $\mathbf{q}_c$ then is a solution of a first-order equation

$$\dot{\mathbf{q}} = \pm\nabla E(\mathbf{q})/2\,. \tag{41.40}$$

However, this is not the end of the story. Indeed, in Section 4.2 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 consider only 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)

$$\mathcal{A}(\mathbf{q}_c) = \Omega^{-1}\big(E(\mathbf{q}_0) - E(0)\big). \tag{41.41}$$

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 $\mathrm{e}^{-\Delta E/\Omega}$ in which ΔE is the variation of the function E:

$$\Delta E = E_{\max} - E_{\min}\,.$$

The time $\tau = O(\mathrm{e}^{\Delta E/\Omega})$ characterizes the exponential decay of the probability of finding $\mathbf{q}(t)$ near the origin when the initial conditions at $t = 0$ are $\mathbf{q}(t = 0) = 0$.

Note finally that in order to complete the calculation of the eigenvalue it is necessary to use multi-instanton techniques of the kind explained in Chapter 43.

Quantum field theory. We now consider a dynamics governed by a purely dissipative Langevin equation, which formally converges towards an equilibrium distribution corresponding to the path or the functional integral of a d dimensional euclidean quantum field theory (see also Section A37.1.3). The dynamic action then is $d + 1$ dimensional. We again examine the problem of an unstable field theory with euclidean action:

$$\mathcal{S}(\phi) = \Omega^{-1}\int \mathrm{d}^d x\, \left[\tfrac{1}{2}(\partial_\mu\phi)^2 + \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{6}g\phi^3\right]. \tag{41.42}$$

We know that a quantum state concentrated around the minimum $\phi(x) = 0$ 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 at leading order reads:

$$\mathcal{A}(\phi) = \tfrac{1}{2}\Omega^{-1}\int \mathrm{d}^d x\, \mathrm{d}t\, \left[\dot{\phi}^2 + \tfrac{1}{4}\left(-\nabla^2\phi + m^2\phi + \tfrac{1}{2}g\phi^2\right)^2\right]. \tag{41.43}$$

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. This means that the relevant maximum of the action instead corresponds to a static instanton configuration. The arguments of Section 41.2.2, then lead to the estimate

$$\tau = O\left(\exp[\mathcal{S}_{\text{inst.}}/\Omega]\right).$$

Note finally that the inequality (41.36) we have used is based on a structure which is typical for scalar *supersymmetric* field theories (see the action (17.88)).

Degenerate minima. Another 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 41.1 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

$$E(q) = q^2(1-q)^2, \tag{41.44}$$

and therefore,

$$E'^2(q) = 4(1-2q)^2 q^2 (1-q)^2. \tag{41.45}$$

This time we look for instanton solutions which connect $q = 0$ 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 will correspond to glue together two instantons. Therefore, the difference between the two leading eigenvalues, which is also the second eigenvalue ϵ_1, is again of the form:

$$\epsilon_1 \sim \exp[-\left(E_{\max} - E_{\min}\right)/\Omega]. \tag{41.46}$$

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.

41.4 Instantons in Stable Boson Field Theories: General Remarks

We now briefly discuss instantons in stable theories, connecting for example degenerate classical minima. The most interesting examples correspond unfortunately to scale invariant classical theories and the evaluation of the instanton contributions at leading order, which formally follows the lines presented in Chapter 40, leads to difficulties both due to UV and IR divergences. Some of them are examined in Chapter 40. Since for the two examples we consider in Sections 41.5,41.6, they have not been satisfactorily solved yet, we restrict ourselves here to classical considerations.

We begin with a few simple remarks about the possible existence of instantons in stable field theories.

Scalar field theories. We first assume that the action has the form:

$$\mathcal{S}(\phi) = \int \left[\tfrac{1}{2} g_{ij}(\phi)\partial_\mu \phi^i \partial_\mu \phi^j + V(\phi)\right] \mathrm{d}^d x\,, \tag{41.47}$$

in which ϕ^i is a multicomponent scalar boson field, $g_{ij}(\phi)$ a positive matrix (positive definite almost everywhere) and

$$\min_{\{\phi\}} V(\phi) = 0\,. \tag{41.48}$$

Equation $(A40.3)$ immediately generalizes:

$$(2-d)\int \tfrac{1}{2}\left(g_{ij}(\phi)\partial_\mu \phi_c^i \partial_\mu \phi_c^j\right) \mathrm{d}^d x = d\int V(\phi_c)\mathrm{d}^d x\,.$$

We see that this equation has no solution for $d > 2$. For $d = 2$, it has only solutions if

$$V(\phi_c(x)) = 0 . \tag{41.49}$$

The condition (41.48) 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:

$$\frac{\delta}{\partial \phi^k(y)} \int \frac{1}{2} g_{ij}(\phi) \partial_\mu \phi^i \partial_\mu \phi^j \mathrm{d}^2 x = 0 .$$

These two equations are in general incompatible, except if $V(\phi)$ vanishes identically. In the latter case the action (41.47) corresponds to a two-dimensional model on a Riemannian manifold. A particular class of such models, models based on homogeneous spaces, has been discussed in Chapters 14,15. Among them, the $CP(N-1)$ models are known to admit instanton solutions and we describe them in Section 41.5.

Gauge theories. If, in addition to scalar fields, the theory contains gauge fields A_μ^a, the gauge invariant action has the form:

$$\mathcal{S}(\phi, \mathbf{A}_\mu) = \int \mathrm{d}^d x \left[\tfrac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + \tfrac{1}{2} \left(\mathrm{D}_\mu \phi_i \right)^2 + V(\phi) \right] . \tag{41.50}$$

We now assume the existence of a finite action solution $\{\phi^c,\ \mathbf{A}_\mu^c\}$ (in which $\mathbf{A}_\mu^c$ is not a pure gauge), and calculate the action for $\lambda \mathbf{A}_\mu^c(\lambda x)$ and $\phi^c(\lambda x)$. After the change of variables $\lambda x \mapsto x$, we obtain

$$\mathcal{S}(\phi^c, \mathbf{A}_\mu^c; \lambda) = \lambda^{4-d} \int \frac{1}{4} \left(\mathbf{F}_{\mu\nu} \right)^2 \mathrm{d}^d x + \lambda^{2-d} \int \frac{1}{2} \left(\mathrm{D}_\mu \phi \right)^2 \mathrm{d}^d x + \lambda^{-d} \int V(\phi) \mathrm{d}^d x . \tag{41.51}$$

Stationarity at $\lambda = 1$ implies

$$(4-d) \int \frac{1}{4} \left(\mathbf{F}_{\mu\nu} \right)^2 \mathrm{d}^d x + (2-d) \int \frac{1}{2} \left(\mathrm{D}_\mu \phi \right)^2 \mathrm{d}^d x - d \int V(\phi) \mathrm{d}^d x = 0 . \tag{41.52}$$

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:

$$V(\phi) = 0 , \tag{41.53a}$$

$$\mathrm{D}_\mu \phi = 0 . \tag{41.53b}$$

Writing the field equations, we conclude that $\mathbf{A}_\mu^c$ is the solution of the pure gauge field equations. As we show in Section 41.6, instantons can indeed be found in pure gauge theories. The equation (41.53*b*), which now is an equation for ϕ^c, then leads to the integrability conditions:

$$[\mathrm{D}_\mu, \mathrm{D}_\nu] = F_{\mu\nu} \Longrightarrow \left(F_{\mu\nu}^a \right)^c t_{ij}^a \phi_j^c = 0 , \tag{41.54}$$

in which the matrices t^a are the generators of the Lie algebra. The conditions (41.54) together with the equation (41.53*a*) show that in general the system has only the trivial solution $\phi^c = 0$.

41.5 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 more detailed analysis.

We consider a set of N complex fields φ_α, subject to the condition

$$\bar{\boldsymbol{\varphi}} \cdot \boldsymbol{\varphi} = 1\,. \tag{41.55}$$

In addition, two vectors $\boldsymbol{\varphi}$ and $\boldsymbol{\varphi}'$ are equivalent if

$$\boldsymbol{\varphi}'(x) = \mathrm{e}^{i\Lambda(x)}\, \boldsymbol{\varphi}(x)\,. \tag{41.56}$$

These conditions characterize the manifold $CP(N-1)$ (for $(N-1)$-dimensional Complex Projective) which is isomorphic to the symmetric space $U(N)/U(1)/U(N-1)$, a complex Grassmannian manifold and one of the symmetric spaces exhibited in Appendix A15.4.3. One form of the unique classical action is

$$\mathcal{S}(\boldsymbol{\varphi}, A_\mu) = \frac{1}{g}\int \mathrm{d}^2 x\, \overline{\mathrm{D}_\mu \boldsymbol{\varphi}} \cdot \mathrm{D}_\mu \boldsymbol{\varphi}\,, \tag{41.57}$$

in which D_μ is the covariant derivative:

$$\mathrm{D}_\mu = \partial_\mu + iA_\mu\,. \tag{41.58}$$

The field A_μ is a gauge field for the $U(1)$ transformations (41.56), and ensures the corresponding equivalence. Since the action contains no kinetic term for A_μ the gauge field is an auxiliary field that can be integrated out, which means that it can be replaced in the action by the solution of the A_μ field equation. Using equation (41.55) one finds

$$A_\mu = i\bar{\boldsymbol{\varphi}} \cdot \partial_\mu \boldsymbol{\varphi}\,. \tag{41.59}$$

After this substitution the field $\bar{\boldsymbol{\varphi}} \cdot \partial_\mu \boldsymbol{\varphi}$ is a composite gauge field.

Note that the $CP(1)$ model is locally isomorphic to the $O(3)$ non-linear σ-model, with the identification

$$\boldsymbol{\phi} = \bar{\varphi}_\alpha \boldsymbol{\sigma}_{\alpha\beta} \varphi_\beta\,. \tag{41.60}$$

Instantons. To prove the existence of locally stable non-trivial minima of the action we use the following inequality (note the analogy with (41.36)):

$$\int \mathrm{d}^2 x\, \left|\mathrm{D}_\mu \boldsymbol{\varphi} \mp i\epsilon_{\mu\nu} \mathrm{D}_\nu \boldsymbol{\varphi}\right|^2 \ge 0\,, \tag{41.61}$$

($\epsilon_{\mu\nu}$ being the antisymmetric tensor, $\epsilon_{12} = 1$). Expanding the expression we obtain

$$\mathcal{S}(\boldsymbol{\varphi}) \ge |Q(\boldsymbol{\varphi})|/g\,, \tag{41.62}$$

with

$$Q(\boldsymbol{\varphi}) = -i\epsilon_{\mu\nu}\int \mathrm{d}^2 x\, \mathrm{D}_\mu \boldsymbol{\varphi} \cdot \overline{\mathrm{D}_\nu \boldsymbol{\varphi}} = i\int \mathrm{d}^2 x\, \epsilon_{\mu\nu} \mathrm{D}_\nu \mathrm{D}_\mu \boldsymbol{\varphi} \cdot \bar{\boldsymbol{\varphi}}\,. \tag{41.63}$$

Then,

$$i\epsilon_{\mu\nu}\mathrm{D}_\nu\mathrm{D}_\mu = \tfrac{1}{2}i\epsilon_{\mu\nu}[\mathrm{D}_\nu,\mathrm{D}_\mu] = \tfrac{1}{2}F_{\mu\nu}\,, \tag{41.64}$$

where $F_{\mu\nu}$ is the curvature

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,.$$

Therefore, using (41.55),

$$Q(\varphi) = \frac{1}{2}\int \mathrm{d}^2x\,\epsilon_{\mu\nu}F_{\mu\nu}\,. \tag{41.65}$$

The integrand is proportional to the two-dimensional abelian chiral anomaly, as shown in Section 20.3.2 (see the general expression (20.101)), and is, therefore, a total divergence

$$\tfrac{1}{2}\epsilon_{\mu\nu}F_{\mu\nu} = \partial_\mu\epsilon_{\mu\nu}A_\nu\,.$$

Substituting this form into equation (41.65) and integrating in a large disk of radius R, we find

$$Q(\varphi) = \lim_{R\to\infty}\oint_{|x|=R}\mathrm{d}x_\mu\,A_\mu(x).$$

$Q(\varphi)$ thus depends only on the behaviour of the classical solution for $|x|$ large and is a topological charge. Finiteness of the action demands that at large distances $\mathrm{D}_\mu\varphi$ vanishes. This implies, from (41.64) that $F_{\mu\nu}$ vanishes and thus A_μ is a pure gauge (and φ a gauge transform of a constant vector)

$$A_\mu = \partial_\mu\Lambda(x) \;\Rightarrow\; Q(\varphi) = \lim_{R\to\infty}\oint_{|x|=R}\mathrm{d}x_\mu\partial_\mu\Lambda(x)\,. \tag{41.66}$$

The topological charge measures the variation of the angle $\Lambda(x)$ on a large circle, which 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 to S_1, which are characterized by an integer n, the winding number, and

$$Q(\varphi) = 2\pi n \implies \mathcal{S}(\varphi) \ge 2\pi|n|/g\,. \tag{41.67}$$

The equality $\mathcal{S}(\varphi) = 2\pi|n|/g$ corresponds to a local minimum and implies that the classical solutions satisfy first-order partial differential (self-duality) equations:

$$\mathrm{D}_\mu\varphi = \pm i\epsilon_{\mu\nu}\mathrm{D}_\nu\varphi\,. \tag{41.68}$$

Vectors, solutions of the equations (41.68) are proportional to holomorphic or anti-holomorphic (depending on the sign) vectors in the variable $z = x_1 + ix_2$ (this reflects the conformal invariance of the classical field theory). Using then the equation (41.55) and the equivalence (41.56) one can cast the holomorphic solution into the form (up to a gauge transformation)

$$\varphi_\alpha = P_\alpha(z)\Big/\sqrt{P\cdot\bar P}\;, \tag{41.69}$$

where $P_\alpha(z)$ are polynomials in z without common roots. The anti-holomorphic solution corresponds to interchange φ and $\bar\varphi$. Translating the $CP(1)$ minimal solution (polynomials of degree 1) in the $O(3)$ σ model language one finds the stereographic mapping of the sphere S_2 onto the plane

$$\phi_1 = \frac{z+\bar z}{1+\bar z z}\,,\quad \phi_2 = i\frac{z-\bar z}{1+\bar z z}\,,\quad \phi_3 = \frac{1-\bar z z}{1+\bar z z}\,.$$

The structure of the semi-classical vacuum. In contrast to our analysis of similar problems in quantum mechanics, we have discussed here the existence of instantons without reference to the structure of the classical vacuum. To find an interpretation of instantons in gauge theories, it is interesting to express the results in the temporal gauge. Then 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:

$$\varphi(x_1) = \mathrm{e}^{i\Lambda(x_1)}\,\mathbf{v}\,, \quad \bar{\mathbf{v}}\cdot\mathbf{v} = 1\,.$$

Moreover, if the vacuum state is invariant under space reflection $\varphi(+\infty) = \varphi(-\infty)$ and thus

$$\Lambda(+\infty) - \Lambda(-\infty) = 2\nu\pi \quad \nu \in \mathbb{Z}\,.$$

Again ν is a topological number that classifies degenerate classical minima, and the semi-classical vacuum thus has a periodic structure. This analysis is consistent with Gauss's law (Section 19.3) which only implies that states are invariant under infinitesimal gauge transformations, and, therefore, under gauge transformations of the class $\nu = 0$ which are continuously connected to the identity.

We now consider a large rectangle with 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 at fixed time. One finds for $R \to \infty$,

$$\Lambda(+\infty,0) - \Lambda(-\infty,0) - [\Lambda(+\infty,T) - \Lambda(-\infty,T)] = 2n\pi\,.$$

Therefore, instantons interpolate between different classical minima. Like in the case of the cosine potential, one projects onto a proper quantum eigenstate, the "θ-vacuum" corresponding to an angle θ by adding, in analogy with the expressions (43.63,43.62), a topological term to the classical action:

$$\mathcal{S}(\varphi) \mapsto \mathcal{S}(\varphi) + i\frac{\theta}{4\pi}\int \mathrm{d}^2x\,\epsilon_{\mu\nu}F_{\mu\nu}.$$

41.6 Instantons in the $SU(2)$ Gauge Theory

We now give an example of instantons in four dimensions. According to the analysis of Section 41.4, 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 $\mathbf{A}_\mu$ is a vector and the gauge action reads:

$$\mathcal{S}(\mathbf{A}_\mu) = \frac{1}{4g}\int [\mathbf{F}_{\mu\nu}(x)]^2\,\mathrm{d}^4x\,, \tag{41.70}$$

with

$$\mathbf{F}_{\mu\nu} = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu + \mathbf{A}_\mu \times \mathbf{A}_\nu\,. \tag{41.71}$$

The existence and some properties of instantons in this theory follow from considerations analogous to those presented for the $CP(N-1)$ model.

We define the dual of the tensor $\mathbf{F}_{\mu\nu}$ by

$$\tilde{\mathbf{F}}_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}\mathbf{F}_{\rho\sigma} \,. \tag{41.72}$$

Then the inequality

$$\int \mathrm{d}^4x \left[\mathbf{F}_{\mu\nu}(x) \pm \tilde{\mathbf{F}}_{\mu\nu}(x)\right]^2 \geq 0 \,, \tag{41.73}$$

implies

$$\mathcal{S}(\mathbf{A}_\mu) \geq |Q(\mathbf{A}_\mu)|/4g \,, \tag{41.74}$$

where $Q(\mathbf{A}_\mu)$ is an expression we have already met in Section 20.3.3 (equation (20.105), here written in $SO(3)$ notation) in the calculation of the axial anomaly

$$Q(\mathbf{A}_\mu) = \int \mathrm{d}^4x \, \mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu} \,. \tag{41.75}$$

We have shown that the quantity $\mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu}$ is a pure divergence (equation (20.106)),

$$\mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu} = \partial_\mu V_\mu \,, \tag{41.76}$$

with

$$V_\mu = 2\epsilon_{\mu\nu\rho\sigma}\left[\mathbf{A}_\nu \cdot \partial_\rho \mathbf{A}_\sigma + \tfrac{1}{3}\mathbf{A}_\nu \cdot (\mathbf{A}_\rho \times \mathbf{A}_\sigma)\right]. \tag{41.77}$$

The integral thus depends only on the behaviour of the gauge field at large distances and its values are quantized (equation (20.112)). Here again, as in 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,

$$-\tfrac{1}{2}i\mathbf{A}_\mu \cdot \boldsymbol{\sigma} = \mathbf{g}(x)\partial_\mu \mathbf{g}^{-1}(x) + O\left(|x|^{-2}\right) \quad |x| \to \infty \,, \tag{41.78}$$

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 the study of the homotopy classes of mappings from S_3 to S_3, which are also classified by an integer, the winding number.

The simplest one to one mapping corresponds to an element $\mathbf{g}(x)$ of the form

$$\mathbf{g}(x) = \frac{x_4 + i\mathbf{x}\cdot\boldsymbol{\sigma}}{r}, \quad r = (x_4^2 + \mathbf{x}^2)^{1/2}, \tag{41.79}$$

and thus

$$A_m^i \underset{r\to\infty}{\sim} 2\left(x_4\delta_{im} + \epsilon_{imk}x_k\right) r^{-2}, \quad A_4^i = -2x_i r^{-2}. \tag{41.80}$$

It follows that

$$\int \mathrm{d}^4x \, \mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu} = \int \mathrm{d}\Omega \, \hat{n}_\mu V_\mu = 32\pi^2, \tag{41.81}$$

in which $\mathrm{d}\Omega$ is the measure on the sphere and $\hat{n}_\mu$ the unit vector normal to the sphere.

If we compare this result with equation (20.112) we see that we have indeed found the minimal action solution. In general, we then expect

$$\int \mathrm{d}^4x \, \mathbf{F}_{\mu\nu} \cdot \tilde{\mathbf{F}}_{\mu\nu} = 32\pi^2 n \,, \tag{41.82}$$

and therefore,

$$\mathcal{S}(\mathbf{A}_\mu) \geq 8\pi^2 |n|/g\,. \tag{41.83}$$

The equality, which corresponds to a local minimum of the action, is obtained for fields satisfying the self-duality equations

$$\mathbf{F}_{\mu\nu} = \pm\tilde{\mathbf{F}}_{\mu\nu}\,, \tag{41.84}$$

which are first-order partial differential equations. The one-instanton solution, which depends on an arbitrary scale parameter λ, is

$$A^i_m = \frac{2}{r^2+\lambda^2}\left(x_4\delta_{im} + \epsilon_{imk}x_k\right), \quad m=1,2,3\,, \quad A^i_4 = -\frac{2x_i}{r^2+\lambda^2}. \tag{41.85}$$

The semi-classical vacuum. In analogy with the analysis of the $CP(N-1)$ model we now introduce 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:

$$-\tfrac{1}{2}i\mathbf{A}_m\cdot\boldsymbol{\sigma} = \mathbf{g}(x_i)\partial_m\mathbf{g}^{-1}(x_i)\,. \tag{41.86}$$

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 $|x|\to\infty$), 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 (41.85) connects minima with different winding numbers. Therefore, as in the case of the $CP(N-1)$ model, to project onto a θ-vacuum, one adds a term to the classical action of gauge theories:

$$\mathcal{S}_\theta(\mathbf{A}_\mu) = \mathcal{S}(\mathbf{A}_\mu) + \frac{i\theta}{32\pi^2}\int d^4x\, \mathbf{F}_{\mu\nu}\cdot\tilde{\mathbf{F}}_{\mu\nu}\,, \tag{41.87}$$

and then integrates over all fields $\mathbf{A}_\mu$ without restriction. At least in the semi-classical approximation, the gauge theory depends on an additional parameter, the angle θ. For non-vanishing values of θ the additional term violates CP conservation, and is at the origin of the strong CP violation problem, because if θ does not vanish experimental bounds are consistent only with unnaturally small values.

Fermions in an instanton background. In QCD (see Section 20.2) gauge fields are coupled to quarks $\mathbf{Q}$, $\bar{\mathbf{Q}}$ with an action (and $SU(3)$ notation):

$$\mathcal{S}(\mathbf{A}_\mu,\bar{\mathbf{Q}},\mathbf{Q}) = -\int \mathrm{d}^4x\left[\frac{1}{4g^2}\operatorname{tr}\mathbf{F}^2_{\mu\nu} + \sum_{f=1}^{N_f}\bar{\mathbf{Q}}_f\left(\not{\mathbf{D}} + m_f\right)\mathbf{Q}_f\right].$$

Two comments can be made. First if the θ term in (41.87) contributes and one fermion field is massless, according to the analysis of Section 20.3.4 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 functional integral and the strong CP violation problem is solved. However, such an hypothesis seems to be inconsistent with experimental data.

Second, as we have already discussed in Section 20.5, 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 non-abelian gauge theories the classical theory is scale invariant. Therefore, solutions depend on a scale parameter which, therefore, is a collective coordinate over which one has to integrate. This leads to difficult problems as the analysis of the massless ϕ_4^4 field theory reveals (see Chapter 40). Both theories are asymptotically free and the problems come from the infrared 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 34,35).

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 alter drastically the physical picture.

Bibliographical Notes

Polyakov has emphasized the double-well as the typical instanton problem in

A.M. Polyakov, *Nucl. Phys.* B120 (1977) 429.

The role of instantons in the question of spontaneous symmetry breaking of discrete symmetries has been stressed in

J. Zinn-Justin, *Basko Polje Summer School 1976*, Saclay preprint DPh-T/76/99, unpublished.

We have used scaling arguments to discuss the existence of instantons

G.H. Derrick, *J. Math. Phys.* 5 (1964) 1252.

References about instantons in the $CP(N-1)$ model include

A. Jevicki *Nucl. Phys.* B127 (1977) 125; D. Förster, *Nucl. Phys.* B130 (1977) 38; M. Lüscher, *Phys. Lett.* 78B (1978) 465; A. D'Adda, P. Di Vecchia and M. Lüscher, *Nucl. Phys.* B146 (1978) 63; B152 (1979) 125; H. Eichenherr, *Nucl. Phys.* B146 (1978) 215; V.L. Golo and A. Perelemov, *Phys. Lett.* 79B (1978) 112; A.M. Perelemov, *Phys. Rep.* 146 (1987) 135.

For instantons in gauge theories see

A.A. Belavin, A.M. Polyakov, A.S. Schwartz and Yu S. Tyupkin, *Phys. Lett.* 59B (1975) 85; G. 't Hooft, *Phys. Rev. Lett.* 37 (1976) 8; *Phys. Rev.* D14 (1976) 3432 (Erratum *Phys. Rev.* D18 (1978) 2199); R. Jackiw and C. Rebbi, *Phys. Rev. Lett.* 37 (1976) 172; C.G. Callan, R.F. Dashen and D.J. Gross, *Phys. Lett.* 63B (1976) 334; A.A. Belavin and A.M. Polyakov, *Nucl. Phys.* B123 (1977) 429; F.R. Ore, *Phys. Rev.* D16 (1977) 2577; S. Chadha, P. Di Vecchia, A. D'Adda and F. Nicodemi, *Phys. Lett.* 72B (1977) 103; T. Yoneya, *Phys. Lett.* 71B (1977) 407; I.V. Frolov and A.S. Schwarz, *Phys. Lett.* 80B (1979) 406; E. Corrigan, P. Goddard and S. Templeton, *Nucl. Phys.* B151 (1979) 93.

We have met several examples of the Bogomolnyi inequality

E.B. Bogomolnyi, *Sov. J. Nucl. Phys.* 24 (1976) 449; M.K. Prasad and C.M. Sommerfeld, *Phys. Rev. Lett.* 35 (1975) 760.

For a discussion of the $U(1)$ problem see

G. 't Hooft, *Phys. Rep.* 142 (1986) 357.

Possible solutions to the strong CP violation problem were discussed in

R.D. Peccei and H.R. Quinn, *Phys. Rev.* D16 (1977) 1791; S. Weinberg, *Phys. Rev. Lett.* 40 (1978) 223; F. Wilczek, *Phys. Rev. Lett.* 40 (1978) 279.

Instantons play also a role in the two-dimensional abelian Higgs model

C.G. Callan, R.F. Dashen and D.J. Gross, *Phys. Lett.* 66B (1977) 375.

APPENDIX A41

A41.1 Trace Formula for Periodic Potentials

We consider a hamiltonian H corresponding to a real periodic potential $V(x)$ with period τ:

$$V(x+\tau) = V(x). \tag{A41.1}$$

Eigenfunctions $\psi_\varphi(x)$ are then also eigenfunctions of the translation operator T:

$$T\psi_\varphi(x) \equiv \psi_\varphi(x+\tau) = \mathrm{e}^{i\varphi}\,\psi_\varphi(x). \tag{A41.2}$$

Let us put the system in a box of size $N\tau$ with periodic boundary conditions. This implies a quantization of φ

$$\mathrm{e}^{iN\varphi} = 1 \Rightarrow \varphi = \varphi_p \equiv \frac{2\pi p}{N}, \quad 0 \le p < N\,. \tag{A41.3}$$

Calling $\psi_{p,n}$ the eigenfunction of H corresponding to the band n and the pseudo-momentum φ_p, we can write the matrix elements of $T\,\mathrm{e}^{-\beta H}$:

$$\langle x'|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle = \sum_{p,n} \psi^*_{p,n}(x')\,\mathrm{e}^{-\beta E_n(\varphi_p)}\,\psi_{p,n}(x+\tau), \tag{A41.4}$$

and therefore,

$$\langle x'|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle = \sum_{p,n} \psi^*_{p,n}(x')\,\mathrm{e}^{-\beta E_n(\varphi_p)+i\varphi_p}\,\psi_{p,n}(x). \tag{A41.5}$$

This implies for the diagonal elements:

$$\langle x|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle = \sum_{p,n} |\psi_{p,n}(x)|^2\,\mathrm{e}^{-\beta E_n(\varphi_p)+i\varphi_p}\,. \tag{A41.6}$$

Because in a translation of a period (equation (*A*41.2)), the eigenfunctions are multiplied by a phase, the equation (*A*41.6) shows that the l.h.s. is a periodic function of x with periodicity τ. Therefore,

$$\int_0^\tau \langle x|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle\,\mathrm{d}x = \sum_{n,p} \mathrm{e}^{i\varphi_p-\beta E_n(\varphi_p)}\,\frac{1}{N}\int_0^{N\tau} |\psi_{p,n}(x)|^2\,\mathrm{d}x\,. \tag{A41.7}$$

The functions $\psi_{p,n}(x)$ are orthonormal over $N\tau$:

$$\int_0^\tau \langle x|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle\,\mathrm{d}x = \frac{1}{N}\sum_{n,p} \mathrm{e}^{i\varphi_p-\beta E_n(\varphi_p)}\,. \tag{A41.8}$$

We now take the large N limit and obtain expression (41.66):

$$\int_0^\tau \langle x|\,T\,\mathrm{e}^{-\beta H}\,|x\rangle\,\mathrm{d}x = \frac{1}{2\pi}\sum_{n}\int_0^{2\pi} \mathrm{e}^{i\varphi-\beta E_n(\varphi)}\,\mathrm{d}\varphi\,, \tag{A41.9}$$

since reality implies

$$E_n(\varphi) = E_n(-\varphi). \tag{A41.10}$$

42 PERTURBATION SERIES AT LARGE ORDERS. SUMMATION METHODS

In Chapter 39, we have discussed the analytic structure of the ground state energy $E(g)$ of the anharmonic oscillator. We have argued that $E(g)$ is analytic in a cut-plane, and calculated by instanton methods its imaginary part on the cut for g small and negative. On the other hand, perturbation theory yields $E(g)$ for g small as a power series in g:

$$E(g) = \sum_{k=0}^{\infty} E_k g^k . \tag{42.1}$$

We explain in this chapter how the behaviour of $\operatorname{Im} E(g)$ for $g \to 0$ is related to the behaviour of the coefficients E_k when the order k becomes large. We then generalize the method to the class of potentials for which we have calculated instanton contributions. The same method can be readily applied to boson field theories, using the results of Chapter 40, while the extension to field theories involving fermions like QED requires, as we show, some additional considerations.

We already know that the expansion (42.1) is divergent for all values of g. This implies that, even for g small, the series does not determine the function $E(g)$ uniquely. We thus examine the implications of the large order behaviour for the problem of the summation of the series. Finally, we describe a few practical methods commonly used to sum divergent series of the type met in quantum mechanics and quantum field theory. Some of these methods have been successfully applied to the $(\phi^2)^2$ field theory in two and three dimensions and have led to the precise predictions of critical exponents displayed in Chapter 29.

42.1 Quantum Mechanics

We first examine two situations where we have already found instantons. We then argue that for other analytic potentials complex solutions to the euclidean equation of motion are relevant.

42.1.1 Real instantons

The anharmonic oscillator. We first consider the ground state energy $E(g)$ of the quartic anharmonic oscillator (39.1). Since $E(g)$ is analytic in the cut-plane and behaves like $g^{1/3}$ for g large, it has a Cauchy representation of the form

$$E(g) = \frac{1}{2} + \frac{g}{\pi} \int_{-\infty}^{0} \frac{\operatorname{Im} E(g')\, dg'}{g'(g'-g)} . \tag{42.2}$$

Expanding the integrand in powers of g, we obtain an integral representation for the coefficients E_k:

$$E_k = \frac{1}{\pi} \int_{-\infty}^{0} \frac{\operatorname{Im} E(g)\, dg}{g^{k+1}} \qquad \text{for } k > 0 . \tag{42.3}$$

When k, the order in the expansion, becomes large, due to the factor g^{-k} the dispersion integral (42.3) is dominated by the small negative g values. In Chapter 39, we have calculated $\operatorname{Im} E(g)$ for g small and negative. We can here use this result to estimate the large k behaviour of E_k:

$$E_k \underset{k\to\infty}{\sim} \frac{1}{\pi}\int^{0-}\left(\frac{8}{\pi}\right)^{1/2}\frac{1}{\sqrt{-g}}\frac{\mathrm{e}^{4/3g}}{g^{k+1}}\left[1+O(g)\right]\mathrm{d}g\,. \tag{42.4}$$

The explicit integration yields

$$E_k = (-1)^{k+1}\left(\frac{6}{\pi^3}\right)^{1/2}\left(\frac{3}{4}\right)^k \Gamma(k+1/2)\left[1+O\left(1/k\right)\right]. \tag{42.5}$$

Successive corrections to the semi-classical result yield a series in powers of g which, integrated, generates a systematic expansion in powers of $1/k$.

General potentials. The same argument is applicable to the situation described in Section 39.5. We can calculate the energy of the unstable 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 above, 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. Let us consider the action

$$\mathcal{S}(q) = \int \mathrm{d}t\left[\tfrac{1}{2}\dot{q}^2(t) + g^{-1}V(q\sqrt{g})\right]. \tag{42.6}$$

The analogue of the dispersion relation (42.3) is

$$E_k \sim \frac{1}{\pi}\int_0^{\infty}\frac{\operatorname{Im} E(g)}{g^{k+1}}\mathrm{d}g\,.$$

The behaviour of $\operatorname{Im} E(g)$ for g small is given by expression (39.103). Integrating near $g=0$, we obtain

$$E_k \sim -\frac{1}{2\pi^{3/2}}x_+ \exp\left[\int_0^{x_+}\left(\frac{1}{\sqrt{2V(x)}}-\frac{1}{x}\right)\mathrm{d}x\right]A^{-(k+1/2)}\Gamma(k+1/2), \tag{42.7}$$

where A is the classical action

$$A = 2\int_0^{x_+}\sqrt{2V(x)}\,\mathrm{d}x\,. \tag{42.8}$$

We now see generic features emerge: at large orders, the perturbative coefficients E_k behave like

$$E_k \underset{k\to\infty}{\sim} Ck^{b-1}k!\,A^{-k}. \tag{42.9}$$

The factor $k!$ is universal and characteristic of the semi-classical or loop expansion. It shows that the perturbation series is a divergent series. The factor A^{-k} depends only on the action, since it is the action of the classical solution; in particular, it also characterizes the behaviour at large orders of the excited energy levels or of correlation functions. The power k^b comes from the power of g in front of the result. It, in particular, depends on the number of continuous symmetries broken by the classical solution, but it would also change if we considered an excited state rather than the ground state. This can be verified by explicitly calculating the imaginary parts of the energy of the excited levels as explained in Chapter 39, and using equation (42.3). The parameter b is in general a half integer. Finally, there is a constant multiplicative factor c which depends in a more complicated way on all the specific features of the expanded quantity.

Discussion. In both examples, we have been able to calculate the large order behaviour of perturbation series from the decay rate due to barrier penetration of an unstable minimum of the potential. For the potentials considered in Section 39.5 the action A is positive and, therefore, all terms in the perturbative expansion have the same sign. The same property holds for the anharmonic oscillator in the unstable case, that is, when g is negative. However, for g positive, in which case perturbation series has been expanded around the stable 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 will help to understand how to obtain the large behaviour in the generic stable case.

42.1.2 Complex instantons

We have up to now 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 were expanding around an absolute minimum of the potential, but which by analytic continuation in the coupling constant could be transformed into the unstable one. We now consider actions of the form (42.6), in which the potential $V(q)$ is an entire function of q and satisfies the condition

$$V(q) = \tfrac{1}{2}q^2 + O\left(q^3\right),$$

and assume that perturbation theory is expanded around $q = 0$, the absolute minimum of the potential. Then, clearly 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 allow an analytic continuation to an unstable situation. We then obtain the large order behaviour from the expression (42.7). We then use the inverse analytic continuation to return to the initial situation. If nothing dramatic happens, the large behaviour of the initial expansion will be given by the analytic continuation of the expression (42.7).

We can now formulate the rules of the large order behaviour calculation directly in the initial theory: to the complex zeros (at finite or infinite distance) of the potential $V(q)$ are associated complex instanton solutions, with, in general, complex (or exceptionally negative) action. These instantons are candidates to contribute to the large order behaviour. In the expression (42.7), 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 dominant contribution to the large order behaviour. Note that the difference we have found between the anharmonic oscillator and the unstable case is generic. In the stable case, the classical action is non-real positive, and the the perturbative coefficients at large order have an order-dependent phase factor.

Such a property plays an essential role in the question of summability of divergent series (see Section 42.6).

A special case. The previous discussion does not immediately apply to the case of potentials with degenerate minima. Let us indeed consider such a potential as the limit of a potential which has two minima at which the values of the potential are very close.

From the explicit form of the action, we see that the classical action has a limit which is twice the action of the instanton which connects the two minima of the potential

$$A = 2\int_0^{x_+} \sqrt{2V(x)}\mathrm{d}x \,.$$

However, the amplitude in front of the expression (42.7) diverges when x_+ is an extremum of the potential. This result can be easily understood. When the values at the two minima approach each other, the time spent close to the second minimum of the potential by the classical trajectory corresponding to the instanton solution diverges. Therefore, fluctuations which tend to change this time leave the action almost stationary. Correspondingly one eigenvalue of the operator $\delta^2 S(q_c)/\delta q\delta q$ goes to zero, and this explains the divergence of expression (42.7) in this case. Only by allowing this time to fluctuate, and introducing an additional time collective coordinate can we obtain the correct answer. Let us also note that here, as in the case of unstable minima, the classical action is positive. This is the source of serious difficulties when one tries to sum the perturbation series, a problem we discuss in Chapter 43.

42.2 Scalar Field Theory

In Chapter 40, we have evaluated the contributions of instantons to the decay rate of metastable states. We can here apply these results to large order behaviour estimates. In the case of the ϕ^4 field theory, the discontinuity across the cut of the n-point function reads (equation (40.22))

$$\operatorname{disc.} Z^{(n)}(x_1,\ldots,x_n) \sim \left(\frac{A}{2\pi}\right)^{d/2} \frac{\mathrm{e}^{-A/g}}{g^{(d+n)/2}} \left(\det M' M_0^{-1}\right)^{-1/2}_{\mathrm{ren}} F_n(x_1,\ldots,x_n) \qquad (42.10)$$

with

$$F_n(x_1,\ldots,x_n) = m^{d+n(d-2)/2} 6^{n/2} \int \mathrm{d}^d x_0 \prod_{i=1}^{n} f\bigl(m(x_i - x_0)\bigr). \qquad (42.11)$$

Using previous arguments, we can immediately translate this result into a large order behaviour estimate for correlation functions

$$\left\{Z^{(n)}(x_1,\ldots,x_n)\right\}_k = \frac{1}{2i\pi}\int \frac{\mathrm{d}g}{g^{k+1}} \operatorname{disc} Z^{(n)}(x_1,\ldots,x_n),$$

and, therefore,

$$\left\{Z^{(n)}(x_1,\ldots,x_n)\right\}_k \underset{k\to\infty}{\sim} \frac{1}{2i\pi}\frac{1}{(2\pi)^{d/2}} F_n(x_1,\ldots,x_n) A^{-n/2-k}\Gamma(k+d/2+n/2). \quad (42.12)$$

In a general scalar boson field theory, if instanton solutions can be found, the same arguments will lead to

$$\left\{Z^{(n)}(x_1,\ldots,x_n)\right\}_k \underset{k\to\infty}{\sim} \sum_{\substack{\text{dominant}\\ \text{saddle points}}} C_n(x_1,\ldots,x_n) k^{b-1} A^{-k} k!\,, \qquad (42.13)$$

in which

(i) A is the instanton action, which is in general complex;
(ii) $b = \frac{1}{2}(n+\delta)$ and δ is the number of symmetries broken by the classical solution;
(iii) $C_n(x_1, \ldots, x_n)$, which does not depend on k, contains the whole dependence in the external arguments.

Example: the renormalization group β-function in the $(\phi^2)^2$ field theory in three dimensions. The large order behaviour has been determined by solving numerically the field equations to find the classical action A and calculating numerically the determinant. The predictions of the asymptotic formulae have been compared with the terms of the series which have been calculated (see Section 29.2). The agreement is quite reasonable and gives us confidence that the large order behaviour estimates are indeed correct (see table 42.1).

Table 42.1

The coefficients β_k of the coupling constant RG function $\beta(g)$ divided by the large order estimate for the $O(N)$ symmetric ϕ_3^4 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

42.3 The ϕ^4 Field Theory in Four Dimensions

As a byproduct of the calculation of the instanton contribution in Sections 40.3–40.6, we can evaluate the semi-classical contribution to the large order behaviour in the ϕ^4 field theory in four dimensions. However, because the theory is exactly renormalizable, we will discover that, as a consequence of their large momenta properties, individual diagrams at order k grow themselves like $k!$, introducing some new complications in the large order behaviour analysis. Moreover, IR singularities in the massless theory also yield contributions of order $k!$, but with a different sign.

42.3.1 Semi-classical contribution

The instanton contribution to the large order behaviour is given by

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k = \frac{1}{\pi}\int_{-\infty}^{0} \frac{\operatorname{Im}\Gamma^{(n)}(p_1, \ldots, p_n)}{g^{k+1}} dg\,. \tag{42.14}$$

This yields a result of the form

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k \underset{k\to\infty}{\sim} C_n(p_1, \ldots, p_n)\int^{0-} \frac{e^{8\pi^2/3g}}{(-g)^{n+5/2}}\frac{dg}{g^{k+1}}\,. \tag{42.15}$$

After integration, we obtain

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k \sim C_n(p_1, \ldots, p_n)(-1)^k\left(\frac{3}{8\pi^2}\right)^{n+3+k}\Gamma(k+n/2+5/2). \tag{42.16}$$

(i) A is the instanton action, which is in general complex;
(ii) $b = \frac{1}{2}(n + \delta)$ and δ is the number of symmetries broken by the classical solution;
(iii) $C_n(x_1, \ldots, x_n)$, which does not depend on k, contains the whole dependence in the external arguments.

Example: the renormalization group β-function in the $(\phi^2)^2$ field theory in three dimensions. The large order behaviour has been determined by solving numerically the field equations to find the classical action A and calculating numerically the determinant. The predictions of the asymptotic formulae have been compared with the terms of the series which have been calculated (see Section 29.2). The agreement is quite reasonable and gives us confidence that the large order behaviour estimates are indeed correct (see table 42.1).

Table 42.1

The coefficients β_k of the coupling constant RG function $\beta(g)$ divided by the large order estimate for the $O(N)$ symmetric ϕ_3^4 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

42.3 The ϕ^4 Field Theory in Four Dimensions

As a byproduct of the calculation of the instanton contribution in Sections 40.3–40.6, we can evaluate the semi-classical contribution to the large order behaviour in the ϕ^4 field theory in four dimensions. However, because the theory is exactly renormalizable, we will discover that, as a consequence of their large momenta properties, individual diagrams at order k grow themselves like $k!$, introducing some new complications in the large order behaviour analysis. Moreover, IR singularities in the massless theory also yield contributions of order $k!$, but with a different sign.

42.3.1 Semi-classical contribution

The instanton contribution to the large order behaviour is given by

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k = \frac{1}{\pi}\int_{-\infty}^{0} \frac{\operatorname{Im}\Gamma^{(n)}(p_1, \ldots, p_n)}{g^{k+1}} dg\,. \tag{42.14}$$

This yields a result of the form

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k \underset{k\to\infty}{\sim} C_n(p_1, \ldots, p_n) \int^{0-} \frac{e^{8\pi^2/3g}}{(-g)^{n+5/2}} \frac{dg}{g^{k+1}}\,. \tag{42.15}$$

After integration, we obtain

$$\left\{\Gamma^{(n)}(p_1, \ldots, p_n)\right\}_k \sim C_n(p_1, \ldots, p_n)(-1)^k \left(\frac{3}{8\pi^2}\right)^{n+3+k} \Gamma(k + n/2 + 5/2). \tag{42.16}$$

From this expression, it is straightforward to derive the large order behaviour of various RG functions in, for example, the fixed momentum subtraction scheme. A comparison between large order behaviour and explicit calculations can be found in table 42.2, in the case of the RG β-function.

Table 42.2

The coefficients β_k of the RG β-function divided by the asymptotic estimate, in the case of the $O(N)$ symmetric ϕ^4_4 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

The large order behaviour of Wilson–Fisher ε-expansion, which is important for the theory of Critical Phenomena, can instead only be guessed at because, as discussed above, the RG functions in the minimal subtraction scheme vanish 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 $g^*(\varepsilon)$ is

$$6g^*(\varepsilon) \sim 48\pi^2\varepsilon/(N+8)\,,$$

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, 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 40.7).

42.3.2 UV and IR (renormalons) contributions

Implicit in the large order behaviour calculation is the assumption 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 quantum mechanics, if we have convincing evidence that it is valid for super-renormalizable theories, it is much more questionable for renormalizable theories, not to mention massless renormalizable theories. We first explain the large momentum problem and then the IR problem of massless theories.

UV singularities: renormalons. If the semi-classical analysis is valid for the 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 naive calculation, non-trivial questions arise about the global RG properties of the theory. 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 $(\phi^2)^2$ field theory in dimension 4, in which ϕ is an N-component vector, and the model has an $O(N)$ symmetry,

$$\mathcal{S}(\phi) = \int \mathrm{d}^4x \left[\frac{1}{2}\left(\partial_\mu\phi\right)^2 + \frac{m^2}{2}\phi^2 + \frac{g}{4}\left(\phi^2\right)^2\right]. \tag{42.17}$$

We have shown in Section 30.5 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

$$F_2(p) = \frac{2g}{(2\pi)^4} \int \frac{\mathrm{d}^4 q}{[(p+q)^2 + m^2]\,[1 + NgB_{\mathrm{r}}(q)]} - \text{subtractions}, \tag{42.18}$$

where the renormalized "bubble" diagram is given by

$$B_{\mathrm{r}}(p) = \frac{1}{(2\pi)^4} \int \frac{\mathrm{d}^4 q}{[(p+q)^2 + m^2]\,(q^2 + m^2)} - \text{subtraction}. \tag{42.19}$$

For large momenta, $B_{\mathrm{r}}(p)$ behaves like

$$B_{\mathrm{r}}(p) \sim \frac{1}{8\pi^2} \ln(m/p), \quad p \to \infty. \tag{42.20}$$

Therefore, the sum of the bubble diagrams which appears in expression (42.18) has a singularity for g small (which justifies the large momentum approximation) and positive at momentum

$$|p| \sim m\, \mathrm{e}^{8\pi^2/Ng} \quad \text{for } g \to 0_+ . \tag{42.21}$$

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 (42.18), 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

$$\int_{|q| \gg 1} \frac{\mathrm{d}q}{q^3} \left[1 + \frac{Ng}{8\pi^2} \ln(m/q)\right]^{-1} + \cdots . \tag{42.22}$$

The change of variables $t = \ln(q/m)$ transforms the expression (42.22) into

$$\int^{\infty} \mathrm{d}t\, \mathrm{e}^{-2t} \frac{1}{1 - Ngt/(8\pi^2)}. \tag{42.23}$$

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 (42.18) 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 a power in 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. Let us for simplicity consider the case of a dimensionless ratio of correlation functions $R(p/m, g)$ without anomalous dimensions,

$$\left(m\frac{\partial}{\partial m} + \beta(g)\frac{\partial}{\partial g}\right) R(p/m, g) = 0. \tag{42.24}$$

The RG equation tells us 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

$$\beta(g)s'(g) = s(g), \tag{42.25}$$

which after integration yields

$$s(g) \sim \exp\left[\int^g \frac{\mathrm{d}g'}{\beta(g')}\right]. \tag{42.26}$$

For g small, $s(g)$ behaves like

$$\beta(g) = \beta_2 g^2 + O\left(g^3\right) \quad \text{with} \quad \beta_2 = \frac{N+8}{8\pi^2}, \tag{42.27}$$

$$s(g) \underset{g\to 0}{\propto} g^{-\beta_3/\beta_2^2}\,\mathrm{e}^{-1/\beta_2 g}. \tag{42.28}$$

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

$$\int_0 \frac{s(g)}{g^{k+1}}\mathrm{d}g \propto (\beta_2/2)^k\Gamma(k+1+2\beta_3/\beta_2^2), \tag{42.29}$$

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 (42.16).

These problems are in fact related to the question of the existence of the renormalized ϕ^4 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.

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, the renormalon singularities cancel. We conjecture on this basis and on the basis of the numerical evidence presented in Chapter 29 that the ε-expansion is free of renormalon singularities.

Massless renormalizable theories. We again 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 (42.19) 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

$$\int^\Lambda \frac{\mathrm{d}^4 q}{q^2\left(1+NgI(q)\right)} = \int \frac{\mathrm{d}^4 q}{q^2\left(1+\frac{N}{8\pi^2}g\ln(\Lambda/q)\right)}.$$

Expanded in powers of g this 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 yielding 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 41.5, 41.6), the Borel transform has also semi-classical singularities on the real positive axis.

42.4 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, unphysical values of the coupling constant. We expect, therefore, some modifications if we consider a system of self-interacting fermions, or of fermions interacting with bosons which themselves have no self-interaction. (Actually, the first case can be reduced to the second one by introducing an auxiliary boson field.) Indeed, the Pauli principle will make the decay of the false vacuum more difficult because several fermions cannot be in 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 will drive the decay of the vacuum, and the fermions will no longer play a role.

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. Let us instead consider a zero-dimensional fermion theory, that is, an integral over a finite number of fermion degrees of freedom:

$$I(\lambda) = \int \prod_{i=1}^{N} \mathrm{d}\bar{\xi}_i \mathrm{d}\xi_i \exp\left[\bar{\xi}_i D_{ij} \xi_j + \lambda C_{ijkl} \bar{\xi}_i \bar{\xi}_j \xi_k \xi_l\right]. \tag{42.30}$$

The quantities ξ_i and $\bar{\xi}_i$ are anticommuting variables and D_{ij} and C_{ijkl} are a set of 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.

42.4.1 Example of a Yukawa-like field theory

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 $\phi(x)$:

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)] \left[\mathrm{d}\bar{\psi}(x)\right] [\mathrm{d}\psi(x)] \exp\left[-\mathcal{S}(\phi, \bar{\psi}, \psi)\right], \tag{42.31}$$

in which the action is

$$\mathcal{S}(\phi, \bar{\psi}, \psi) = \int \mathrm{d}^d x \left[-\bar{\psi}\left(\not{\partial} + M + \sqrt{g}\phi\right)\psi + \tfrac{1}{2}\left(\partial_\mu \phi\right)^2 + \tfrac{1}{2} m^2 \phi^2\right]. \tag{42.32}$$

The parameter g is a loop expansion parameter. Since a fermion field has no classical limit, the expression (42.31) is not very well 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 we 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. After the integration over ψ and $\bar{\psi}$, we obtain

$$\mathcal{Z} = \int [\mathrm{d}\phi(x)] \det [\not{\partial} + M + \sqrt{g}\phi(x)] \exp \left[-\frac{1}{2}\int \mathrm{d}^d x \left((\partial_\mu \phi)^2 + m^2\phi^2\right)\right]. \qquad (42.33)$$

We are faced with a new difficulty arising from the integration, the effective action now is non-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 $\sqrt{g}$. As a consequence, essential singularities can only be generated by the infinite range of the ϕ-integration. It is thus sufficient to calculate the contribution to the functional integral of large fields $\phi(x)$. 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 allows for cancellations.

42.4.2 Evaluation of the fermion determinant for large fields

As an exercise we first solve a similar problem in which, however, some slight additional complications due to the spin structure are absent.

The Fredholm determinant of a Schrödinger operator for large potentials. We first evaluate, in the limit of large (smooth) potentials $V(x)$, the Fredholm determinant of the Schrödinger operator:

$$D(V) = \det\left[-\nabla^2 + \mu^2 + V(x)\right]\left[-\nabla^2 + \mu^2\right]^{-1}. \qquad (42.34)$$

On intuitive grounds, we expect the determinant to converge towards a local functional. To derive this property, we write the determinant using the identity

$$\operatorname{tr}\ln HH_0^{-1} = -\operatorname{tr}\int_0^\infty \frac{\mathrm{d}t}{t}\left[\mathrm{e}^{-Ht} - \mathrm{e}^{-H_0 t}\right], \qquad (42.35)$$

applied to the case

$$H = -\nabla^2 + \mu^2 + V(x)\,, \qquad H_0 = -\nabla^2 + \mu^2. \qquad (42.36)$$

When $|V(x)|$ becomes large, the integral over t is dominated by the small t region. The evaluation for t small of the evolution operator e^{-tA}, corresponding to the Schrödinger operator A, is a problem we have solved in Section 2.2, in the construction of the path integral representation. Using directly equation (2.15) and replacing for t small the argument in the potential V by a constant x we find (it is also the leading order in equation (2.69))

$$\operatorname{tr}\mathrm{e}^{-Ht} - \operatorname{tr}\mathrm{e}^{-H_0 t} \underset{t\to 0}{\sim} \frac{1}{(4\pi t)^{d/2}}\int \mathrm{d}^d x \left[\mathrm{e}^{-t[V(x)+\mu^2]} - \mathrm{e}^{-t\mu^2}\right]. \qquad (42.37)$$

As we have indicated there, this evaluation is valid only if $V(x)$ is at least continuous. It then follows

$$\operatorname{tr}\ln HH_0^{-1} \sim -\frac{1}{(4\pi)^{d/2}}\int_0^\infty \frac{\mathrm{d}t}{t^{1+d/2}}\int \mathrm{d}^d x \left[\mathrm{e}^{-t[V(x)+\mu^2]} - \mathrm{e}^{-t\mu^2}\right]. \tag{42.38}$$

For $d < 2$, we can also integrate over t and finally obtain

$$\operatorname{tr}\ln HH_0^{-1} \sim -\frac{\Gamma(-d/2)}{(4\pi)^{d/2}}\int \mathrm{d}^d x \left[\left(V(x)+\mu^2\right)^{d/2} - \mu^d\right]. \tag{42.39}$$

For $d \geq 2$, we know that the quantity (42.34), which has the form of a one-loop diagram in a scalar field theory, has to be renormalized. For $d = 2$, we have to add a mass counterterm. Since $V(x)$ is large we can neglect μ. We then obtain the evaluation

$$\begin{aligned}\ln D(V) &\sim \lim_{d\to 2}\left\{-\frac{1}{(4\pi)^{d/2}}\int \mathrm{d}^d x\left[\Gamma(-d/2)V^{d/2}(x) + \Gamma(1-d/2)V(x)\right]\right\},\\ &\sim -\frac{1}{4\pi}\int \mathrm{d}^2 x\, V(x)\ln V(x)\,. \end{aligned} \tag{42.40}$$

In the same limit, we obtain for $d = 3$

$$\ln D(V) \sim -\frac{1}{6\pi}\int \mathrm{d}^3 x\, V^{3/2}(x)\,. \tag{42.41}$$

For $d = 4$ a counterterm quadratic in $V(x)$ is required. Then, we find

$$\ln D(V) \sim \frac{1}{32\pi^2}\int \mathrm{d}^4 x\, V^2(x)\ln V(x)\,. \tag{42.42}$$

The fermion determinant. The same method can be used to evaluate the fermion determinant

$$D(g) \equiv \det\left(\not{\partial} + M + \sqrt{g}\phi(x)\right)\left(\not{\partial} + M\right)^{-1}. \tag{42.43}$$

Using reflection symmetry we first rewrite $D(g)$:

$$\ln D(g) = \tfrac{1}{2}\operatorname{tr}\ln\left[\not{\partial} + M + \sqrt{g}\phi(x)\right]\left[-\not{\partial} + M + \sqrt{g}\phi(x)\right]\left(-\partial^2 + M^2\right)^{-1},$$

(the order of factors does not matter in a determinant). Evaluating the product, we find

$$\ln D(g) = \tfrac{1}{2}\operatorname{tr}\ln\left[-\partial^2 + \left(M + \sqrt{g}\phi\right)^2 + \sqrt{g}\not{\partial}\phi\right]\left(-\partial^2 + M^2\right)^{-1}.$$

We then again use identity (42.35) with now

$$H = -\partial^2 + \left(M + \sqrt{g}\phi\right)^2 + \sqrt{g}\not{\partial}\phi\,, \tag{42.44}$$

$$H_0 = -\partial^2 + M^2. \tag{42.45}$$

For g large and $\phi(x)$ smooth enough, we can neglect the term $\sqrt{g}\not{\partial}\phi$ in (42.44). Then, the trace over γ matrices yields a factor $N = \operatorname{tr}\mathbf{1}$ and the remaining part of the calculation is identical to the case of the Schrödinger equation with the potential

$$V(x) = g\phi^2(x) + 2\sqrt{g}M\phi(x)\,. \tag{42.46}$$

Substituting this expression into equation (42.39), we obtain the large field behaviour:

$$\ln D(g) \sim -\frac{N}{2}\frac{\Gamma(-d/2)}{(4\pi)^{d/2}}\int \mathrm{d}^d x\left[\left(M + \sqrt{g}\phi(x)\right)^d - M^d\right]. \tag{42.47}$$

42.4.3 The large order behaviour

We can now study the essential singularity of the theory at $g = 0$ small from the properties of the effective local action:

$$\mathcal{S}_{\text{eff}}(\phi) = \int \mathrm{d}^d x \left[\frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2} m^2 \phi^2 + \frac{N}{2}\frac{\Gamma(-d/2)}{(4\pi)^{d/2}} g^{d/2} \phi^d(x)\right]. \tag{42.48}$$

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 the 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 to factorize the dependence on g in front of the classical action:

$$\phi(x) \mapsto \phi(x) g^{-d/2(d-2)}. \tag{42.49}$$

The classical action calculated for a solution takes thus the form

$$\mathcal{S}(\phi_c) = (A/g)^{d/d-2}, \tag{42.50}$$

where A does not depend on g. Introducing this form into the Cauchy representation, we find

$$\mathcal{Z}_k \underset{k\to\infty}{\sim} \int_0 \frac{\mathrm{e}^{-(A/g)^{d/d-2}}}{g^{k-1}} \mathrm{d}g\,. \tag{42.51}$$

The integration yields the large order estimate

$$\mathcal{Z}_k \sim A^{-k} \Gamma\left[k(d-2)/d\right]. \tag{42.52}$$

We observe that, as expected, this theory is less divergent than a purely bosonic field theory. The boson result is recovered (in a cut-off field theory) for d large, because the Pauli principle becomes decreasingly when the dimension increases. For $d = 2$, the expression (42.52) becomes

$$\mathcal{Z}_k \sim A^{-k} (\ln k)^k, \tag{42.53}$$

in agreement with rigorous bounds that yield

$$|\mathcal{Z}_k| < (k!\,)^{\varepsilon} \quad \text{for all } \varepsilon > 0\,. \tag{42.54}$$

Remark. To compare the contributions of boson and fermion interactions, we have implicitly assumed a loop expansions. 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 (42.32) 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.

42.4.4 The QED example

The potentially most interesting application of the preceding analysis is QED. The action has formally the same structure, but one additional complication then arises. The fermion integration yields the determinant

$$D(e) = \det\left(\partial\!\!\!/ + m + ie A\!\!\!/\right). \tag{42.55}$$

To estimate $D(e)$ for large charge e we can use the equation (we assume d even see, for example, Appendix A18.3)

$$D^2(e) = \det\left(m^2 - \mathrm{D}_\mu^2 - \tfrac{1}{2}e\sigma_{\mu\nu}F_{\mu\nu}\right). \tag{42.56}$$

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:

$$\ln D(e) \sim \tfrac{1}{2}N_d \operatorname{tr}\ln\left(m^2 - \mathrm{D}_\mu^2\right). \tag{42.57}$$

We then use the representation (42.35) to evaluate the determinant of the electromagnetic Schrödinger operator. The path integral representation of e^{-tA} has been given in Section 3.2. The determination of the large coupling constant behaviour is, however, more subtle than before. The electromagnetic term in the path integral depends only on the geometry of the loop one integrates along and not on the time spent on the loop trajectory (see equation (3.18)). Therefore, the large coupling constant limit does not select the short time contributions in the representation (42.35) and the determinant no longer leads to a local effective interaction. A direct calculation of the determinant has not been performed. The difficulty seems to be related to the fact that, due to gauge invariance, the gauge degree of freedom of the gauge field cannot be considered as slowly varying. It has, therefore, been conjectured, on the basis of studying the determinant for special gauge fields, that the determinant is equivalent for large e to

$$\ln D(e) \sim C(d)\int \mathrm{d}^d x\, \left|e[A_\mathrm{T}]_\mu\right|^d ; \quad C^{-1}(d) = d(4\pi)^{(d-1)/2}\Gamma\left((d+1)/2\right),$$

where $[A_\mathrm{T}]_\mu$ is the transverse part of A_μ:

$$[A_\mathrm{T}]_\mu(x) = A_\mu(x) - \partial^{-2}\partial_\mu\partial_\nu A_\nu(x).$$

This result is gauge invariant, as it should, but non-local except in the gauge $\partial_\mu A_\mu = 0$. It agrees for $d = 2$ with the exact result (32.60) 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 by Balian *et al.*, or Bogolmony and Fateev. Taking the minimal action solution one obtains an evaluation of the form

$$\mathcal{Z}_k \sim (-1)^k A^{-k}\Gamma(k/2), \quad A = 4.886, \tag{42.58}$$

the expansion parameter being $\alpha = e^2/4\pi$. It is worth mentioning that this evaluation is probably not very useful as a practical mean 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 well with numerical estimates.

42.4.4 The QED example

The potentially most interesting application of the preceding analysis is QED. The action has formally the same structure, but one additional complication then arises. The fermion integration yields the determinant

$$D(e) = \det\left(\partial\!\!\!/ + m + ie A\!\!\!/\right). \tag{42.55}$$

To estimate $D(e)$ for large charge e we can use the equation (we assume d even see, for example, Appendix A18.3)

$$D^2(e) = \det\left(m^2 - \mathrm{D}_\mu^2 - \tfrac{1}{2} e\sigma_{\mu\nu} F_{\mu\nu}\right). \tag{42.56}$$

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:

$$\ln D(e) \sim \tfrac{1}{2} N_d \operatorname{tr} \ln\left(m^2 - \mathrm{D}_\mu^2\right). \tag{42.57}$$

We then use the representation (42.35) to evaluate the determinant of the electromagnetic Schrödinger operator. The path integral representation of e^{-tA} has been given in Section 3.2. The determination of the large coupling constant behaviour is, however, more subtle than before. The electromagnetic term in the path integral depends only on the geometry of the loop one integrates along and not on the time spent on the loop trajectory (see equation (3.18)). Therefore, the large coupling constant limit does not select the short time contributions in the representation (42.35) and the determinant no longer leads to a local effective interaction. A direct calculation of the determinant has not been performed. The difficulty seems to be related to the fact that, due to gauge invariance, the gauge degree of freedom of the gauge field cannot be considered as slowly varying. It has, therefore, been conjectured, on the basis of studying the determinant for special gauge fields, that the determinant is equivalent for large e to

$$\ln D(e) \sim C(d) \int \mathrm{d}^d x \left|e[A_\mathrm{T}]_\mu\right|^d; \quad C^{-1}(d) = d(4\pi)^{(d-1)/2}\Gamma\left((d+1)/2\right),$$

where $[A_\mathrm{T}]_\mu$ is the transverse part of A_μ:

$$[A_\mathrm{T}]_\mu(x) = A_\mu(x) - \partial^{-2}\partial_\mu\partial_\nu A_\nu(x).$$

This result is gauge invariant, as it should, but non-local except in the gauge $\partial_\mu A_\mu = 0$. It agrees for $d = 2$ with the exact result (32.60) 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 by Balian *et al.*, or Bogolmony and Fateev. Taking the minimal action solution one obtains an evaluation of the form

$$\mathcal{Z}_k \sim (-1)^k A^{-k} \Gamma(k/2), \quad A = 4.886, \tag{42.58}$$

the expansion parameter being $\alpha = e^2/4\pi$. It is worth mentioning that this evaluation is probably not very useful as a practical mean 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 well with numerical estimates.

42.5 Divergent Series, Borel Summability

Asymptotic series. Let us consider a function $f(z)$, analytic in a sector S:

$$|\operatorname{Arg} z| \le \alpha/2\,, \qquad |z| \le |z_0|. \tag{42.59}$$

We assume that $f(z)$ has in S the following asymptotic expansion:

$$f(z) = \sum_0^{\infty} f_k z^k. \tag{42.60}$$

This means that the series (42.60) diverges for all non-trivial values of z and that in S the following bound is satisfied:

$$\left| f(z) - \sum_{k=0}^{N} f_k z^k \right| \le C_{N+1} |z|^{N+1} \quad \text{for all } N\,, \tag{42.61}$$

in which

$$C_N |z|^N \underset{\text{for } N \to +\infty}{\to} \infty \quad \forall\, z \ne 0\,.$$

Though the series (42.60) diverges, it is possible to use it to estimate the function $f(z)$ for $|z|$ small. At $|z|$ fixed, we can look for a minimum in the bound (42.61) when N varies. If $|z|$ is small enough, the bound first decreases with N and then, since the series is divergent, finally increases. If we truncate the series at the minimum, we get the best possible estimate of $f(z)$, with a finite error $\varepsilon(z)$. Let us assume for definiteness that the coefficients C_N have the form

$$C_N = M\, A^{-N} (N!)^{\beta}. \tag{42.62}$$

We can then estimate $\varepsilon(z)$ explicitly and find

$$\varepsilon(z) = \min_{\{N\}} C_N |z|^N \sim \exp\left[-\beta (A/|z|)^{1/\beta} \right]. \tag{42.63}$$

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 (42.59) and smaller than $\varepsilon(z)$ in the whole sector. The new function still satisfies the condition (42.61). However, there is one situation in which the asymptotic series defines a unique function. If the angle α satisfies: $\alpha \ge \pi\beta$, then a classical theorem about analytic functions tells us that a function analytic in the sector and bounded by $\varepsilon(z)$ in the whole sector vanishes identically. In particular, in the special case $\beta = 1$, which is typical for perturbative expansions, we find

$$\alpha \ge \pi\,. \tag{42.64}$$

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 specialize from now on to $\beta = 1$, since it is the most useful example, but the generalization to arbitrary β is straightforward. Under the condition (42.64), the function $f(z)$ is uniquely defined by the series. In addition, there then exist

42.5 Divergent Series, Borel Summability

Asymptotic series. Let us consider a function $f(z)$, analytic in a sector S:

$$|\operatorname{Arg} z| \le \alpha/2\,, \qquad |z| \le |z_0|. \tag{42.59}$$

We assume that $f(z)$ has in S the following asymptotic expansion:

$$f(z) = \sum_0^\infty f_k z^k. \tag{42.60}$$

This means that the series (42.60) diverges for all non-trivial values of z and that in S the following bound is satisfied:

$$\left| f(z) - \sum_{k=0}^{N} f_k z^k \right| \le C_{N+1} |z|^{N+1} \quad \text{for all } N\,, \tag{42.61}$$

in which

$$C_N |z|^N \underset{\text{for } N\to+\infty}{\to} \infty \quad \forall\, z \neq 0\,.$$

Though the series (42.60) diverges, it is possible to use it to estimate the function $f(z)$ for $|z|$ small. At $|z|$ fixed, we can look for a minimum in the bound (42.61) when N varies. If $|z|$ is small enough, the bound first decreases with N and then, since the series is divergent, finally increases. If we truncate the series at the minimum, we get the best possible estimate of $f(z)$, with a finite error $\varepsilon(z)$. Let us assume for definiteness that the coefficients C_N have the form

$$C_N = M\, A^{-N} (N!)^\beta. \tag{42.62}$$

We can then estimate $\varepsilon(z)$ explicitly and find

$$\varepsilon(z) = \min_{\{N\}} C_N |z|^N \sim \exp\left[-\beta (A/|z|)^{1/\beta}\right]. \tag{42.63}$$

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 (42.59) and smaller than $\varepsilon(z)$ in the whole sector. The new function still satisfies the condition (42.61). However, there is one situation in which the asymptotic series defines a unique function. If the angle α satisfies: $\alpha \ge \pi\beta$, then a classical theorem about analytic functions tells us that a function analytic in the sector and bounded by $\varepsilon(z)$ in the whole sector vanishes identically. In particular, in the special case $\beta = 1$, which is typical for perturbative expansions, we find

$$\alpha \ge \pi\,. \tag{42.64}$$

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 specialize from now on to $\beta = 1$, since it is the most useful example, but the generalization to arbitrary β is straightforward. Under the condition (42.64), the function $f(z)$ is uniquely defined by the series. In addition, there then exist methods to "sum" the series, which means that one can reconstruct the function from the knowledge of the terms of the series. One set of methods is based upon the Borel transformation.

The Borel transform $B_f(z)$ of $f(z)$ is defined by

$$B_f(z) = \sum_0^\infty B_k z^k \equiv \sum_0^\infty \frac{f_k}{k!} z^k. \tag{42.65}$$

The bound (42.61) and the estimate (42.62) give us a bound on the coefficients f_k:

$$|f_k/k!| < M\, A^{-k}. \tag{42.66}$$

Thus $B_f(z)$ is analytic at least in a circle of radius A and uniquely defined by the series. Furthermore, in the sense of power series

$$f(z) = \int_0^\infty \mathrm{e}^{-t}\, B_f(zt)\mathrm{d}t\,. \tag{42.67}$$

As a consequence of the inequality (42.64), it can be shown that $B_f(z)$ is also analytic in a sector

$$|\operatorname{Arg} z| \in [0, \tfrac{1}{2}(\alpha - \pi)[\,, \tag{42.68}$$

and does not increase faster than an exponential in the sector, so that integral (42.67) converges for $|z|$ small enough and inside the sector

$$|\operatorname{Arg} z| < \alpha/2\,.$$

In addition, it can be shown that the r.h.s. of equation (42.67) satisfies a bound of type (42.61). Hence, this integral representation yields the unique function which has the asymptotic expansion (42.60) in the domain (42.59).

42.6 Large Order Behaviour and Borel Summability

We have learned that, for a large class of potentials in quantum mechanics and for a number of field theories, instanton contributions for small values of the loop expansion parameter g behave like

$$C g^{-b}\, \mathrm{e}^{-a/g}\,. \tag{42.69}$$

The corresponding contribution to the perturbative coefficients for large order k of the loop expansion is then,

$$(C/\pi) k^{b-1} a^k k!\,. \tag{42.70}$$

Therefore, the coefficients B_k of the Borel transform $B(z)$ (equation (42.65)) behave as

$$B_k \sim (C/\pi) k^{b-1} a^k. \tag{42.71}$$

This asymptotic estimate tells us that the singularity of $B(z)$ closest to the origin is located at the point $z = 1/a$. More precisely, the Borel transform $B(z)$ has an algebraic singularity of the form

$$\frac{C}{\pi}\int_0 \frac{\mathrm{d}g\, \mathrm{e}^{-a/g}}{g^{b+1}} \sum_k \frac{1}{k!}\left(\frac{z}{g}\right)^k = \frac{C}{\pi}\int_0 \frac{\mathrm{d}g\, \mathrm{e}^{-(a-z)/g}}{g^{b+1}} = (C/\pi)\Gamma(b)(a-z)^{-b}.$$

Therefore, the integral (42.67) does not exist if the classical action $A = 1/a$ is positive. The perturbation series in such theories is not Borel summable. Let us, in the light of this result, discuss the various situations we have encountered:

(i) The field equations have no real instanton solutions. This is, in particular, the case if we have 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. It is only a presumption because various features of the perturbative expansion, invisible at large orders, could prevent Borel summability. The perturbative expansion could contain for instance contributions all of the same sign, growing faster than any exponential of the order k, but much smaller than $k!$ (for example $\sqrt{k!}$). 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) We have found real instantons in the theory because we expanded around a relative minimum of the potential: the perturbative expansion is not Borel summable.

However, in this case, we can provide one additional piece of information useful for determining the solution: the unstable situation can be considered as coming from a stable situation by analytic continuation. Therefore, a possible solution could be to integrate in the Borel transform just above the cut which is on the real positive axis. As a consequence, from a real perturbative expansion we would obtain a complex result, but this is exactly what we expect. It is easy to verify that the imaginary part is for g small exactly what we 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) There are real instantons connecting degenerate classical minima.

The theory is not Borel summable. Integration above or below the axis yields a complex result for a real quantity. This cannot be the correct prescription. The half sum of the integral above and below is real, but even in the simple example of the double well-potential, one can verify numerically, and argue analytically, that it is not the correct solution. We show in Chapter 43, for one-dimensional potentials, that the additional information needed to determine the sum of the perturbative expansion is provided by the consideration of many instanton contributions . The corresponding problem has not been solved in field theory examples yet.

Remarks. We have given field theory examples of such a situation in Sections 41.5, 41.6: the two-dimensional $CP(N-1)$ models and four-dimensional $SU(2)$ gauge theory. In these models real instantons connect degenerate minima of the classical action and the corresponding classical action is positive. Therefore, the coefficients of the perturbative expansion contain a non-Borel summable contribution. This contribution does not necessarily dominate the large order behaviour, because, as the example of of the ϕ^4_4 massless field theory (see Section 42.3.2) illustrates, when a field theory is classically scale invariant, the perturbative expansion might be dominated by contributions unobtainable by semi-classical methods, and related to the UV or IR singularities.

42.7 Practical Summation Methods

Various practical summation methods rely upon a Borel transformation.

The Borel transformation reduces the problem of determining the function to the analytic continuation of the Borel transform. The Borel transform is given by a Taylor series in a circle and an analytic continuation of the series in a neighbourhood of the real positive axis is required. This analytic continuation can be performed by many methods and the optimal choice depends somewhat on the additional information one possesses about the function. Let us give two examples:

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. From the series, one derives Padé approximants which are rational functions P_M/Q_N satisfying

$$B_f(z) = \frac{P_M(z)}{Q_N(z)} + O\left(z^{N+M+1}\right), \tag{42.72}$$

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 $[M,N]$ Padé approximants with $N+M \le K$. This method is well adapted to meromorphic functions. The main disadvantage of the method is that for a rather general class of functions Padé approximants are known to converge only in measure and thus spurious poles may occasionally appear close to or on the real positive axis.

Even if Padé approximants converge, this property may be the source of some instabilities in the results, and, therefore, make the empirical evaluation of errors difficult.

Conformal mapping. If we know the domain of analyticity of the Borel transform, we can find a mapping which preserves the origin, and maps the domain of analyticity onto a circle. In the transformed variable, the new series converges in the whole domain of analyticity. Let us explain the method on an example.

If the Borel transform is analytic in a cut-plane, the cut running along the real negative axis from $-\infty$ to $-1/a$, we can map the cut-plane onto a circle of radius 1:

$$z \mapsto u, \quad u(z) = \frac{\sqrt{1+az}-1}{\sqrt{1+az}+1}. \tag{42.73}$$

From the original series for the Borel transform, we derive a series in powers of the new variable u:

$$B_f(z) = \sum \frac{f_k}{k!} z^k, \qquad B_f[z(u)] = \sum_0^\infty B_k u^k. \tag{42.74}$$

Introducing this expansion in the Borel transformation, we obtain a new expansion for $f(z)$,

$$f(z) = \sum_0^\infty B_k I_k(z)\,, \tag{42.75}$$

in which the functions $I_k(z)$ have the integral representation:

$$I_k(z) = \int_0^\infty \mathrm{e}^{-t}\,[u(zt)]^k\,\mathrm{d}t\,. \tag{42.76}$$

It is possible to study the natural domain of convergence of this new expansion. Using for $u(z)$ the explicit expression (42.73), we evaluate $I_k(z)$ for k large by the steepest descent method. The saddle point equation is

$$-1 + \frac{k}{t}\frac{1}{\sqrt{1+azt}} = 0\,, \tag{42.77}$$

which for k large yields

$$t \sim k^{2/3}/(az)^{1/3}. \tag{42.78}$$

It follows that $I_k(z)$ behaves for k large as

$$I_k(z) \sim \exp\left[-3k^{2/3}/(az)^{1/3}\right]. \tag{42.79}$$

Three situations can now arise:

(i) The coefficients B_k either decrease or at least do not grow too rapidly,

$$|B_k| < M\,\mathrm{e}^{\varepsilon k^{2/3}} \quad \text{for all } \varepsilon > 0\,.$$

Then, the expansion (42.75) converges at least in the region

$$\operatorname{Re} z^{-1/3} > 0 \;\Rightarrow\; |\mathrm{Arg}| < 3\pi/2\,. \tag{42.80}$$

This, in particular, implies that the function $f(z)$ must be analytic in the corresponding region which contains a part of the second sheet.

(ii) The coefficients behave like

$$B_k \sim \exp\left(ck^{2/3}\right) \quad \text{for } k \text{ large}. \tag{42.81}$$

The domain of convergence is, then,

$$\operatorname{Re} z^{-1/3} > \tfrac{1}{3} c a^{1/3}. \tag{42.82}$$

This condition implies analyticity in a finite domain containing a part of the second sheet since for $|z|$ small, the r.h.s. 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 (42.74) 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 should map a smaller part of the domain of analyticity onto a circle.

Application to the calculation of critical exponents. In the summation method based on Borel transformation and mapping it is easy to incorporate the information coming from the large order behaviour analysis. This is one reason why it has been used quite systematically in the framework of the ϕ^4 field theory to calculate critical exponents and other universal quantities. In Chapter 29, we have given values for various critical exponents, obtained by applying variants of the Borel summation method to the known terms of the perturbative expansion, that is, six successive terms in fixed dimension 3 and up to order ε^5 for the ε-expansion.

Let us now summarize the information available for in the ϕ^4 field theory that justifies the use of this summation method.

(i) The Borel summability of perturbation theory in ϕ^4_2 and ϕ^4_3 has been rigorously established.

(ii) The large order behaviour has been determined in all cases and compares favourably with the first terms of the series available (see Section 42.2).

(iii) Since all known instanton solutions in the ϕ^4 theory give negative actions, it is plausible that the Borel transform is analytic in a cut-plane, the location and nature of the singularity closest to the origin being given by the large order estimates.

Consequently, the methods based upon a Borel transformation and a conformal mapping of the cut-plane onto a circle, have appeared as excellent candidates to sum the perturbation series and the ε-expansion.

For completeness, let us finally give at least one example of a summation method not based on a Borel transformation.

Order dependent mappings (ODM). The ODM method requires, to be applicable, some knowledge of the analyticity properties of the function itself. As we have discussed, the series diverge because the function has singularities accumulating at the origin. However, the strengths of the singularities have to decrease fast enough for the function to have a series expansion. In the examples we have met, the discontinuity of the function decreases exponentially near the origin. The idea is then to pretend that the function is analytic, in addition to its true domain of analyticity, in a small circle 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 our problem would be solved. Since the original series is in fact only asymptotic, the series in the transformed variable is also asymptotic. However, as a result of this transformation, the coefficients of the new series now depend on an adjustable parameter ρ.

Let us assume for instance that $f(z)$ is analytic in a cut-plane. We then use the mapping

$$z = 4\rho u/(1-u)^2 . \tag{42.83}$$

The transformed series has the form

$$f\bigl(z(u)\bigr) = \sum_0^{\infty} P_k(\rho) u^k , \tag{42.84}$$

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

$$z = \rho\, h(u) , \qquad h(u) = O(u) . \tag{42.85}$$

The kth 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 roughly 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 the modulus of the last term taken into account is minimal. By introducing an additional parameter, one modifies the situation: one can choose first the order of truncation and then try to adjust the parameter ρ in such a way that, at z again fixed, the last term taken into account is minimal.

The kth order approximant has the form

$$\{f(z)\}_k = \sum_{l=0}^{k} P_l(\rho_k) \left[u(z)\right]^l , \qquad P_k(\rho_k) = 0 . \tag{42.86}$$

It can be shown under some conditions that if the terms f_k of the original series grow like $(k!)^\beta$ then the sequence ρ_k decreases like $1/k^\beta$. Such a method has been successfully applied to test problems like the quartic anharmonic oscillator with a mapping

$$z = \rho u/(1-u)^{3/2},$$

and to one physical example, the hydrogen atom in a strong magnetic field.

Bibliographical Notes

The connection between barrier penetration effects in the semi-classical limit and large behaviour in perturbation theory was first realized by

F.J. Dyson, Phys. Rev. 85 (1952) 631; C.M. Bender and T.T. Wu, *Phys. Rev.* 184 (1969) 1231; D7 (1973) 1620; *Phys. Rev. Lett.* 37 (1976) 117; T. Banks, C.M. Bender and T.T. Wu, *Phys. Rev.* D8 (1973) 3346; T. Banks and C.M. Bender, *Phys. Rev.* D8 (1973) 3366;

These authors calculate the barrier penetration coefficients in quantum mechanics using standard WKB methods. The idea of using instead path integrals can be found in an early paper

C.S. Lam, *Nuovo Cimento*, 55A (1968) 258;

The field theory calculation (the ϕ_4^4 theory) is due to Lipatov, already quoted,

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.

The method has been generalized in

E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev.* D15 (1977) 1558,

and then applied to many theories like the abelian gauge field coupled to scalars

C. Itzykson, G. Parisi and J.-B. Zuber, *Phys. Rev. Lett.* 38 (1977) 306; A.P. Bukhvostov and L.N. Lipatov, *Zh. Eksp. Teor. Fiz.* 73 (1977) 1658;

to the ϕ^3 field theory with Potts symmetry,

A. Houghton, J.S. Reeve and D.J. Wallace, *Phys. Rev.* B17 (1978) 2956;

to theories with degenerate minima (see Chapter 43)

E. Brézin, G. Parisi and J. Zinn-Justin, *Phys. Rev.* D16 (1977) 408;

non-abelian gauge theories

E.B. Bogomolny and V.A. Fateyev, *Phys. Lett.* 71B (1977) 93; L.N. Lipatov, A.P. Bukhvostov and E.I. Malkov, *Phys. Rev.* D19 (1979) 2974.

The idea of evaluating the large order behaviour by relating it the discontinuity of functions is due to

G. Parisi, *Phys. Lett.* 66B (1977) 167; E.B. Bogomolny, *Phys. Lett.* 67B (1977) 193.

Let us also quote the review article

E.B. Bogomolny, V.A. Fateyev and L.N. Lipatov, *Sov. Sci. Rev.* section A: *Phys. Rev.* 2 (1980) 247, I.M. Khalatnikov ed.

and the reprint volume

Large Order Behaviour of Perturbation Theory, Current Physics vol. 7, J.C. Le Guillou and J. Zinn-Justin eds. (North-Holland, Amsterdam 1990).

Large order in field theories with fermions has first studied by

G. Parisi, *Phys. Lett.* 66B (1977) 382.

QED has been considered in

C. Itzykson, G. Parisi and J.-B. Zuber, *Phys. Rev.* D16 (1977) 996; R. Balian, C. Itzykson, J.-B. Zuber and G. Parisi, *Phys. Rev.* D17 (1978) 1041; E.B. Bogomolny and

V.A. Fateyev, *Phys. Lett.* 76B (1978) 210; E.B. Bogomolny and Yu A. Kubyshin, *Yad. Fiz.* 34 (1981) 1535; 35 (1982) 202 (*Sov. J. Nucl. Phys.* 34 (1981) 853; 35 (1982) 114).

The Yukawa interaction in two dimensions has been discussed in

M.P. Fry, *Phys. Lett.* 80B (1978) 65.

The mathematical discussion of divergent series can be found in

E. Borel, *Leçon sur les Séries Divergentes*, 2nd edn (Gauthier-Villars, Paris 1928); G. Hardy, *Divergent Series* (Oxford University Press, Oxford 1949).

For a textbook concerning Padé approximants, see

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).

Summation of series by Borel transformation and mapping was proposed by

J.J. Loeffel, *Saclay Report*, DPh-T/76/20 unpublished.

The summation method based on order-dependent mappings has been studied in

R. Seznec and J. Zinn-Justin, *J. Math. Phys.* 20 (1979) 1398; J.C. Le Guillou and J. Zinn-Justin, *Ann. Phys. (NY)* 147 (1983) 57,

and proofs of convergence in special cases given in

R. Guida, K. Konishi and H. Suzuki, *Ann. Phys. (NY)* 241 (1995) 152.

The Borel summability of the ϕ^4 in two and three dimensions has been established in

J.P. Eckmann, J. Magnen and R. Sénéor, *Commun. Math. Phys.* 39 (1975) 251; J.S. Feldman and K. Osterwalder, *Ann. Phys. (NY)* 97 (1976) 80; J. Magnen and R. Sénéor, *Commun. Math. Phys.* 56 (1977) 237; J.-P. Eckmann and H. Epstein, *Commun. Math. Phys.* 68 (1979) 245.

For applications of differential approximants and references, see

A.J. Guttmann and G.S. Joyce, *J. Phys. A: Math. Gen.* 5 (1972) L81; M.E. Fisher and J.H. Chen, *Phase Transitions*, Cargèse 1980, vol. B72, M. Lévy, J.C. Le Guillou and J. Zinn-Justin eds. (Plenum, New York 1982).

APPENDIX A42

A42.1 Large Order Behaviour for Simple Integrals

Consider the integral of Chapter 39:

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left[-\left(\tfrac{1}{2}x^2 + \tfrac{1}{4}gx^4\right)\right] \mathrm{d}x\,. \tag{A42.1}$$

We can expand $I(g)$ in power series,

$$I(g) = \sum_0^\infty I_k g^k. \tag{A42.2}$$

The coefficients I_k count the number of vacuum Feynman diagrams with the proper weights in a ϕ^4 field theory. We have argued in Section 39.2 that the imaginary part of $I(g)$ for g negative and small was dominated by the non-trivial saddle points:

$$x^2 = -1/g\,. \tag{A42.3}$$

The contribution of the saddle points yields

$$\operatorname{Im} I(g) \underset{g\to 0_-}{\sim} 2^{-1/2}\, \mathrm{e}^{1/4g}\,. \tag{A42.4}$$

Therefore, I_k behaves for k large as

$$I_k \underset{k\to\infty}{\sim} \frac{1}{\pi\sqrt{2}}(-4)^k (k-1)!\,. \tag{A42.5}$$

This result leads to the following interpretation of the large order behaviour formulae obtained in Chapter 42: 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 $(4A)^{-k}$ where k is the order but also, up to an additive constant, the number of loops and A the classical action.

A42.2 Non-Loop Expansions

Although we have discussed large order behaviour estimates only for loop expansions, it is possible to generalize the analysis for perturbative expansions in different parameters. For example, let us consider the action

$$\mathcal{S}(q) = \int \left[\tfrac{1}{2}\dot{q}^2 + \tfrac{1}{2}q^2 + \lambda V(q)\right] \mathrm{d}t\,, \tag{A42.6}$$

in which $V(q)$ is a polynomial

$$V(q) = \sum_2^{2N} V_n q^n, \tag{A42.7}$$

and we are interested in the large order behaviour in powers of λ.

The divergent large order behaviour is a consequence of the infinite range of the q integration. Therefore, it is dominated by the large q behaviour of V and is related to the instanton solutions of an action in which $V(q)$ has been replaced by its term of highest degree. Then, by rescaling, we see that the classical solution $q_c(t)$ has the form

$$q_c(t) = \lambda^{-1/(2N-2)} f(t) \,. \tag{A42.8}$$

The term of degree n in $V(q)$ then gives a contribution to the classical action proportional to $\lambda^{1-n/(2N-2)}$. For $\lambda \to 0$, 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 $n \geq 2N-2$.

A42.3 Linear Differential Approximants

Padé approximants provide the simplest example of a general class of approximants, which are obtained as solutions to equations (algebraic or differential) with polynomial coefficients. 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 let us give the example of the linear differential approximants.

Let $f(z)$ be a function for which we know a power series expansion. We can construct approximants $\bar{f}(z)$ to this function by looking for solutions of the differential equation

$$\sum_{n=0}^{N} P_n(z) \left(\frac{\partial}{\partial z}\right)^n \bar{f}(z) = R(z), \tag{A42.9}$$

in which the polynomials $P_n(z)$ and $R(z)$ form a set of polynomials of the lowest possible total degree chosen such that

$$f(z) - \bar{f}(z) = O\left(z^k\right). \tag{A42.10}$$

In the generic situation, the degrees $[P_n]$ and $[R]$ of the polynomials P_n and R satisfy

$$\sum_{n=0}^{N} [P_n] + [R] = k \,. \tag{A42.11}$$

The advantage of these kinds of approximants is that they are extremely flexible. It is possible to use a lot of additional information one possesses 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. There is, of course, a price to pay: this approximation is much more unstable. It is necessary to select among the large number of approximants one can construct, those for which one has some reasons to believe that they are especially well adapted to the original function one wants to approximate.

Due to the generality of the problem, a systematic study of this class of approximants is lacking. Notice that the method can be generalized to power series in more than one variable. One then writes partial differential equations with polynomial coefficients in several variables.

43 MULTI-INSTANTONS IN QUANTUM MECHANICS

The concepts of multi-instanton configurations and contributions are non-trivial because classical equations are non-linear and thus in general a linear combination of solutions is not a solution. However, in the limit in which all instantons are largely separated, such configurations render the action almost stationary because each instanton solution differs from a constant solution only by exponentially small corrections at large distances (in a field theory this is only true if the theory is massive). We here examine in the context of Quantum Mechanics the significance of such multi-instanton ***quasi-solutions***. The generalization to Field Theory, however, is non-trivial and has still to be worked out.

We have found in preceding chapters several situations in which multi-instantons can be expected to play a role. In the case in which instantons are found, when calculating the contribution to $\operatorname{tr} e^{-\beta H}$ at finite β we always have kept only the solution which 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 infinite β limit, these configurations have the properties we expect from a n-instanton. However, there is a subtlety: naively, we expect these configurations to give a contribution 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

$$E_0 = E_0^{(0)} + E_0^{(1)} + \cdots ,$$

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

$$\operatorname{tr} e^{-\beta H} \sim e^{-\beta E_0} \sim e^{-\beta E_0^{(0)}} \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \left(E_0^{(1)}\right)^n . \tag{43.1}$$

Thus, the existence of an 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 the estimation of the large order behaviour of perturbation theory for potentials with degenerate minima. We have seen that when we start from a situation in which the minima are almost degenerate, we obtain in the degenerate limit a typical two-instanton contribution but with an infinite multiplicative coefficient. This divergence has the following interpretation: In the degenerate limit, the classical solution decomposes into the superposition of an instanton and an anti-instanton infinitely separated and thus fluctuations which tend to change the distance between the instanton and the anti-instanton induce a vanishingly small variation of the action. It follows that, to properly study 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, we know that n-instanton contributions do exist. However, these contributions do not in general correspond to solutions of the classical equation of motion. They correspond to configurations of largely separated instantons connected in a way which has to be examined, and become solutions of the equation of motion only asymptotically, in the limit of infinite separation. These configurations depend on n times more collective coordinates than the one-instanton configuration.

In Sections 43.1,43.2, we first study explicitly two examples which we have already considered in Chapter 41: the *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. Finally, from the results obtained for the many instanton contributions, we are led to conjecture the exact form of the semi-classical expansion for potentials with degenerate minima, generalizing the exact Bohr–Sommerfeld quantization condition.

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.

43.1 The Double-Well Potential

We first consider the hamiltonian of the double-well potential, already discussed in Section 41.1:

$$H = -\tfrac{1}{2}\mathrm{d}_q^2 + V(q\sqrt{g})/g, \quad V(q) = \tfrac{1}{2}q^2(1-q)^2. \tag{43.2}$$

In the infinite β limit, the instanton solutions are (equation (41.10))

$$q_\pm(t) = \frac{1}{\sqrt{g}} f\bigl(\mp(t-t_0)\bigr), \tag{43.3}$$

$$f(t) = 1/(1+\mathrm{e}^t) = 1 - f(-t)\,, \tag{43.4}$$

where the constant t_0 characterizes the instanton position.

43.1.1 The two-instanton configuration

We first construct the two-instanton configuration (really an instanton–anti-instanton). We look for a configuration depending on an additional time parameter, the separation between instantons, which in the limit of infinite separation decomposes into two instantons, and which for large separations minimizes the variation of the action (figure 43.1). 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 (for details, see Sections 39.4.1 and A43.2). We use instead a method which, at least at leading order, is simpler and shows that the result is universal.

We consider a configuration $q_c(t)$ which is the sum of instantons separated by a distance θ, up to an additive constant adjusted in such a way as to satisfy the boundary conditions. It is convenient to introduce some notation:

$$u(t) = f(t-\theta/2), \tag{43.5}$$

$$\underline{u}(t) = f(-t-\theta/2)\,, \tag{43.6}$$

$$v(t) = 1 - \underline{u}(t) = u(t+\theta)\,, \tag{43.7}$$

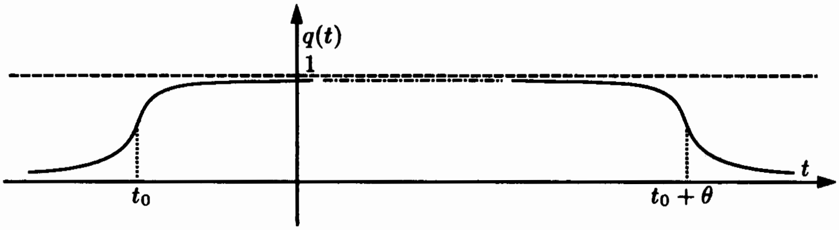


Fig. 43.1 The two-instanton configuration.

where $f(t)$ is the function (43.4). We then take the path

$$q_c(t)\sqrt{g} = u + \underline{u} - 1 = u - v\,. \tag{43.8}$$

(Again time translation $t \mapsto t + t_0$ generates a set of degenerate configurations.) This 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 $\mathrm{e}^{-\theta}$. Although the calculation of the corresponding action is straightforward, we perform it stepwise to show the generality of the ansatz. The action of the path (43.8) is

$$\begin{aligned}\mathcal{S}(q_c) &= \int \mathrm{d}t \left(\frac{1}{2}\dot{q}_c^2 + \frac{1}{g}V(q_c\sqrt{g})\right)\\ &= 2\times\frac{1}{6g} + \frac{1}{g}\int \mathrm{d}t\left[\dot{\underline{u}}\dot{u} + V(\underline{u}+u-1) - V(\underline{u}) - V(u)\right].\end{aligned} \tag{43.9}$$

The parity of q_c allows to restrict the integration to the region $t > 0$, where v is small. After an integration by parts of the term $\dot{u}\dot{\underline{u}} = -\dot{v}\dot{u}$, we find

$$\mathcal{S}(q_c) = \frac{1}{3g} + \frac{2}{g}\left\{v(0)\dot{u}(0) + \int_0^{+\infty}\mathrm{d}t\left[v\ddot{u} + V(u-v) - V(u) - V(v)\right]\right\}. \tag{43.10}$$

We now expand in powers of v, and take into account the equation of motion for u. The leading terms is of order $\mathrm{e}^{-\theta}$ and we thus stop at order v^2. We obtain

$$\mathcal{S}(q_c) = \frac{1}{3g} + \frac{2}{g}v(0)\dot{u}(0) + \frac{2}{g}\left\{\int_0^{+\infty}\mathrm{d}t\left[\tfrac{1}{2}v^2V''(u) - V(v)\right]\right\}. \tag{43.11}$$

The function v decreases exponentially away from the origin so the main contributions to the integral come from the neighbourhood of $t = 0$, where $V''(u) \sim 1$. Moreover, $V(v)$ can be replaced at leading order by $\frac{1}{2}v^2$. Therefore, the two terms cancel and we are left with the integrated contribution

$$v(0)\dot{u}(0) \sim -\mathrm{e}^{-\theta},$$

and thus

$$\mathcal{S}(q_c) = g^{-1}\left[\tfrac{1}{3} - 2\,\mathrm{e}^{-\theta} + O\left(\mathrm{e}^{-2\theta}\right)\right]. \tag{43.12}$$

It will become clearer later why we need the classical action only up to order $\mathrm{e}^{-\theta}$. It is also useful to keep β large but finite in the calculation. Symmetry between θ and $\beta - \theta$ then implies

$$\mathcal{S}(q_c) = g^{-1}\left[\tfrac{1}{3} - 2\,\mathrm{e}^{-\theta} - 2\,\mathrm{e}^{-(\beta-\theta)}\right]. \tag{43.13}$$

As a check, we can calculate the extremum of $\mathcal{S}(q)$ at β fixed and we obtain

$$\theta_c = \beta/2, \Rightarrow \mathcal{S}(q_c) = g^{-1}\left[\tfrac{1}{3} - 4\,\mathrm{e}^{-\beta/2} + O\left(\mathrm{e}^{-\beta}\right)\right].$$

In Chapter 41, for the same hamiltonian, we have found for β large (equation (41.12)):

$$\mathcal{S}(q_c) = g^{-1}\left[\tfrac{1}{6} - 2\,\mathrm{e}^{-\beta} + O\left(\mathrm{e}^{-2\beta}\right)\right]. \tag{43.14}$$

Both results are consistent. Indeed to compare them, we have to replace β by $\beta/2$ in equation (43.14) 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 $\mathrm{e}^{-\theta}$ and the variation of the action of order $\mathrm{e}^{-2\theta}$ at least. Setting

$$q(t) = q_c(t) + r(t), \tag{43.15}$$

we find, expanding the action up to second order in $r(t)$,

$$\begin{aligned}\mathcal{S}(q_c + r) = \mathcal{S}(q_c) &+ \int \left[\dot q_c(t)\dot r(t) + \frac{1}{\sqrt{g}} V'\left(q_c(t)\sqrt{g}\right) r(t)\right] \mathrm{d}t \\ &+ \frac{1}{2}\int \mathrm{d}t \left[\dot r^2(t) + V''(q_c\sqrt{g}) r^2(t)\right] + O\left([r(t)]^3\right).\end{aligned} \tag{43.16}$$

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 quadratic in $r(t)$, we replace V'' by one, 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 $\mathrm{e}^{-\theta}$ while the quadratic term is of order 1. Shifting r to eliminate the linear term would then give a contribution of order $\mathrm{e}^{-2\theta}$ and thus negligible at the order we consider.

43.1.2 The n-instanton configuration

We now consider a succession of n instantons, separated by times θ_i,

$$\textstyle\sum_{i=1}^{n} \theta_i = \beta. \tag{43.17}$$

At leading order, we need consider "interactions" only between nearest neighbours. This is an essential simplifying feature of one-dimensional quantum mechanics. The classical action $\mathcal{S}_c(\theta_i)$ can then be directly inferred from expression (43.13):

$$\mathcal{S}_c(\theta_i) = \frac{1}{g}\left[\frac{n}{6} - 2\sum_{i=1}^{n} \mathrm{e}^{-\theta_i} + O\left(\mathrm{e}^{-(\theta_i+\theta_j)}\right)\right]. \tag{43.18}$$

Other interactions are negligible because they are of higher order in $\mathrm{e}^{-\theta}$.

Note that for n even the n-instanton configuration contributes to $\operatorname{tr}\mathrm{e}^{-\beta H}$, while for n odd it contributes to $\operatorname{tr} P\,\mathrm{e}^{-\beta H}$ (P is the reflection operator). Then, calculating

$$\mathcal{Z}_\epsilon = \tfrac{1}{2}\operatorname{tr}\left[(1+\epsilon P)\,\mathrm{e}^{-\beta H}\right], \tag{43.19}$$

we obtain for $\epsilon = +1$ and $\epsilon = -1$ contributions to the even and odd eigenstate energies respectively.

Remark. Since we keep in the action all terms of order $\mathrm{e}^{-\beta}$ we expect to find the contributions not only to the two lowest energies but also to all energies which remain finite when g goes to zero (see the remark after equation (39.103) in Section 39.7).

43.1.3 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 configuration is not a solution of the equation of motion, we have constructed it in such a way that we can neglect the linear terms in the gaussian integration. The second derivative of the action at the classical path

$$M(t',t) = \left[-\mathrm{d}_t^2 + V''\left(\sqrt{g}q_c(t)\right)\right]\delta(t-t') \tag{43.20}$$

has the form of a quantum hamiltonian with a potential which consists of n wells almost identical to the well arising in the one-instanton problem, and which are largely separated. At leading order, the corresponding spectrum is, therefore, the spectrum arising in the one-instanton problem n-times degenerate. Corrections are exponentially small in the separation (for details see Appendix A43.1). Moreover, by introducing n collective time variables, we suppress n times the zero eigenvalue, and generate the jacobian of the one-instanton case to the power n. We can, therefore, write the n-instanton contribution to $\mathcal{Z}_\epsilon$ (equation (43.19)) as

$$\mathcal{Z}_\epsilon^{(n)} = \mathrm{e}^{-\beta/2}\,\frac{\beta}{n}\left(\frac{\mathrm{e}^{-1/6g}}{\sqrt{\pi g}}\right)^n \int_{\theta_i\ge 0}\delta\left(\sum\theta_i-\beta\right)\prod_i \mathrm{d}\theta_i \exp\left[\frac{2}{g}\sum_{i=1}^n \mathrm{e}^{-\theta_i}\right]. \tag{43.21}$$

All the 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 $\mathrm{e}^{-\beta/2}$ is the usual normalization factor.

We define the quantity λ, the "fugacity" of the instanton gas,

$$\lambda = \frac{\epsilon}{\sqrt{\pi g}}\,\mathrm{e}^{-1/6g}, \tag{43.22}$$

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:

$$\Sigma(\beta,g) = \mathrm{e}^{-\beta/2} + \sum_{n=1}^{\infty}\mathcal{Z}_\epsilon^{(n)}(\beta,g). \tag{43.23}$$

If we neglect the instanton interaction (the dilute gas approximation), we can integrate over the θ_i's and calculate the sum

$$\Sigma(\beta,g) = \mathrm{e}^{-\beta/2}\left[1+\beta\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\frac{\beta^{n-1}}{(n-1)!}\right] = \mathrm{e}^{-\beta(1/2-\lambda)}. \tag{43.24}$$

We recognize the perturbative and one-instanton contribution, at leading order, to $E_\epsilon(g)$, the ground state and the first excited state energies:

$$E_\epsilon(g) = \frac{1}{2} + O(g) - \frac{\epsilon}{\sqrt{\pi g}}\,\mathrm{e}^{-1/6g}\left(1+O(g)\right). \tag{43.25}$$

Discussion. To go beyond the one-instanton approximation, we have to take into account the interaction between instantons. We then face a problem: examining expression

(43.21), we discover that the interaction between instantons is *attractive*. For g small, the dominant contributions to the integral, therefore, come from configurations in which the instantons are close. For such configurations, the concept of instanton is no longer valid, since such configurations cannot be distinguished from the fluctuations around the constant or the one-instanton solution.

We should not be completely surprised by this phenomenon. Indeed, the large order behaviour analysis 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. But if the perturbative expansion is ambiguous at the two-instanton order, we should not expect to be able to calculate a contribution of the same order or smaller. To proceed any further we first have to make a choice about how we define the sum of perturbation theory. It is possible to show that the perturbation series is Borel summable for g negative by relating it to the perturbative expansion of the $O(2)$ anharmonic oscillator. We, therefore, 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 is *repulsive*, and the expression (43.21) becomes meaningful. We, therefore, 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 non-Borel summable and has to be summed by the same procedure.

43.1.4 The calculation

We again write the n-instanton contribution

$$\mathcal{Z}_\epsilon^{(n)} \sim \frac{\beta}{n} \mathrm{e}^{-\beta/2} \lambda^n \int_{\theta_i \geq 0} \delta\left(\sum \theta_i - \beta\right) \prod_{i=1}^n \mathrm{d}\theta_i \exp\left[\frac{2}{g} \sum_{i=1}^n \mathrm{e}^{-\theta_i}\right]. \tag{43.26}$$

To factorize the integral over the variables θ_i, we replace the δ-function by an integral representation:

$$\delta\left(\sum_{i=1}^n \theta_i - \beta\right) = \frac{1}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} \mathrm{d}s \exp\left[-s\left(\beta - \sum_{i=1}^n \theta_i\right)\right], \quad \eta > 0. \tag{43.27}$$

In terms of the function

$$I(s) = \int_0^{+\infty} \mathrm{e}^{s\theta - \mu\,\mathrm{e}^{-\theta}}\, \mathrm{d}\theta\,, \tag{43.28}$$

the integral (43.26) can be rewritten as

$$\mathcal{Z}_\epsilon^{(n)} \sim \frac{\beta\,\mathrm{e}^{-\beta/2}}{2i\pi} \frac{\lambda^n}{n} \int_{-i\infty-\eta}^{i\infty-\eta} \mathrm{d}s\, \mathrm{e}^{-\beta s} \left[I(s)\right]^n, \quad \text{with } \mu = -2/g\,. \tag{43.29}$$

By giving to s a small negative real part, we have ensured the convergence of the integral (43.28). To evaluate the integral (43.28), we set

$$\mu\,\mathrm{e}^{-\theta} = t\,, \tag{43.30}$$

and the integral becomes

$$I(s) = \int_0^{\mu} \frac{\mathrm{d}t}{t} \left(\frac{\mu}{t}\right)^s \mathrm{e}^{-t} = \int_0^{+\infty} \frac{\mathrm{d}t}{t} \left(\frac{\mu}{t}\right)^s \mathrm{e}^{-t} + O\left(\mathrm{e}^{-\mu}/\mu\right), \tag{43.31}$$

for μ positive and large, that is, $g \to 0_-$. Up to an exponentially small correction we thus obtain

$$I(s) \sim \mu^s \Gamma(-s)\,. \tag{43.32}$$

The generating function $\Sigma(\beta, g)$ of the leading order multi-instanton contributions (43.23) then becomes

$$\begin{aligned} \Sigma(\beta, g) &= -\frac{\beta\,\mathrm{e}^{-\beta/2}}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} \mathrm{d}s\, \mathrm{e}^{-\beta s} \sum_n \frac{\lambda^n}{n} \mu^{ns} \left[\Gamma(-s)\right]^n \\ &= -\frac{\beta\,\mathrm{e}^{-\beta/2}}{2i\pi} \int_{-i\infty-\eta}^{i\infty-\eta} \mathrm{d}s\, \mathrm{e}^{-\beta s} \ln\left[1 - \lambda\mu^s \Gamma(-s)\right]. \end{aligned} \tag{43.33}$$

We set

$$E = s + 1/2\,, \qquad \phi(E) = 1 - \lambda \mu^{E-1/2} \Gamma(1/2 - E) \tag{43.34}$$

We then integrate $\beta\,\mathrm{e}^{-\beta s}$ by parts and obtain

$$\Sigma(\beta, g) = -\frac{1}{2i\pi} \int_{-i\infty}^{+i\infty} \mathrm{d}E\, \mathrm{e}^{-\beta E} \frac{\phi'(E)}{\phi(E)}. \tag{43.35}$$

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 $\mathrm{Re}(E) > 0$. Integrating we obtain a sum of residues

$$\Sigma(\beta, g) = \sum_{N \geq 0} \mathrm{e}^{-\beta E_N}, \tag{43.36}$$

where the energies E_N are solutions of the equation

$$\phi(E) = 1 - \lambda \mu^{E-1/2} \Gamma(1/2 - E) = 0\,. \tag{43.37}$$

Since λ is small, a zero E of this equation is close to a pole of $\Gamma(1/2 - E)$:

$$E_N = N + \tfrac{1}{2} + O(\lambda)\,, \quad N \geq 0\,. \tag{43.38}$$

We then expand the solutions of equation (43.37) in a power series of λ:

$$E_N(g) = \sum E_N^{(n)}(g) \lambda^n. \tag{43.39}$$

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 (43.37) as

$$\frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}} \left(-\frac{2}{g}\right)^E \Gamma(\tfrac{1}{2} - E) = -\epsilon i \Leftrightarrow \frac{\cos \pi E}{\pi} = \epsilon i \frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}} \left(-\frac{2}{g}\right)^E \frac{1}{\Gamma(\frac{1}{2} + E)}. \tag{43.40}$$

For example, the one-instanton contribution is

$$E_N^{(1)}(g) = -\frac{\epsilon}{N!}\left(\frac{2}{g}\right)^{N+1/2} \frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}}\left(1 + O(g)\right). \tag{43.41}$$

The two-instanton contribution is then

$$E_N^{(2)}(g) = \frac{1}{(N!)^2}\left(\frac{2}{g}\right)^{2N+1} \frac{\mathrm{e}^{-1/3g}}{2\pi}\left[\ln(-2/g) - \psi(N+1) + O\left(g\ln g\right)\right], \tag{43.42}$$

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 $g^{-1}\,\mathrm{e}^{-\theta}$ 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

$$E_N^{(n)}(g) = -\left(\frac{2}{g}\right)^{n(N+1/2)}\left(\frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}}\right)^n \left[P_n^N\left(\ln(-g/2)\right) + O\left(g\left(\ln g\right)^{n-1}\right)\right], \tag{43.43}$$

in which $P_n^N(\sigma)$ is a polynomial of degree $n-1$. For example, for $N=0$, we find

$$P_2(\sigma) = \sigma + \gamma\,, \qquad P_3(\sigma) = \frac{3}{2}(\sigma+\gamma)^2 + \frac{\pi^2}{12}\,, \tag{43.44}$$

in which γ is Euler's constant $\gamma = -\psi(1) = 0.577215\ldots$.

Discussion. When we now perform our analytic continuation from g negative to g positive, two things happen: the Borel sums become complex and get an imaginary part exponentially smaller by about a factor $\mathrm{e}^{-1/3g}$ than the real part. Simultaneously, the function $\ln(-2/g)$ also becomes complex and gets an imaginary part $\pm i\pi$. Since the sum of all the contributions is real, the imaginary parts should cancel. 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 derive

$$\operatorname{Im} E^{(0)}(g) \sim \frac{1}{\pi g}\,\mathrm{e}^{-1/3g}\operatorname{Im}\left[P_2\left(\ln(-g/2)\right)\right], \tag{43.45}$$

and, therefore,

$$\operatorname{Im} E^{(0)}(g) \sim -\frac{1}{g}\,\mathrm{e}^{-1/3g}\,. \tag{43.46}$$

Expanding now $E^{(0)}(g)$ in perturbation series

$$E^{(0)}(g) = \sum_k E_k^{(0)} g^k, \tag{43.47}$$

and again using a dispersion relation to calculate the coefficients $E_k^{(0)}$, we obtain the large order behaviour of the perturbative expansion:

$$E_k^{(0)} \underset{k\to\infty}{\sim} \frac{1}{\pi}\int_0^\infty \operatorname{Im}\left[E^{(0)}(g)\right]\frac{\mathrm{d}g}{g^{k+1}}, \Rightarrow\ E_k^{(0)} \sim -\frac{1}{\pi}3^{k+1}k!\,. \tag{43.48}$$

From the imaginary part of P_3, we derive the large order behaviour of the expansion of

$$E^{(1)}(g) = -\frac{1}{\sqrt{\pi g}}\,\mathrm{e}^{-1/6g}\left(1 + \sum^{\infty} E_k^{(1)} g^k\right). \tag{43.49}$$

We express that the imaginary parts of $E^{(1)}(g)$ and $E^{(3)}(g)$ cancel at leading order:

$$\operatorname{Im} E^{(1)}(g) \sim -\left(\frac{\mathrm{e}^{-1/6g}}{\sqrt{\pi g}}\right)^3 \operatorname{Im}\left[P_3\left(\ln(-g/2)\right)\right]. \tag{43.50}$$

For the coefficients $E_k^{(1)}$,we write a dispersion integral

$$E_k^{(1)} = \frac{1}{\pi}\int_0^\infty \left\{\operatorname{Im}\left[E^{(1)}(g)\right]\sqrt{\pi g}\,\mathrm{e}^{1/6g}\right\}\frac{\mathrm{d}g}{g^{k+1}}. \tag{43.51}$$

Then, using equations (43.44) and (43.50), we find

$$E_k^{(1)} \sim -\frac{1}{\pi}\int_0^\infty 3\left[\ln(2/g)+\gamma\right]\mathrm{e}^{-1/3g}\,\frac{\mathrm{d}g}{g^{k+2}}. \tag{43.52}$$

At leading order for k large, we can replace g by its saddle point value $1/3k$ in $\ln g$ and finally obtain

$$E_k^{(1)} = -\frac{3^{k+2}}{\pi}k!\left[\ln 6k + \gamma + O\left(\frac{\ln k}{k}\right)\right]. \tag{43.53}$$

Both results (43.48) and (43.53) have been checked against the numerical behaviour of the corresponding series for which about 100 terms are known.

The behaviour of the real part of P_2 for g has also been verified numerically.

43.2 The Periodic Cosine Potential

To avoid the proliferation of big integer factors it is convenient to use a specific normalization of the coupling constant. We write the hamiltonian

$$H = -\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2 + \frac{1}{16g}\left(1-\cos 4q\sqrt{g}\right). \tag{43.54}$$

As we have already discussed in Chapter 41, to each state of the harmonic oscillator is associated for g small a band of the periodic potential. Indeed 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 hamiltonian and can thus be diagonalized simultaneously. A state in a band is characterized by the eigenvalue $\mathrm{e}^{i\varphi}$ of T and the corresponding eigenfunction $\psi_{N,\varphi}(q)$ of H is thus such that

$$T\psi_{N,\varphi} = \mathrm{e}^{i\varphi}\,\psi_{N,\varphi}\,, \quad H\psi_{N,\varphi} = E_N(g,\varphi)\psi_{N,\varphi}\,, \tag{43.55}$$

where the eigenvalue $E_N(g,\varphi)$ is a periodic function of the angle φ and $E_N(g,\varphi) = N + 1/2 + O(g)$.

The partition function in the φ-sector. We define the partition function in the sector corresponding to an angle φ as the sum

$$\mathcal{Z}(\beta,g,\varphi) = \sum_N \mathrm{e}^{-\beta E_N(g,\varphi)}, \tag{43.56}$$

In particular, for β large,

$$\mathcal{Z}(\beta,g,\varphi) \underset{\beta\to\infty}{\sim} \mathrm{e}^{-\beta E_0(g,\varphi)}. \tag{43.57}$$

We also define the quantity

$$\begin{aligned}\mathcal{Z}_l(\beta,g) \equiv \operatorname{tr}\left(T^l\,\mathrm{e}^{-\beta H}\right) &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\varphi \sum_N \mathrm{e}^{-\beta E_N(g,\varphi)} \int \mathrm{d}q\, \psi^*_{N,\varphi}(q)\psi_{N,\varphi}(q + l\pi/2\sqrt{g}) \\ &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\varphi \sum_N \mathrm{e}^{-\beta E_N(g,\varphi)}\,\mathrm{e}^{il\varphi} = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\varphi\, \mathcal{Z}(\beta,g,\varphi)\,\mathrm{e}^{il\varphi}.\end{aligned} \tag{43.58}$$

Inverting this last relation, we find

$$\mathcal{Z}(\beta,g,\varphi) = \sum_{l=-\infty}^{+\infty} \mathrm{e}^{-il\varphi}\,\mathcal{Z}_l(\beta,g). \tag{43.59}$$

This is the representation we now use to calculate $\mathcal{Z}(\beta,g,\varphi)$. The path integral representation of $\mathcal{Z}_l(\beta,g)$ is

$$\mathcal{Z}_l(\beta,g) = \int_{q(\beta/2)=q(-\beta/2)+l\pi/2\sqrt{g}} [\mathrm{d}q(t)] \exp\left[-\mathcal{S}(q)\right] \tag{43.60}$$

with

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \mathrm{d}t \left[\frac{1}{2}\dot{q}^2(t) + \frac{1}{16g}\left(1-\cos 4q\sqrt{g}\right)\right]. \tag{43.61}$$

Note that the factor $\mathrm{e}^{-il\varphi}$ can be incorporated in the path integral. Indeed since

$$-il\varphi = -\frac{2\sqrt{g}}{\pi}\varphi\big(q(\beta/2) - q(-\beta/2)\big) = -\frac{2i\sqrt{g}}{\pi}\int_{-\beta/2}^{+\beta/2} \mathrm{d}t\, \dot{q}(t),$$

this corresponds to add to $\mathcal{S}(q)$ the integral of a local density

$$\mathcal{S}(q) \mapsto \mathcal{S}(q) + \frac{2i\sqrt{g}}{\pi}\int_{-\beta/2}^{+\beta/2} \mathrm{d}t\, \dot{q}(t). \tag{43.62}$$

The sum over l then is obtained by summing over all periodic trajectories. Note that for β large the configurations for which $q(\beta/2) - q(-\beta/2)$ is not a multiple of the period are suppressed in the large β limit, since their classical action is necessarily infinite. The sum (43.59) can thus be written without restriction on the path $q(t)$:

$$\mathrm{e}^{-\beta E_0(g,\varphi)} \underset{\beta\to\infty}{\sim} \int [\mathrm{d}q(t)] \exp\left[-\mathcal{S}(q) - i\frac{2\sqrt{g}}{\pi}\varphi\int_{-\beta/2}^{+\beta/2} \mathrm{d}t\, \dot{q}(t)\right]. \tag{43.63}$$

These expressions have natural generalizations in the case of the θ-vacuum of the $CP(N-1)$ model and gauge theories (see Sections 41.5,41.6).

Remark. If we expand an eigenvalue $E_N(g,\varphi)$ in a Fourier series

$$E_N(g,\varphi) = \sum_{l=-\infty}^{+\infty} E_{N,l}(g)\, \mathrm{e}^{il\varphi}, \qquad E_{N,l} = E_{N,-l}\,, \tag{43.64}$$

then for g small $E_{N,l}(g)$ is dominated by l-instanton contributions. In particular for the ground state energy in the φ sector, $E_0(g,\varphi)$, the $l=1$ term behaves like

$$E_{0,l=1}(g) \sim \frac{1}{\sqrt{\pi g}}\, \mathrm{e}^{-1/2g}\,. \tag{43.65}$$

The n-instanton configurations. There is one important difference between this example, and the double-well potential. In the previous case, each configuration was a succession of instantons and anti-instantons. Here instead at each step the instanton can go to the next minimum of the potential or the preceding one. We assign, therefore, a sign $\epsilon = +1$ to an instanton and a sign $\epsilon = -1$ to an anti-instanton. A straightforward calculation, similar to the calculation presented above (for details see Appendix A43.2) yields the following interaction term between two consecutive instantons of types ϵ_1 and ϵ_2 separated by a distance θ_{12},

$$\frac{2\epsilon_1\epsilon_2}{g}\, \mathrm{e}^{-\theta_{12}}\,. \tag{43.66}$$

The force between instantons of the same kind is repulsive, while it is attractive for different kinds. We call λ the one-instanton contribution at leading order,

$$\lambda = \frac{1}{\sqrt{\pi g}}\, \mathrm{e}^{-1/2g}\,. \tag{43.67}$$

With this notation the n-instanton contribution reads

$$\mathcal{Z}^{(n)}(\beta,g,\varphi) = \beta\, \mathrm{e}^{-\beta/2}\, \frac{\lambda^n}{n} \int_{\theta_1 \ge 0} \delta\left(\sum_{i=1}^{n} \theta_i - \beta\right) J_n(\theta_i) \tag{43.68}$$

with

$$J_n(\theta_i) = \sum_{\epsilon_i = \pm 1} \exp\left(\sum_{i=1}^{n} -\frac{2}{g}\epsilon_i\epsilon_{i+1}\, \mathrm{e}^{-\theta_i} - i\epsilon_i\varphi\right). \tag{43.69}$$

The additional term $-i\epsilon_i\varphi$ comes from the formula (43.59). 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, then, we obtain

$$\begin{aligned} \mathcal{Z}^{(n)}(\beta,g,\varphi) = {} & \frac{\beta}{2i\pi}\, \mathrm{e}^{-\beta/2}\, \frac{\lambda^n}{n} \oint \mathrm{d}s\, \mathrm{e}^{-\beta s}\, \Gamma^n(-s) \\ & \times \sum_{\{\epsilon_i = \pm 1\}} \exp\left[\sum_{i=1}^{n} -i\epsilon_i\varphi - s\ln(\epsilon_i\epsilon_{i+1}g/2)\right]. \end{aligned} \tag{43.70}$$

We introduce the notation

$$\sigma = \ln(g/2)\,, \tag{43.71}$$

and decide to make the analytic continuation from above so that

$$\ln\left(\tfrac{1}{2}g\epsilon_i\epsilon_{i+1}\right) = \sigma - \tfrac{1}{2}i\pi(1-\epsilon_i\epsilon_{i+1}). \tag{43.72}$$

The expression (43.70) can then be written as

$$\begin{aligned}\mathcal{Z}^{(n)}(\beta,g,\varphi) \sim {} & \frac{\beta}{2i\pi}\,\mathrm{e}^{-\beta/2}\,\frac{\lambda^n}{n}\oint \mathrm{d}s\,\mathrm{e}^{-\beta s}\left[\Gamma(-s)\,\mathrm{e}^{-\sigma s}\right]^n \\ & \times \sum_{\{\epsilon_i=\pm1\}} \exp\left[\sum_{i=1}^{n} -i\epsilon_i\varphi + \tfrac{1}{2}i\pi s(1-\epsilon_i\epsilon_{i+1})\right].\end{aligned} \tag{43.73}$$

The summation over the set $\{\epsilon_i\}$ corresponds to the calculation of the partition function of a one-dimensional Ising model whose transfer matrix $\mathbf{M}$ is

$$\mathbf{M} = \begin{bmatrix} \mathrm{e}^{-i\varphi} & \mathrm{e}^{i\pi s} \\ \mathrm{e}^{i\pi s} & \mathrm{e}^{i\varphi}\end{bmatrix}. \tag{43.74}$$

The sum then is $\operatorname{tr}\mathbf{M}^n$. The eigenvalues $m_\pm$ of $\mathbf{M}$ are

$$m_\pm = \cos\varphi \pm \left(\mathrm{e}^{2i\pi s} - \sin^2\varphi\right)^{1/2}. \tag{43.75}$$

The expression (43.73) can then be written as

$$\mathcal{Z}^{(n)}(\beta,g,\varphi) \sim \frac{\beta\,\mathrm{e}^{-\beta/2}}{2i\pi}\frac{\lambda^n}{n}\oint \mathrm{d}s\,\mathrm{e}^{-\beta s}\left[\Gamma(-s)\,\mathrm{e}^{-\sigma s}\right]^n\left(m_+^n + m_-^n\right). \tag{43.76}$$

The sum $\Sigma(\beta,g)$ of all leading order multi-instanton contributions can now be calculated:

$$\begin{aligned}\Sigma(\beta,g) = \mathrm{e}^{-\beta/2} - \frac{\beta\,\mathrm{e}^{-\beta/2}}{2i\pi}\oint \mathrm{d}s\,\mathrm{e}^{-\beta s}\ln\big\{ & \left[1-\lambda\Gamma(-s)\,\mathrm{e}^{-\sigma s}\,m_+(s)\right] \\ & \times\left[1-\lambda\Gamma(-s)\,\mathrm{e}^{-\sigma s}\,m_-(s)\right]\big\}.\end{aligned} \tag{43.77}$$

The argument of the logarithm can also be written as

$$\begin{aligned}&\left[1-\lambda\Gamma(-s)\,\mathrm{e}^{-\sigma s}\,m_+(s)\right]\left[1-\lambda\Gamma(-s)\,\mathrm{e}^{-\sigma s}\,m_-(s)\right] \\ &\quad = 1-\lambda\Gamma(-s)\,\mathrm{e}^{-\sigma s}\left[2\cos\varphi + \lambda\frac{2i\pi\,\mathrm{e}^{(i\pi-\sigma)s}}{\Gamma(1+s)}\right].\end{aligned} \tag{43.78}$$

An integration by parts of $\beta\,\mathrm{e}^{-\beta s}$ in integral (43.77) yields our final result:

$$\Sigma(\beta,g) = \sum_{N=0}^{\infty}\mathrm{e}^{-\beta E_N(g)}, \tag{43.79}$$

in which $E_N(g) = \frac{1}{2} + s_N(\lambda,\sigma)$ is a solution expandable in powers of λ of the equation

$$\left(\frac{2}{g}\right)^{-E}\frac{\mathrm{e}^{1/2g}}{\Gamma(\frac{1}{2}-E)} + \left(-\frac{2}{g}\right)^{E}\frac{\mathrm{e}^{-1/2g}}{\Gamma(\frac{1}{2}+E)} = \frac{2\cos\varphi}{\sqrt{2\pi}}. \tag{43.80}$$

Note the symmetry in the change $g,E \mapsto -g,-E$. This symmetry is, however, slightly misleading because the equation is actually quadratic in $\Gamma(\frac{1}{2}-E)$ and only one root, corresponding to m_+, is relevant for $g>0$.

43.3 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:

$$\begin{aligned} V(x) &= \tfrac{1}{2}x^2 + O\left(x^3\right), \\ V(x) &= \tfrac{1}{2}\omega^2(x-x_0)^2 + O\left((x-x_0)^3\right). \end{aligned} \tag{43.81}$$

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 calculate 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 take the large β limit of this expression. Using the expressions derived in Section 39.5, we find

$$\left\{\operatorname{tr} \mathrm{e}^{-\beta H}\right\}_{(n)} = i(-1)^n \frac{\beta}{n\sqrt{\pi g}}\sqrt{\frac{\omega C}{n(1+\omega)}}\, \mathrm{e}^{-\omega\beta/[2n(1+\omega)]}\, \mathrm{e}^{-nA(\beta)/g} \tag{43.82}$$

with the definitions

$$C = x_0^2 \omega^{2/(1+\omega)} \exp\left\{\frac{2\omega}{1+\omega}\left[\int_0^{x_0} \mathrm{d}x \left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} - \frac{1}{\omega(x_0-x)}\right)\right]\right\}, \tag{43.83}$$

and

$$A(\beta) = 2\int_0^{x_0} \sqrt{2V(x)}\mathrm{d}x - 2C\frac{(1+\omega)}{\omega}\, \mathrm{e}^{-(\beta/n)\omega/(1+\omega)} + \cdots . \tag{43.84}$$

Note that n has not the same meaning here as in Section 43.1. Since ω is different from 1, the one-instanton configuration does not contribute and n instead counts the number of instanton anti-instanton pairs in the language of Section 43.1. Therefore, n here actually corresponds to $2n$ in the $\omega = 1$ limit.

43.3.1 The n-instanton action

We now call θ_i the successive amounts of time the classical trajectory spends near x_0, and φ_i near the origin. The n-instanton action takes the form

$$A(\theta_i, \varphi_j) = na - 2\sum_{i=1}^{n} \left(C_1\, \mathrm{e}^{-\omega\theta_i} + C_2\, \mathrm{e}^{-\varphi_i}\right) \tag{43.85}$$

with $\sum_{i=1}^{n}(\theta_i + \varphi_i) = \beta$ and

$$a = 2\int_0^{x_0} \sqrt{2V(x)}\mathrm{d}x\,. \tag{43.86}$$

By comparing the value of the action at the saddle point

$$\theta_i = \frac{\beta}{n(1+\omega)}, \qquad \varphi_i = \frac{\omega\beta}{n(1+\omega)} \tag{43.87}$$

with the expression (43.84), we see that we can choose

$$C_1 = C/\omega\,, \qquad C_2 = C \tag{43.88}$$

by adjusting the definitions of θ and φ.

43.3.2 The n-instanton contribution

The n-instanton contribution then has the form

$$\left\{\operatorname{tr} \mathrm{e}^{-\beta H}\right\}_{(n)} = \beta\, \mathrm{e}^{-\beta/2} \frac{\mathrm{e}^{-na/g}}{(\pi g)^n} N_n \int_{\theta_i, \varphi_i \geq 0} \delta\left(\sum_i \theta_i + \varphi_i - \beta\right) \times \exp\left[\sum_{i=1}^{n} \frac{1}{2}(1-\omega)\theta_i - \frac{1}{g} A(\theta, \varphi)\right]. \tag{43.89}$$

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 performing a steepest descent integration over the variables θ_i and φ_i and compare the result with expression (43.82). The result is

$$N_n = \frac{\left(C\sqrt{\omega}\right)^n}{n}. \tag{43.90}$$

The factor $1/n$ comes from the symmetry of the action under cyclic permutations of the θ_i and φ_i.

We now set

$$\lambda = \frac{\mathrm{e}^{-a/g}}{\pi g} C\sqrt{\omega}\,, \qquad \mu = -\frac{2C}{g}\,. \tag{43.91}$$

As in Section 43.1, we introduce an integral representation for the δ-function:

$$\delta\left(\sum_{i=1}^{n} (\theta_i + \varphi_i) - \beta\right) = \frac{1}{2i\pi} \int_{-i\infty-\eta}^{+i\infty+\eta} ds \exp\left[-s\beta + s\sum_{i=1}^{n} (\theta_i + \varphi_i)\right] \tag{43.92}$$

with $n > 0$.

The expression (43.89) can be rewritten as

$$\left\{\operatorname{tr} \mathrm{e}^{-\beta H}\right\}_{(n)} = \beta\, \mathrm{e}^{-\beta/2} \frac{\lambda^n}{n} \frac{1}{2i\pi} \int_{-i\infty-\eta}^{+i\infty-\eta} ds\, \mathrm{e}^{-s\beta} \left[I(s) J(s)\right]^n, \tag{43.93}$$

where we now have defined

$$I(s) = \int_0^{+\infty} \mathrm{e}^{\theta s - \mu\, \mathrm{e}^{-\theta}}\, d\theta\,, \tag{43.94}$$

$$J(s) = \int_0^{+\infty} \exp\left\{\left[\tfrac{1}{2}(1-\omega) + s\right]\theta - \frac{\mu}{\omega} \mathrm{e}^{-\omega\theta}\right\} d\theta\,. \tag{43.95}$$

In the small g limit, the integrals can be evaluated as

$$I(s) = \Gamma(-s)\mu^s, \tag{43.96}$$

$$J(s) = \frac{1}{\sqrt{\mu\omega}} \Gamma\left(\tfrac{1}{2} - (s+\tfrac{1}{2})/\omega\right) \left(\frac{\mu}{\omega}\right)^{(s+1/2)/\omega}. \tag{43.97}$$

We call as before $\Sigma(\beta, g)$ 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:

$$\Sigma(\beta, g) = \sum_{\alpha} \mathrm{e}^{-\beta E_\alpha}, \tag{43.98}$$

in which the values $E_\alpha = \frac{1}{2} + s_\alpha$ are the solutions expandable for g small of the equation

$$\left(\frac{\mu}{\omega}\right)^{E/\omega} \Gamma(\tfrac{1}{2} - E/\omega)\mu^E \Gamma(\tfrac{1}{2} - E)\frac{\mathrm{e}^{-a/g}}{2\pi} = -1\,. \tag{43.99}$$

We now notice that we find two series of energy levels corresponding to the poles of the two Γ-functions:

$$E_N = N + \tfrac{1}{2} + O(\lambda)\,, \tag{43.100}$$
$$E_N = (N + \tfrac{1}{2})\omega + O(\lambda)\,. \tag{43.101}$$

The same expression contains the instanton contributions to the two different sets of levels.

One can verify that the many-instanton contributions are singular for $\omega = 1$. But if one directly sets $\omega = 1$ in equation (43.99), one obtains

$$\mu^{2E}\Gamma^2(\tfrac{1}{2} - E)\frac{\mathrm{e}^{-a/g}}{2\pi} = -1\,,$$

equation which can be rewritten as

$$\mu^{E}\Gamma(\tfrac{1}{2} - E)\frac{\mathrm{e}^{-a/(2g)}}{\sqrt{2\pi}} = \pm i\,. \tag{43.102}$$

This is exactly the set of two equations obtained in Section 43.1.

43.3.3 Large order estimates of perturbation theory

The expression (43.99) can be used to determine the large order behaviour of perturbation theory by calculating the imaginary part of the leading instanton contribution and writing a dispersion integral as we have done in Section 42.1.1. For the energy $E_N(g) = N + \frac{1}{2} + O(g)$, one finds

$$\operatorname{Im} E_N(g) \sim K_N g^{-(N+1/2)(1+1/\omega)}\,\mathrm{e}^{-a/g} \tag{43.103}$$

with

$$\begin{aligned} K_N = {} & \frac{(-1)^{N+1}}{2\pi N!}\omega^{-(N+1/2)/\omega}\,(2C)^{(N+1/2)(1+1/\omega)}\sin[\pi(N + \tfrac{1}{2})(1 + 1/\omega)] \\ & \times \Gamma[\tfrac{1}{2} - (N + \tfrac{1}{2})/\omega]. \end{aligned} \tag{43.104}$$

From the imaginary part of $E_N(g)$, one deduces at large order k:

$$E_{Nk} \sim K_N \frac{\Gamma\big(k + (N + 1/2)(1 + 1/\omega)\big)}{a^{k+(N+1/2)(1+1/\omega)}}\big(1 + O\left(k^{-1}\right)\big). \tag{43.105}$$

Note that this expression, in contrast with the instanton contribution to the real part, is uniform in the limit $\omega = 1$ in which the result (43.48) is recovered.

43.4 The $O(\nu)$ Symmetric Anharmonic Oscillator

It is interesting to consider a last example, the analytic continuation of the energy levels of the $O(\nu)$ symmetric anharmonic oscillator for $g < 0$:

$$H = -\tfrac{1}{2}\nabla^2 + \tfrac{1}{2}\mathbf{q}^2 + g\left(\mathbf{q}^2\right)^2 . \tag{43.106}$$

We discuss the $\nu = 2$ example, but the generalization to arbitrary values of ν is simple.

The $O(2)$ anharmonic oscillator. The instanton solution can be written as

$$\mathbf{q}(t) = \mathbf{u} f(t) , \tag{43.107}$$

in which $\mathbf{u}$ is a fixed unit vector. The one-instanton contribution to the ground state energy is

$$E^{(1)}(g) = \frac{4i}{g}\, \mathrm{e}^{1/3g}\left(1 + O(g)\right) \quad \text{for } g \to 0_- . \tag{43.108}$$

From a calculation of the instanton interaction, one infers the n-instanton action:

$$A(\theta_i) = -\frac{1}{3}n - 4\sum_i \mathrm{e}^{-\theta_i}\cos\varphi_i , \tag{43.109}$$

in which θ_i is the distance between two successive instantons and φ_i the angle between them:

$$\cos\varphi_i = \mathbf{u}_i \cdot \mathbf{u}_{i+1} . \tag{43.110}$$

It is useful to consider the quantity (Sections 3.3,23.5)

$$\operatorname{tr}\left[R(\alpha)\,\mathrm{e}^{-\beta H}\right] = \int [\mathrm{d}\mathbf{q}(t)] \exp\left(-\mathcal{S}(q)\right) \quad \text{with} \quad \hat{\mathbf{q}}(-\beta/2)\cdot\hat{\mathbf{q}}(\beta/2) = \cos\alpha . \tag{43.111}$$

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 r.h.s. of equation (43.111) can be rewritten as

$$\operatorname{tr}\left[R(\alpha)\,\mathrm{e}^{-\beta H}\right] = \sum_{l,N} \mathrm{e}^{-il\alpha - \beta E_{l,N}} . \tag{43.112}$$

In this expression, l is the angular momentum. The boundary condition imposed on the path integral (43.111) implies a constraint on the many-instanton configuration:

$$\sum_{i=1}^{n} \varphi_i = \alpha , \tag{43.113}$$

constraint which can be implemented through the identity

$$\delta\left(\sum_{i=1}^{n}\varphi_i - \alpha\right) = \frac{1}{2\pi}\sum_{l=-\infty}^{+\infty} \exp\left[il\left(\sum_{i=1}^{n}\varphi_i\right) - i\alpha l\right] . \tag{43.114}$$

The n-instanton contribution to expression (43.112) then takes the form

$$\left\{\operatorname{tr}\left[R(\alpha)\,\mathrm{e}^{-\beta H}\right]\right\}_{(n)} \sim \frac{\lambda^n}{2i\pi n}\beta\,\mathrm{e}^{-\beta}\int \mathrm{d}s\;\mathrm{e}^{-s\beta}\sum_{l=-\infty}^{+\infty}\mathrm{e}^{-il\alpha}\left[I_l(s)\right]^n, \tag{43.115}$$

where we have set

$$\lambda = \frac{4i}{g}\,\mathrm{e}^{1/3g}, \qquad \mu = -\frac{4}{g}, \tag{43.116}$$

$$I_l(s) = \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\varphi\int_{-\infty}^{+\infty}\mathrm{d}\theta\,\exp\left(s\theta + il\varphi - \mu\,\mathrm{e}^{-\theta}\cos\varphi\right). \tag{43.117}$$

We call $\Sigma_l(\beta,g)$ the generating function of n-instanton contributions at fixed angular momentum l:

$$\Sigma_l(\beta,g) = \sum_n \frac{\lambda^n}{2i\pi n}\beta\,\mathrm{e}^{-\beta}\int \mathrm{d}s\;\mathrm{e}^{-s\beta}\left[I_l(s)\right]^n. \tag{43.118}$$

Let us evaluate $I_l(s)$. The integration over θ yields

$$I_l(s) = \mu^s\Gamma(-s)\int_0^{2\pi}\frac{\mathrm{d}\varphi}{2\pi}\,\mathrm{e}^{il\varphi}\left(\cos\varphi\right)^s. \tag{43.119}$$

Performing the last integral and using various relations among Γ functions, we finally obtain

$$I_l(s) = \mu^s\,\mathrm{e}^{\frac{1}{2}i\pi(s+1)}\frac{\Gamma(\frac{1}{2}(l-s))2^{-s-1}}{\Gamma(1+\frac{1}{2}(l+s))}. \tag{43.120}$$

It is easy to verify that

$$I_l(s) = I_{-l}(s). \tag{43.121}$$

From equation (43.118), we now derive the result

$$\operatorname{tr}\mathrm{e}^{-\beta H_l} = \sum_{N=0}^{\infty}\mathrm{e}^{-\beta E_{N,l}} \tag{43.122}$$

with $E_{N,l} = s_{N,l} + 1$ the solution of the equation

$$\mathrm{e}^{1/(3g)}\left(-\frac{2}{g}\right)^E \mathrm{e}^{i\pi(E+l)/2}\frac{\Gamma\big(\frac{1}{2}(l+1-E)\big)}{\Gamma\big(\frac{1}{2}(l+1+E)\big)} = 1, \tag{43.123}$$

which satisfies

$$E_{N,l} = l + 2N + 1 + O(g), \quad N \ge 0. \tag{43.124}$$

Note that checks about these expressions are provided by the surprising perturbative relation between the $O(2)$ anharmonic oscillator with negative coupling and the double-well potential.

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 (43.123) the corresponding substitution, one obtains

$$i\,\mathrm{e}^{1/(3g)}\left(-\frac{2}{g}\right)^E \mathrm{e}^{i\pi(E+l+\nu/2)/2}\frac{\Gamma\big(\frac{1}{2}(l+\nu/2-E)\big)}{\Gamma\big(\frac{1}{2}(l+\nu/2+E)\big)} = 1, \tag{43.125}$$

At leading order in λ, one recovers the imaginary part of the energy levels for g small and negative:

$$\operatorname{Im} E_{N,l} \underset{g\to 0_-}{=} -\frac{1}{N!}\frac{1}{\Gamma(\frac{1}{2}\nu+l+N)}\left(\frac{2}{g}\right)^{(\nu/2)+l+2N} \mathrm{e}^{1/3g}\left(1+O(g)\right). \tag{43.126}$$

Using the Cauchy formula, one can derive from this expression large order estimates for perturbation theory. At next order in λ, 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.

43.5 Generalized Bohr–Sommerfeld Quantization Formula

Up to now, we have considered instanton contributions only at leading order. However, the form of the result is extremely suggestive and has led us to conjecture the general form of the semi-classical expansion for potentials with degenerate minima.

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 Appendix A43.4).

We introduce the function

$$D(E,g) = E + \sum_{k=1}^{\infty} g^k D_{k+1}(E), \tag{43.127}$$

in terms of which the perturbation expansion for an energy level $E_N^{(0)}$ can be obtained by inverting

$$N+\tfrac{1}{2} = D(E^{(0)},g). \tag{43.128}$$

One verifies that as a power series $D(E,g) = -D(-E,-g)$.

This is the form of the usual exact Bohr–Sommerfeld quantization condition.

To take into account instanton contributions, we write instead another equation, which generalize Bohr–Sommerfeld quantization condition to the double-well potential,

$$\frac{1}{\sqrt{2\pi}}\Gamma(\tfrac{1}{2}-D)\,(-2/g)^{D(E,g)}\,\mathrm{e}^{-A(g,E)/2} = \epsilon i \tag{43.129}$$

with

$$A(g,E) = \frac{1}{3g} + \sum_{k=1}^{\infty} g^k A_{k+1}(E), \tag{43.130}$$

where again in the sense of power series $A(g,E) = -A(-g,-E)$. The functions $A(E,g)$, $D(E,g)$ can be calculated by expanding the corresponding WKB series, which is an expansion at gE fixed, for E small (for details see Appendix A43.4). The coefficients $D_k(E)$ and $A_k(E)$ are polynomials of degree k in E.

If we solve (43.129) in the one-instanton approximation and substitute into equation (43.127), we find

$$E = E^{(0)}(g) - \epsilon\left(\frac{2}{g}\right)^{N}\frac{1}{N!}\frac{\mathrm{e}^{-A(g,E^{(0)})/2}}{\sqrt{\pi g}}\frac{\partial D}{\partial E}\left(E^{(0)}\right)^{-1}. \tag{43.131}$$

We see that 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 (43.129), we find for the energy level $E_N(g) = N + 1/2 + O(g)$ the following expansion,

$$E_N(g) = \sum_0^{\infty} E_l^{(0)} g^l + \sum_{n=1}^{\infty} \frac{1}{g^{Nn}} \left(\frac{1}{\sqrt{\pi g}}\, \mathrm{e}^{-1/6g}\right)^n \sum_{k=0}^{n-1} (\ln(-2/g))^k \sum_{l=0}^{\infty} e_{nkl} g^l . \quad (43.132)$$

All the series in powers of g appearing in this expansion are determined by the perturbative expansion of A and D. This phenomenon has later found an explanation in the framework of the theory of *resurgent* functions.

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 quantum mechanics, simplify the problem of the summation of the many-instanton contributions.

In Appendix A43.4, we give the generalized Bohr–Sommerfeld quantization formulae for other potential with degenerate minima.

Bibliographical Notes

The two-instanton contribution and its application to large order behaviour for potential with degenerate minima has been discussed in

E. Brézin, G. Parisi and J. Zinn-Justin, *Phys. Rev.* D16 (1977) 408; E.B. Bogomolny, *Phys. Lett.* 91B (1980) 431.

The multi-instanton contributions have been calculated in

J. Zinn-Justin, *Nucl. Phys.* B192 (1981) 125; B218 (1983) 333; *J. Math. Phys.* 22 (1981) 511; 25 (1984) 549.

In these articles a relation between the $O(2)$ anharmonic oscillator and the double-well potential was conjectured, which was later proven in

A.A. Andrianov, *Ann. Phys. (NY)* 140 (1982) 82.

Multi-instantons in the $O(3)$ non-linear σ-model have been considered in

V.A. Fateev, I.V. Frolov and A.S. Schwarz, *Nucl. Phys.* B154 (1979) 1, *Yad. Fiz.* 30 (1979) 1134; B. Berg and M. Lüscher, *Commun. Math. Phys.* 69 (1979) 57; A.P. Bukhvostov and L.N. Lipatov, *Nucl. Phys.* B180[FS2] (1981) 116.

The peculiar equations (43.127,43.129) which imply relations between the perturbative expansions around multi-instantons have been recently proven using the theory of *resurgent* functions in

F. Pham, *Resurgence, Quantized Canonical Transformation and Multi-Instanton*, Algebraic Analysis, vol. II (1988), *C.R. Acad. Sci. Paris* 309 (1989) 999; E. Delabaere et H. Dillinger, thesis Université de Nice (Nice 1991).

For the WKB method as discussed in the appendix, see also

A. Voros, *Annales IHP* 29 (1983) 3.

APPENDIX A43

A43.1 Multi-Instantons: The Determinant

We can write the operator M defined by equation (43.20) as

$$M = -\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 + 1 + \sum_{i=1}^{n} v(t-t_i), \tag{A43.1}$$

in which $v(t)$ is a potential localized around $t = 0$, and t_i are the positions of the instantons:

$$v(t) = O\left(\mathrm{e}^{-|t|}\right) \quad \text{for } |t| \to \infty\,. \tag{A43.2}$$

We want to calculate

$$\det M\, M_0^{-1} = \det\left\{1 + \left[-(\mathrm{d}/\mathrm{d}t)^2 + 1\right]^{-1} \sum_{i=1}^{n} v(t-t_i)\right\}. \tag{A43.3}$$

Using the identity $\ln\det = \operatorname{tr}\ln$, we expand the r.h.s. in powers of $v(t)$:

$$\begin{aligned}\ln\det M M_0^{-1} = {} & \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \int \Bigg[\Delta(u_1-u_2) \sum_{i_1=1}^{n} v(u_2 - t_{i_1})\Delta(u_2-u_3)\ldots\Delta(u_k-u_1) \\ & \times \sum_{i_k=1}^{n} v(u_1 - t_{i_k})\Bigg] \prod_{j=1}^{k} \mathrm{d}u_j \end{aligned} \tag{A43.4}$$

with the definition

$$\Delta(t) = \left\langle 0 \right| \left[-(\mathrm{d}/\mathrm{d}t)^2 + 1\right]^{-1} \left| t \right\rangle \sim \frac{1}{2}\mathrm{e}^{-|t|} \quad \text{for } 1 \ll t \ll \beta\,. \tag{A43.5}$$

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 from each potential the same instanton contribution survive. Therefore,

$$\ln\det M M_0^{-1} = n \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \int \Delta(u_1-u_2)v(u_2)\cdots\Delta(u_k-u_1)v(u_1) \prod_{j=1}^{k} \mathrm{d}u_j \quad \text{for } |t_i - t_j| \gg 1\,. \tag{A43.6}$$

We recognize n times the logarithm of the one-instanton determinant.

A43.2 The Instanton Interaction

We assume, as in Section 43.3, that the potential has two degenerate minima at the points $x = 0$ and $x = x_0$ with

$$\begin{aligned} V(x) &= \tfrac{1}{2}x^2 + O\left(x^3\right) \\ V(x) &= \tfrac{1}{2}\omega^2\left(x - x_0\right)^2 + O\left((x - x_0)^3\right). \end{aligned} \tag{A43.7}$$

The one-instanton solution $q_c(t)$ which goes from 0 to $q_0 = x_0/\sqrt{g}$ can be written as

$$q_c(t) = f(t)/\sqrt{g}\,. \tag{A43.8}$$

We choose the function $f(t)$ in such a way that it satisfies

$$\begin{aligned} x_0 - f(t) &\sim \sqrt{C}\,\mathrm{e}^{-\omega t}/\omega \quad \text{for } t \to +\infty \\ f(t) &\sim \sqrt{C}\,\mathrm{e}^{t} \quad \text{for } t \to -\infty\,. \end{aligned} \tag{A43.9}$$

By solving the equation of motion, it is easy to calculate the constant

$$C = x_0^2\omega^{2/(1+\omega)}\exp\left\{\frac{2\omega}{1+\omega}\left[\int_0^{x_0}\mathrm{d}x\left(\frac{1}{\sqrt{2V(x)}} - \frac{1}{x} - \frac{1}{\omega(x_0 - x)}\right)\right]\right\}. \tag{A43.10}$$

We recognize the constant (43.83).

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 Chapter 43, we take as a two-instanton configuration:

$$q_1(t) = \frac{1}{\sqrt{g}}\left(f_+(t) + \epsilon f_-(t)\right), \qquad \epsilon = \pm 1 \tag{A43.11}$$

with

$$f_+(t) = f(t - \theta/2), \qquad f_-(t) = f(-t - \theta/2). \tag{A43.12}$$

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. Finally for the classical trajectory which goes instead from the origin to q_0 and back we take

$$q_2(t) = \left[f(t + \theta/2) + f(\theta/2 - t) - x_0\right]/\sqrt{g}\,. \tag{A43.13}$$

We now 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:

$$\mathcal{S}(q_1) = \mathcal{S}_+(q_1) + \mathcal{S}_-(q_1) \tag{A43.14}$$

with

$$\begin{aligned} \mathcal{S}_+(q_1) &= \int_0^{+\infty} \left[\tfrac{1}{2}\dot{q}_1^2 + V\left(\sqrt{g}q_1(t)\right)/g\right] \mathrm{d}t, \\ \mathcal{S}_-(q_1) &= \int_{-\infty}^{0} \left[\tfrac{1}{2}\dot{q}_1^2 + V\left(\sqrt{g}q_1(t)\right)/g\right] \mathrm{d}t. \end{aligned} \tag{A43.15}$$

The value $t = 0$ of the separation point is somewhat arbitrary and can be replaced by any value which remains finite when θ becomes infinite. 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

$$\begin{aligned} \mathcal{S}_+(q_1) = \frac{1}{g}\int_0^{+\infty} \mathrm{d}t \Big\{ \left[\tfrac{1}{2}\dot{f}_+^2(t) + V\left(f_+(t)\right)\right] + \epsilon\left[\dot{f}_-(t)\dot{f}_+(t) + V'\left(f_+(t)\right)f_-(t)\right] \\ + \tfrac{1}{2}\left[\dot{f}_-^2(t) + V''\left(f_+(t)\right)f_-^2(t)\right]\Big\}. \end{aligned} \tag{A43.16}$$

Since $f_-(t)$ decreases exponentially, only values of t small compared to $\theta/2$ contribute to the last term of equation (A43.16) which is proportional to V''. For such values of t, we have

$$\tfrac{1}{2}V''\left(f_+(t)\right)f_-^2 \sim V\left(f_-(t)\right). \tag{A43.17}$$

For the terms linear in $f_-(t)$, we integrate by parts the kinetic term and use the equation of motion

$$\ddot{f}(t) = V'\left[f(t)\right]. \tag{A43.18}$$

Only the integrated term survives and yields

$$\int_0^{+\infty} \mathrm{d}t \left[\dot{f}_-(t)\dot{f}_+(t) + V'\left(f_+(t)\right)f_-(t)\right] = -\dot{f}(-\theta/2)f(-\theta/2). \tag{A43.19}$$

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

$$\mathcal{S}(q_1) = \frac{1}{g}\left[2a - 2\epsilon f(-\theta/2)\dot{f}(-\theta/2) + \cdots\right] \tag{A43.20}$$

with

$$a = \int_0^{x_0} \sqrt{2V(x)}\mathrm{d}x\,. \tag{A43.21}$$

Replacing, for θ large, f by its asymptotic form (A43.9), we finally obtain the classical action

$$\mathcal{S}(q_1) = g^{-1}\left[2a - 2C\epsilon\,\mathrm{e}^{-\theta} + O\left(\mathrm{e}^{-2\theta}\right)\right], \tag{A43.22}$$

and, therefore, the instanton interaction.

The calculation of the classical action corresponding to $q_2(t)$ follows the same steps and one finds

$$\mathcal{S}(q_2) = \frac{1}{g}\left\{2a - 2\left[f(\theta/2) - x_0\right]\dot{f}(\theta/2) + \cdots\right\}, \tag{A43.23}$$

which for θ large is equivalent to

$$\mathcal{S}(q_2) = \frac{1}{g}\left[2a - 2(C/\omega)\,\mathrm{e}^{-\omega\theta}\right]. \tag{A43.24}$$

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 $q = q_0$:

$$\mathcal{S}(q) = g^{-1}\left[2a - 2C\left(\mathrm{e}^{-\beta+\theta} + \mathrm{e}^{-\omega\theta}/\omega\right)\right], \qquad (A43.25)$$

in agreement with equations (43.85,43.88).

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 Appendix 39.4.1. The main problem with such a method is to find a system of constraints which are both theoretically reasonable, 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):

$$1 = \int \prod_{i=1}^{n} \left[\int \mathrm{d}t\, \dot{q}^2_{\epsilon_i}(t-t_i)\right] \delta\left[\int \mathrm{d}t\, \dot{q}_{\epsilon_i}(t-t_i)\big(q(t) - q_{\epsilon_i}(t-t_i)\big)\right] \mathrm{d}t_i\,, \qquad (A43.26)$$

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

$$\left(\frac{\|\dot{q}_+\|^2}{2i\pi}\right)^n \int \prod_{i=1}^{n} \mathrm{d}t_i \mathrm{d}\lambda_i \int [\mathrm{d}q(t)] \prod_{i=1}^{n} \exp\left[-\mathcal{S}(q,\lambda_i)\right] \quad \text{with}$$
$$\mathcal{S}(q,\lambda_i) = \mathcal{S}(q) + \sum_{i=1}^{n} \lambda_i \int \mathrm{d}t\, \dot{q}_{\epsilon_i}(t-t_i)\big(q(t) - q_{\epsilon_i}(t-t_i)\big). \qquad (A43.27)$$

The arguments of Appendix 39.4.1 can then be generalized to recover the results of Section 43.1.

A43.3 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 quantum mechanics which we have studied in Chapter 43. We consider the function

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}q \, \exp\left[-\frac{1}{g} V(q\sqrt{g})\right], \qquad (A43.28)$$

where $V(x)$ is an entire function with an absolute minimum at $x = 0$, $V(0) = 0$. For g small, $I(g)$ can be calculated by steepest descent, expanding V around $q = 0$:

$$I(g) = \sum_{k \geq 0} I_k g^k. \qquad (A43.29)$$

One can write a finite dimensional integral of the form (A43.28) as a generalized Borel or Laplace transform,

$$I(g) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}q\, \mathrm{d}t\, \delta\left[V(q\sqrt{g}) - t\right] \mathrm{e}^{-t/g}\,. \qquad (A43.30)$$

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 $q = q_0$:

$$\mathcal{S}(q) = g^{-1}\left[2a - 2C\left(\mathrm{e}^{-\beta+\theta} + \mathrm{e}^{-\omega\theta}/\omega\right)\right], \tag{A43.25}$$

in agreement with equations (43.85,43.88).

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 Appendix 39.4.1. The main problem with such a method is to find a system of constraints which are both theoretically reasonable, 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):

$$1 = \int \prod_{i=1}^{n} \left[\int \mathrm{d}t\, \dot q^2_{\epsilon_i}(t-t_i)\right] \delta\left[\int \mathrm{d}t\, \dot q_{\epsilon_i}(t-t_i)\bigl(q(t) - q_{\epsilon_i}(t-t_i)\bigr)\right] \mathrm{d}t_i\,, \tag{A43.26}$$

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

$$\begin{gathered}\left(\frac{\|\dot q_+\|^2}{2i\pi}\right)^n \int \prod_{i=1}^{n} \mathrm{d}t_i \mathrm{d}\lambda_i \int [\mathrm{d}q(t)] \prod_{i=1}^{n} \exp\left[-\mathcal{S}(q,\lambda_i)\right] \quad \text{with} \\ \mathcal{S}(q,\lambda_i) = \mathcal{S}(q) + \sum_{i=1}^{n} \lambda_i \int \mathrm{d}t\, \dot q_{\epsilon_i}(t-t_i)\bigl(q(t) - q_{\epsilon_i}(t-t_i)\bigr).\end{gathered} \tag{A43.27}$$

The arguments of Appendix 39.4.1 can then be generalized to recover the results of Section 43.1.

A43.3 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 quantum mechanics which we have studied in Chapter 43. We consider the function

$$I(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}q\, \exp\left[-\frac{1}{g} V(q\sqrt{g})\right], \tag{A43.28}$$

where $V(x)$ is an entire function with an absolute minimum at $x = 0$, $V(0) = 0$. For g small, $I(g)$ can be calculated by steepest descent, expanding V around $q = 0$:

$$I(g) = \sum_{k \ge 0} I_k g^k. \tag{A43.29}$$

One can write a finite dimensional integral of the form ($A43.28$) as a generalized Borel or Laplace transform,

$$I(g) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}q\, \mathrm{d}t\, \delta\left[V(q\sqrt{g}) - t\right] \mathrm{e}^{-t/g}\,. \tag{A43.30}$$

We integrate over q,

$$I(g)=\frac{1}{\sqrt{2\pi g}}\int_0^{\infty}\mathrm{d}t\,\mathrm{e}^{-t/g}\sum_i\frac{1}{|V'[x_i(t)]|},\qquad (A43.31)$$

in which $\{x_i(t)\}$ are the solutions of the equation

$$V[x_i(t)]=t\,.\qquad (A43.32)$$

When the function $V(x)$ is monotonous both for x positive and negative, equation (A43.32) has two solutions for all values of t and equation (A43.31) is directly the Borel representation of the function $I(g)$, which has a Borel summable power series expansion.

We now assume instead that $V(x)$ has a second local minimum which gives a negligible contribution to $I(g)$ for g small. A simple example is

$$V(x)=\tfrac{1}{2}x^2-\tfrac{1}{3a}x^3(1+a)+\tfrac{1}{4a}x^4,\quad \tfrac{1}{2}<a<1\,,\qquad (A43.33)$$

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:

$$I_k\underset{k\to\infty}{\propto}\Gamma(k)\,[V(a)]^{-k}\,,\qquad V(a)>0\,,\qquad (A43.34)$$

(in the example (A43.33) $V(a)=a^2(1-a/2)/6$) and the series is not Borel summable.

The *naive* Borel transform of $I(g)$ is obtained by retaining in equation (A43.31) only the roots of equation (A43.32) which exists for t small. The singularities of the Borel transform then correspond to the zeros of $V''(x)$.

For the potential (A43.33), the expression (A43.31) has the form (θ is the step function)

$$I(g)=\frac{1}{\sqrt{2\pi g}}\int_0^{+\infty}\mathrm{d}t\,\mathrm{e}^{-t/g}\left[\frac{1}{\left|V'(x_1(t))\right|}+\frac{\theta(V(a)-t)}{\left|V'(x_2(t))\right|}+\frac{\theta(V(a)-t)\theta(t-V(1))}{\left|V'(x_3(t))\right|}\right.$$
$$\left.+\frac{\theta(t-V(1))}{\left|V'(x_4(t))\right|}\right],\qquad (A43.35)$$

with the definitions (see figure 43.2): $x_1(t)\le 0\le x_2(t)\le a\le x_3(t)\le 1\le x_4(t)$.

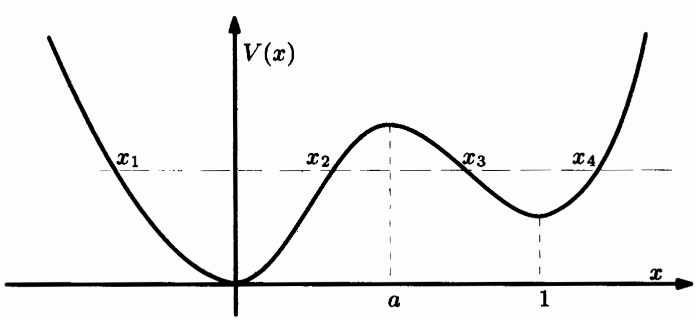


Fig. 43.2 The four roots of equation (A43.32).

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:

$$V[x_2(t)] = t \pm i\epsilon \,. \qquad (A43.36)$$

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 $t > V(a)$. Since $x_2(t)$ and $x_3(t)$ meet at $t = V(a)$, the analytic continuation will correspond to take for $x_3(t)$ the other solution:

$$V[x_3(t)] = t \mp i\epsilon \,. \qquad (A43.37)$$

We, therefore, 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.

We have, therefore, succeeded in writing expression ($A43.35$) as the sum of three saddle point contributions (see figure 43.3). There is some arbitrariness in this decomposition which here corresponds to the choice $\epsilon = \pm 1$.

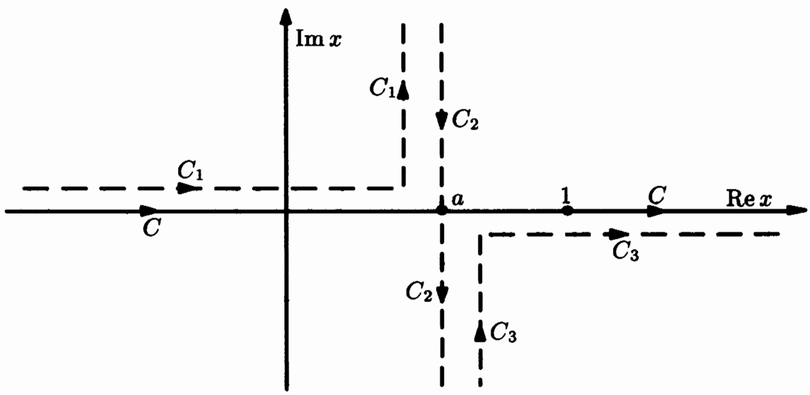


Fig. 43.3 The different contours in the x plane.

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 $0, a, 1$.

A43.4 Multi-Instantons and WKB Approximation

We now give a more general form of the conjecture presented in Section 43.5 and indicate how a few properties of the functions $A(E, g)$ and $D(E, g)$ can be obtained from the WKB expansion.

A43.4.1 The conjecture

The double-well potential. For the double-well potential we have conjectured equation (43.129),

$$\frac{1}{\sqrt{2\pi}}\Gamma\left(\tfrac{1}{2} - D\right)(-2/g)^{D(E,g)}\, \mathrm{e}^{-A(g,E)/2} = \epsilon i \,, \qquad (A43.38)$$

where the functions $D(E, g)$ and $A(E, g)$ have the expansions (43.127) and (43.130), respectively. This conjecture has been generalized to a potential with two asymmetric wells:

$$\frac{1}{2\pi}\Gamma\left(\tfrac{1}{2}-D_1(E,g)\right)\left(-\frac{2C_1}{g}\right)^{D_1(E,g)}\Gamma\left(\tfrac{1}{2}-D_2(E,g)\right)\left(-\frac{C_2}{g}\right)^{D_2(E,g)}\mathrm{e}^{-A(g,E)}=-1\,, \tag{A43.39}$$

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 is

$$\left(\frac{2}{g}\right)^{-D}\frac{\mathrm{e}^{A(E,g)/2}}{\Gamma(\frac{1}{2}-D)}+\left(\frac{-2}{g}\right)^{D}\frac{\mathrm{e}^{-A(g,E)/2}}{\Gamma(\frac{1}{2}+D)}=\frac{2\cos\varphi}{\sqrt{2\pi}}. \tag{A43.40}$$

The $O(\nu)$-symmetric anharmonic oscillator. For the $O(\nu)$ anharmonic oscillator, the conjecture corresponds to the expansion of the energy levels for $g < 0$ and, in particular, yields the instanton contributions to the large order behaviour.

$$i\,\mathrm{e}^{-A(E,g)}\left(-\frac{2}{g}\right)^{D}\mathrm{e}^{i\pi(D+l+\nu/2)/2}\frac{\Gamma\left(\frac{1}{2}(l+\nu/2-D)\right)}{\Gamma\left(\frac{1}{2}(l+\nu/2+D)\right)}=1\,. \tag{A43.41}$$

A43.4.2 The WKB approximation

We consider the Schrödinger equation

$$-\tfrac{1}{2}\psi''+\left(V(\sqrt{g}q)/g\right)\psi=E\psi\,. \tag{A43.42}$$

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

$$-g^2\psi''+2V(x)\psi=2gE\psi\,. \tag{A43.43}$$

The WKB expansion is an expansion for $g \to 0$ at Eg fixed, in contrast with the perturbative expansion where E is fixed. It can be constructed by introducing the corresponding Riccati equation, setting

$$S(x)=-g\psi'/\psi\,, \tag{A43.44}$$

where S satisfies

$$gS'(x)-S^2(x)+S_0^2(x)=0\,,\qquad S_0^2(x)=2V(x)-2gE\,. \tag{A43.45}$$

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:

$$S(x,g,E)=S_+(x,g,E)+S_-(x,g,E)\,,\qquad S_\pm(x,-g,-E)=\pm S_\pm(x,g,E)\,. \tag{A43.46}$$

It follows

$$gS_-'-S_+^2-S_-^2+S_0^2=0\,, \tag{A43.47}$$

$$gS_+'-2S_+S_-=0\,, \tag{A43.48}$$

which allows us to express the wave function in terms of S_+ only:

$$\psi = (S_+)^{-1/2} \exp\left[-\frac{1}{g}\int_{x_0}^{x} \mathrm{d}x'\, S_+(x')\right]. \tag{A43.49}$$

The spectrum can then be determined by the condition

$$\frac{1}{2i\pi}\oint_C \mathrm{d}z \frac{\psi'(z)}{\psi(z)} = N\,, \tag{A43.50}$$

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 figure 43.2). In terms of S_+ equation ($A43.50$) becomes

$$-\frac{1}{2i\pi g}\oint_C \mathrm{d}z\, S_+(z) = N + \tfrac{1}{2}\,. \tag{A43.51}$$

If we replace S_+ by its WKB expansion and expand each term in a power series of Eg we get the function $D(E,g)$ (the perturbative expansion):

$$-\frac{1}{2i\pi g}\oint_C \mathrm{d}z\, S_+(z) = D(E,g) = -D(-E,-g)\,. \tag{A43.52}$$

Potentials with degenerate minima. In the case of potentials with degenerate minima, two functions D_1 and D_2 (in the notation ($A43.39$)) 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 $A(g,E)$:

$$\begin{aligned}\frac{1}{g}\oint_{C'} \mathrm{d}z\, S_+(z) = A(E,g) + \ln(2\pi) - \sum_{i=1}^{2} \ln\Gamma\left(\tfrac{1}{2} - D_i(E,g)\right) \\ + D_i(E,g)\ln(-2g/C_i)\,,\end{aligned} \tag{A43.53}$$

where C' encloses $[x_2, x_3]$ in figure 43.2. 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

$$\mu + 2N + 1 = D(E,g,\mu), \qquad \mu = l + \nu/2 - 1\,, \tag{A43.54}$$

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 ($A43.46$)) of $-g(\psi_l'/\psi_l + (\nu-1)/2|q|)$. The following properties can then verified

$$D(E,g,\mu) = -D(-E,-g,\mu), \qquad D(E,g,\mu) = D(E,g,-\mu),$$

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.

INDEX

INDEX