θ RENORMALIZATION, SUPERUNIVERSALITY, AND ELECTRON-ELECTRON INTERACTIONS IN THE THEORY OF THE QUANTUM HALL EFFECT

Promotiecommissie

Promotor	prof. dr. A. M. M. Pruisken
Overige leden	dr. M. A. Baranov
	prof. dr. E. Brezin
	prof. dr. H. W. Capel
	prof. dr. A. M. Finkelstein
	prof. dr. K. Schoutens
	prof. dr. J. Smit

The research described in this thesis was carried out at the Institute for Theoretical Physics of the University of Amsterdam, Valckeneirstraat 65, 1018 XE Amsterdam, The Netherlands.

Printed by PrintPartners Ipskamp, Enschede

ISBN 90-5776-154-8

θ RENORMALIZATION, SUPERUNIVERSALITY, AND ELECTRON-ELECTRON INTERACTIONS IN THE THEORY OF THE QUANTUM HALL EFFECT

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus,
prof. mr. P. F. van der Heijden,
ten overstaan van een door het college voor promoties ingestelde
commissie in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 24 mei 2006, te 10:00 uur.

door

Igor Burmistrov
geboren te Moskou, Rusland

Prof. dr. A. M. M. Pruisken (promotor)

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

Chapter 1

Introduction

1.1 A bit of history

The integral and fractional quantum Hall effects [1, 2] are remarkable and richly complex phenomena of nature with a level of significance that is comparable to that of superconductivity and superfluidity. The robust quantization of the Hall conductance is observed in experiments on two dimensional electron systems at low temperatures ($50mK$ - $4K$) and strong perpendicular magnetic fields ($1 - 20T$). This quantization phenomenon originally came as a complete surprise in Physics and appeared to be in fundamental conflict with the prevailing ideas on electron transport in metals, especially the semi classical theory that explains the ordinary Hall effect. [3, 4, 5] Although discovered relatively recently [6, 7] the quantum Hall effect has already led to two Nobel prizes in Physics, one in 1985 for the discovery of the integral quantum Hall effect [8] and one in 1998 for the discovery of the fractional quantum Hall effect. [9, 10, 11]

In spite of the huge number of theoretical and experimental papers that have appeared over the years, and the dramatic progress that has been made, our microscopic understanding of the quantum Hall effect is still far from being complete. From the experimental side, the list of observed but unexplained transport phenomena is still growing and the subject matter, as it now stands, goes well beyond any of the theoretical scenario's that were originally proposed to explain the quantum Hall plateaus. The main reason why a microscopic theory of the quantum Hall effect is complicated is the fact the problem generally lacks a "small parameter". Any attempt to force the phenomenon in the mold of "mean-field" theory is doomed to fail. A satisfactory theory of the quantum Hall effect must treat the various different microscopic factors on an equal footing. These factors are i) strong magnetic fields, ii) a random impurity potential and iii) the effects of electron-electron interactions. Besides all these factors the theory should also be designed to deal with such phenomena as the linear response in electromagnetic potentials.

Traditionally, the integral and fractional quantum Hall effect have been studied largely separately. The reason for this otherwise artificial distinction is as follows. The integral quantum Hall effect can be understood, at least qualitatively, in terms of *localized* and *extended* single electron states of the random impurity problem in strong magnetic fields. On the other hand, the *electron-electron interaction* is known to be one of the prerequisites for the existence of the fractional quantum Hall effect.

1.1.1 Old paradigm: Anderson *de*-localization

The random impurity problem, otherwise known as the theory of *Anderson localization*, has already had a very long history even before the quantum Hall effects had been observed. A dramatic breakthrough took place at about the same time as von Klitzing's discovery. It was then understood, by various different schools of thought, that the noninteracting two dimensional electron gas in the presence of a random potential fluctuations is always an *insulator*, at least for sample sizes that are large enough. [12, 13, 14, 15, 16, 17] These famous results of what is called the "scaling theory" of Anderson localization were in spectacular conflict with the existence of the quantum Hall effect, however. In order for the quantum Hall effect to exist one had to assume that the strong magnetic field adds something very dramatic to the theory of Anderson localization that was previously unknown. This *something* turned out to be a peculiar *topological* feature of replica field theory [18, 19] that is invisible in the standard diagrammatic techniques of Anderson localization. The glorious idea proposed in Ref. [19] provides the mechanism by which "extended" states appear in the problem and, hence, the true reason why the quantum Hall effect actually exists. [20, 21] Moreover, these topological concepts led directly to the prediction of *quantum criticality* of the quantum Hall plateau *transitions* with interesting scaling behavior of the magneto resistance data with varying temperature and magnetic field. [22, 23] Soon after the first laboratory experiments on quantum criticality had been conducted various different schemes for numerical simulation on the disordered free electron gas were developed. At present there is an impressive stock of numerical exponent values obtained by many different authors and this includes the multi-fractal singularity spectrum of density fluctuations. [24, 25]

1.1.2 Grand unification

Contrary to the integral quantum Hall effect, electron-electron interactions have always been the main objective in the studies of the fractional quantum Hall effect. The physics of random potential fluctuations is completely excluded from the traditional considerations which are phenomenological in nature. Within this paradigm a number of approaches have emerged over the years and only four of them have survived. These are the Laughlin wave function, [26] Jain's composite fermion picture, [27] the Chern-Simons statistical gauge field approach [28] and effective theories for the quantum Hall edge. [29]

The most ambitious program would be to incorporate all the advances made in both the integral and fractional quantum Hall effect in a single microscopic formalism that would lead to a *grand unified* theory of the quantum Hall effect. For this purpose it is important to know that amongst the various different approaches that presently exist, there is only one theory that deals successfully with localization and interaction effects in metals, namely the one originally developed by Finkelstein. [30] The

Finkelstein theory is notorious for its mathematical complexity, however, and much of the original structure of the theory was previously not at all understood. It therefore did not occur to anyone to use the Finkelstein theory in order lay the foundation for a unified theory of the quantum Hall effect. Opposite to all expectations in the field, however, it was argued [31] that by incorporating the aforementioned topological features in the Finkelstein theory one would have the very same non-perturbative mechanism for electron "de-localization" as the one that had previously been discovered, in the theory of free electrons. Moreover, several very basic advances have been made that facilitate the study of the Finkelstein approach as a field theory. I mention, in particular, the discovery of a new interaction symmetry that had previously remained concealed, namely $\mathcal{F}$-invariance. This symmetry is intimately related to the usual electrodynamic $U(1)$ gauge invariance of the interacting electron gas and, among many other things, it permits the coupling of the Finkelstein action to external electromagnetic potentials. [32, 33] The foundations for a unified theory of the quantum Hall effect were finally laid in Refs [34, 35]. In this theory the fractional quantum Hall regime emerges from exactly same formalism as the one was used for the integral regime.

1.1.3 New paradigm: Super universality

In spite of enormous progress achieved in Refs [32, 33, 34, 35], the unification of the integral and fractional quantum Hall effects with the theory disordered metals was still far from being complete at the time this PhD project had started. For one thing, detailed computations had to be performed that would demonstrate the renormalizability of the unified action both on a *perturbative* and *non-perturbative* level. For another, some of the most important new insights in the field, in particular that of *super universality* of topological principles in quantum field theory, was only beginning to see daylight at that time. The statement of super universality essentially says that the fundamental aspects of the quantum Hall effect are all quite generally displayed by the *instanton vacuum* in asymptotically free field theory, independent of the specific application that one is interested in such as the number of field components in the theory.

The super universality concept fundamentally upsets the prevailing ideas and expectations in the field, in particular the many conflicting claims on the θ dependence as well as topological excitations (instantons and instanton gases) that have resulted from the historical studies [36, 37, 38] of the large N expansion of the CP^{N-1} model. These claims [39, 40, 41] have promoted, for a very long time, the wrong physical ideas and incorrect mathematical objectives in the literature, for example the longstanding belief which says that quantum Hall physics is merely a feature of the theory in the *replica limit*. This, in turn, has motivated a never ending series of utterly incorrect mathematical papers advocating the "failure of the replica trick", the "superiority of supersymmetry" etc. etc. The complete lack of physical insight in the literature is in sharp contrast to the general idea with which the θ parameter originally was introduced in the theory of the quantum Hall effect. Following the original papers one takes the experiment on the quantum Hall regime as an important laboratory where the strong coupling problems previously encountered in QCD can be explored and investigated in great detail.

ϵ expansion

The first open problem, to be addressed in this thesis, is that of the renormalizability of the unified theory beyond one loop order. This issue, which has important consequences for the field of localization and interaction effects as a whole, was partially addressed in Ref. [33] where the scheme of dimensional regularization [42] was introduced in the problem containing the so-called *singlet interaction amplitude*. However, a detailed understanding of the problem in $2+\epsilon$ dimensions goes well beyond the pioneering work of Ref. [33]. A complete theory to order ϵ^2 is important, especially since it is the only place where one can get explicit information on longstanding issues such as *dynamical scaling* of the interacting electron gas. Moreover, following up on the seminal work [43] by E. Brezin and J. Zinn-Justin in the seventies, the ϵ expansion generally facilitates the extraction of explicit scaling functions for the physical observables of the theory, in particular the equation of state for the macroscopic conductances. For the first time, explicit scaling results are being obtained and the "mobility edge problem" in $2+\epsilon$ dimensions in the presence of electron-electron interactions is solved completely.

θ renormalization

The second major topic, studied in this thesis, is the peculiar *mechanism* that is responsible for generating "extended states" or "massless excitations" at $\theta = \pi$ in the *interacting* electron gas in the presence of strong magnetic fields. Notice that the highly nontrivial idea which says that this mechanism should be identically the same for both free electrons and interacting systems is firmly supported by the results obtained from numerical simulations [24, 25] and those taken in the laboratory. [23] To obtain a quantitative assessment of this mechanism I develop, in this thesis, a theory of topological excitations (*instantons*) for both interacting and non-interacting systems. This theory is based, to a large extend, on the methodology originally developed for QCD by 'tHooft. [44] However, there are several very basic advances as compared to the traditional theories and views on instantons and instanton effects.

First of all, unlike the usual free energy computations which do not teach us anything about the singularity structure of the theory with varying θ, our main interest is in the non-perturbative consequences for the renormalization group, in particular the concept of θ renormalization by instantons. This concept is important since it demonstrates that instantons are the fundamental topological objects of the theory that establish the cross-over between the weak coupling Goldstone phase at short distances and the strong coupling quantum Hall phase that generally appears at much larger distances only. This scenario of quantum Hall physics which, by the way, one can quite generally associate with the θ parameter in asymptotically free field theory, is based on several basic principles that have remarkably emerged in recent years only. I mention in particular the *massless chiral edge excitations* that are quite generally displayed by an instanton vacuum, [34, 35] as well as the intimately related theory of *physical observables* (in the context of Gross) that generally defines the renormalization of the system. This thesis is largely devoted to a study of the consequences of these new principles.

Spatially varying masses

As a second technical advancement I extend the theory of instantons to generally include *mass terms.* Mass terms play an extremely important role in condensed matter applications, especially since it is well known that they determine to a major extend the physics of the problem. It is also well known that mass terms dramatically complicate the subject of topological excitations and the matter usually goes by the name of "constrained instantons". Even though the construction of a constrained instanton formally solves the semiclassical problem, [45] to our knowledge there has been nobody until to date who has been able to solve the full quantum theory, for otherwise very good reasons. It has already been noticed previously that the fundamental problem with mass terms is that they generally spoil the *geometrical* features of the harmonic oscillator problem or instanton determinant. [31] This clearly indicates that the very idea of constrained instantons is rather unnatural and it does not facilitate an analysis of the renormalization behavior of the theory. A different way of saying this is that many of the difficulties encountered in quantum field theory have actually been borne out of pursuing the wrong physical objectives.

The problem can be avoided all together by following up on the procedure of *spatially varying masses* which essentially transforms the mass terms defined in *flat* space into those defined in *curved* space. [31] This procedure permits a complete solution of the harmonic oscillator problem while retaining the ultraviolet singularity structure of the theory. Moreover, it is possible to return back to the problem of mass terms in *flat* space all the way in the end, by employing the various tricks in evaluating instanton determinants as introduced first by 'tHooft. [44] The methodology of spatially varying masses ultimately provides the most important piece of information that - to our knowledge - cannot be obtained in any different manner, namely non-perturbative (instanton) contributions the anomalous dimension or γ function associated with mass terms.

Experimental predictions

Having established a complete understanding of the various distinctly different aspects of the renormalization group it is next important to move on to the third major topic of this thesis which is to translate the results into clear predictions for the experiment. As a first test I revisit the renormalization theory of the free electron gas which has created this field of research in the first place. Even though this theory has original set the stage for the experiments on quantum criticality in the quantum Hall regime, the results were previously not understood well enough to provide serious numerical estimates for the critical exponents. At present I shall greatly benefit from the numerical exponent values that are known from an extensive series of computer simulations on the free electron gas by many others. [24, 25] In fact, the predictions of this thesis that are based on instantons agree remarkably well with the numerical estimates obtained from the experiments on the free electron gas taken from the computer. This includes not only the correlation (or localization) length exponent but also the entire multi fractal singularity spectrum of the plateau transitions. For the first time, therefore, a complete microscopic theory of the quantum Hall effect has been established and, as an integral part of this thesis, I shall outline the various different reasons why instanton calculus works so well, at least as far as the free electron gas is concerned. Starting from the surprising results obtained for free electrons it becomes easier to understand the fundamentally different physics of the Coulomb interaction problem and, in particular, the reasons why Fermi liquid principles are fundamentally wrong.

The results of this thesis, along with their physical objectives, may in general be regarded as some of the basic advancements which have led to the aforementioned concept of super universality of quantum Hall physics. At the time of writing of this thesis this concept has been investigated in a variety of completely different physical contexts. These include, besides the interacting and non-interacting electron gas, also the exactly solvable θ dependence of the large N expansion of the CP^{N-1} model, [46, 47] the theory of quantum spin chains [48] as well as the Ambegaokar-Eckern-Schön model [49] for the Coulomb blockade problem. [50] Each of these cases stands for a distinctly different realization of the θ parameter in a scale invariant theory. The basic features are nevertheless identically the same as those described in this thesis. These include the existence of *robust* topological quantum numbers that explain the precision and stability of the quantum Hall plateaus, as well as *gapless* excitations at $\theta = \pi$ that generally facilitate a transition to take place between adjacent quantum Hall plateaus.

1.2 What *is* the quantum Hall effect?

Two dimensional electron gas is realized on interface between two semiconductors, e.g. Si and SiO_2, GaAs and $Al_xGa_{1-x}As$, or $In_xGa_{1-x}As$ and InP. [5] The samples used in experiments on scaling typically have a low mobility $\mu \sim 10^4 - 10^5 cm^2/V \cdot s$ in the absence of magnetic field. The transport measurements are performed with the help of the AC lock-in technique as sketched in Fig. 1.1. When a current I is applied to the electron system a voltage V_0 is measured (see Fig. 1.1). In the presence of a perpendicular magnetic field B a voltage drop, usually referred to as the Hall voltage, appears in the direction perpendicular to the current flow. According to the classical theory of metals, [4] the Hall resistance $R_H = V_H/I$ is related to filling factor of the Landau levels $\nu_f = nch/eB$ where n denotes the electron density, c the speed of light, h the Plank constant and e the electron charge. The following simple relation is obtained

$$R_H^{-1} = \frac{e^2}{h}\nu_f \tag{1.2.1}$$

which means that the Hall resistance increases linearly with the magnetic field B.

What has been remarkably observed in the original experiments by von Klitzing, Dorda and Pepper, [6] however, is that Eq. (1.2.1), although true at room temperature, transforms at very low temperatures in the range $0 - 4K$ into a series of quantum

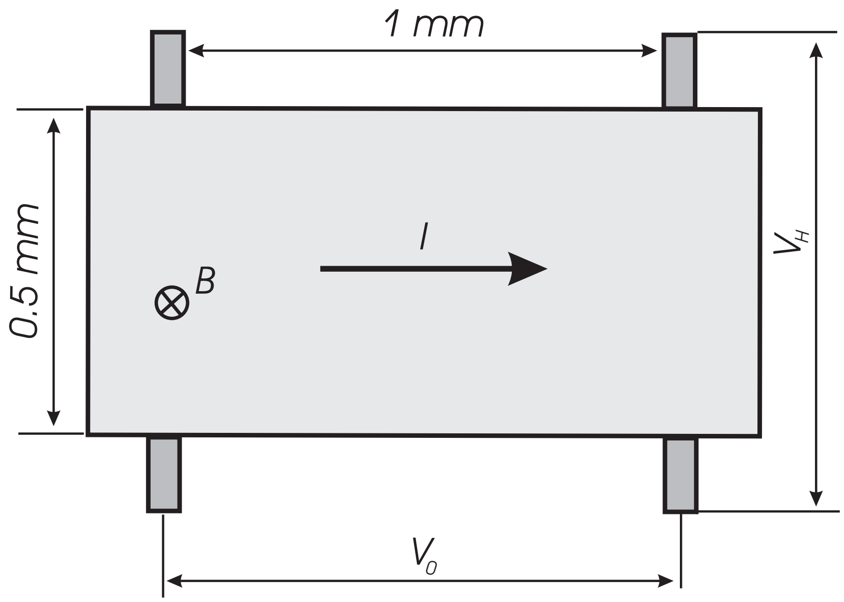


Figure 1.1: Schematic diagram of Hall bar.

Hall plateaus that can be represented as follows

$$R_H^{-1} = \frac{e^2}{h} \sum_{n=0}^{\infty} \vartheta(\nu_f - n - 1/2). \tag{1.2.2}$$

Here, $\vartheta(x)$ stands for the Heaviside step function. For magnetic field strengths in the range $n - 1/2 \lesssim \nu_f \lesssim n + 1/2$ the Hall resistance is accurately quantized according to $R_H^{-1}[h/e^2] = n$. When the temperature approaches absolute zero, the steps between adjacent quantum Hall plateaus become infinitely sharp whereas at finite temperatures the steps are smoothed out. The quantization phenomenon of the Hall resistance is now referred to as the integral quantum Hall effect.

In better quality, higher mobility samples and at high magnetic field strengths (usually more than $10\,T$) it has been discovered experimentally by Tsui, Störmer and Gossard [7] that the Hall resistance is fractionally quantized around filling factors $\nu_f = p/q$ according to $R_H^{-1} h/e^2 = p/q$ with p an arbitrary integer and q an odd integer. This quantization is usually termed the fractional quantum Hall effect. The quantum Hall effects, both integral and fractional, can be observed in a single sample at low temperatures by varying the applied magnetic field [51] as shown in Fig. 1.2. An important general feature of the quantum Hall effect is that the longitudinal resistance $R_0 = V_0/I$ vanishes when the Hall resistance is in the plateau regime.

At a later stage it was theoretically predicted [22] and experimentally verified by measurements taken from low mobility $In_xGa_{1-x}As/InP$ heterostructures [23] that both the longitudinal resistance R_0 and the Hall R_H resistance, with varying temperature and magnetic field, display the following scaling behavior at the *transition*

($\nu_f \approx n + 1/2$) between adjacent quantum Hall plateaus

$$R_{0,H} = f_{0,H}\left(T^{-\kappa}(\nu_f - n - 1/2)\right). \tag{1.2.3}$$

Here, $f_0(X)$ and $f_H(X)$ are regular functions of their argument. The critical exponent κ is universal which means that its numerical value is the same for all Landau levels n and sample independent. The first measurements [23] resulted in an experimental value $\kappa = 0.42 \pm 0.04$.

However, during the period of almost two decades that followed experimentalists were mainly struggling with the conceptual difficulties associated with quantum phase transitions in disordered systems. Some of the more difficult aspects of the theory relating to the *observability* of quantum criticality were generally discarded and/or overlooked. Nevertheless, it was already understood from the very first beginning that scaling results like Eq. (1.2.3) are valid only when the temperature approaches absolute zero. The characteristic temperature or length scale where scaling occurs first depends strongly on the details such as the type of random potential fluctuations in the sample, [22] the presence of macroscopic inhomogeneities etc. [52] As has already been stated at many different places elsewhere, these complications generally impose serious constraints on the experimental design and the technology of growing samples.

In experimental samples the random potential is created by the dopants that can be situated at some distance d from two dimensional electron system. If the distance d is large compared to the magnetic field length $l_H = \sqrt{c\hbar/eB}$, $d \gg l_H$ then the dopants cause long-ranged potential fluctuations. This is what happens in GaAs/Al_xGa_{1-x}As heterostructures that generally do not display scaling (1.2.3) in the range of experimental temperatures. On the other hand, in In_xGa_{1-x}As/InP samples the dopants are known to be located near the interface such that $d \ll l_H$ and is the main reason why these samples were chosen in the original experiments on quantum criticality.

Macroscopic inhomogeneities such as gradients in the electron density, misalignments of the Hall bar contacts etc. generally prevent an accurate measurement of the critical exponent κ taken from the plateau-plateau transitions. [52] It so turned out that only the measurements on the plateau-insulator provide a reasonable accuracy and latest results on In_xGa_{1-x}As/InP samples [53] indicate that the critical exponent equals $\kappa = 0.57 \pm 0.02$ rather than 0.42 ± 0.04.

Only very recently detailed studies have begun on a series of state of the art samples with a well defined disorder by the group of D.C. Tsui in Princeton. [54] The results seem to be in favor of the value $\kappa = 0.42$ as originally reported by H.P. Wei *et al.* [23] Unfortunately, however, the new data were presented using the old Fermi liquid ideas that were introduced at the time of the original experiments. [22] In this thesis I will show that our microscopic understanding of the Coulomb interaction problem [32, 33] has dramatically advanced in recent years and phenomenological ideas based on Fermi liquid principles are no longer satisfactory, by any standard.

To specify the effects of the electron-electron interaction on the plateau transitions we consider the Fourier transform of the pair potential $U(q)$. If $U(q) \to \infty$ as $q \to 0$ we name the electron-electron interaction long-ranged. When $U(q) \to \text{const}$ as $q \to 0$ the electron-electron interaction is short-ranged. In Ref. [32] it has been established that the system with long-ranged electron-electron interaction possesses

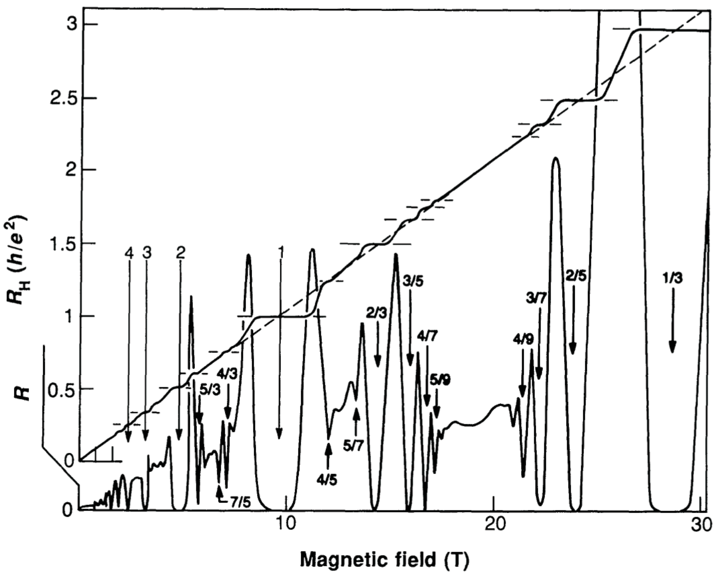


Figure 1.2: Hall resistance R_H and resistance R_0 as functions of an applied magnetic field. Figure is taken from Ref. [51]

$\mathcal{F}$-invariance. This symmetry is absent or broken for short-ranged interactions as well as free electrons. This remarkable result implies that the critical exponent κ is different depending on the range of the interaction potential. In realistic samples the electrons interact through the Coulomb potential which is long-ranged. Therefore, it is an experimental challenge to design a metallic gate close to the two dimensional electron gas which renders the effective electron-electron interaction short-ranged. [31] A different value of κ will then observed and this value can be compared with the analytical results reported in this thesis.

In summary I can say that the quantum Hall regime is an outstanding laboratory for investigating quantum phase transitions and interaction effects, some of the leading topics of current interest.

1.3 Outline of this thesis

Chapter 2 is devoted to the Grassmannian $U(m+n)/U(m) \times U(n)$ non-linear sigma model in the presence of the θ term. This theory in the limit $m = n = 0$ is known to describe the disordered free electron gas in a strong magnetic field. [19] I start out elaborating on the general consequences of the massless chiral edge excitations in the theory with arbitrary values of m and n. [46] This new ingredient of the theory can be employed to define the general renormalization behavior of the system in an unambiguous manner. In the second part of this chapter I first revisit the instanton methodology developed in Refs [20, 21] and introduce the methodology of spatially varying masses. Based on the non-perturbative results for the renormalization group β and γ I discuss the quantum critical behavior of the theory at $\theta = \pi$. In the range $0 \leqslant m, n \lesssim 1$ I find a second order transition with exponents that vary continuously with varying values of m and n. The predictions of the theory with $m = n = 0$ are being compared with correlation length and multi fractal exponents that are known from numerical simulations on the free electron gas.

In Chapter 3 I embark on the spin polarized (or spinless) disordered electron gas in $2+\epsilon$ dimensions in the presence of the Coulomb interaction. I present the results of a computation to second order in an expansion in ϵ which is one order higher than what was known previously. The complete scaling behavior is obtained for quantities like the conductivity and the specific heat near the metal-insulator transition or mobility edge. Finally, we employ the results in the construction of a generalized scaling diagram for the interacting electron gas in the conductance versus c plain where the parameter $0 \leq c \leq 1$ denotes the range of the electron-electron interaction.

Chapter 4 is devoted to instanton computations in the theory of interacting electrons in two dimensions and in strong magnetic fields. In the first part I introduce the theory of physical observables and formulate the general topological principles by which the Hall conductance is robustly quantized. In the second and main part of this Chapter I generalize the theory of instantons to include the interacting electron gas and derive the non-perturbative contributions to the renormalization group β and γ functions. The results can be represented in terms of a three dimensional renormalization group diagram spanned by the conductances σ_{xx} and σ_{xy} and the aforementioned parameter c.

Finally, in Chapter 5 I derive the various different aspects of scaling in the weak and strong coupling regimes of the interacting electron gas. In the weak coupling regime I explain the "topological oscillations" of the magnetoresistance components that have recently been observed experimentally [55] and predict that similar oscillations occur the specific heat of the electron gas. In the strong coupling regime I derive general scaling results for the quantum Hall plateau transitions and discuss the fundamental consequences of the Fermi-liquid and non-Fermi liquid critical fixed points for the experiments on the quantum Hall regime.

Chapter 2

θ renormalization in generalized CP^{N-1} models

2.1 Introduction

2.1.1 Super universality

The quantum Hall effect has remained one of the most beautiful and outstanding experimental realizations of the *instanton vacuum* concept in non-linear sigma models. [1, 22] Although originally introduced in the context of Anderson (de-)localization in strong magnetic fields, [18, 19] the topological ideas in quantum field theory have mainly been extended in recent years to include a range of physical phenomena and applications that are much richer and broader than what was previously anticipated. What remarkably emerges is that the aforementioned topological concepts retain their significance also when the Coulomb interaction between the disordered electrons is taken into account. [31] A detailed understanding of interaction effects is vitally important not only for conducting experiments on *quantum criticality* of the plateau transitions, [23, 63, 64, 65] but also for the long standing quest for a *unified action* that incorporates the low energy dynamics of both the integral and fractional quantum Hall states. [32, 33, 34, 35]

Perhaps the most profound advancement in the field has been the idea which says that the instanton vacuum generally displays *massless chiral edge excitations*. [34, 66] These provide the resolution of the many *strong coupling* problems that historically have been associated with the instanton vacuum concept in scale invariant theories. [37, 38, 67] The physical significance of the *edge* is most clearly demonstrated by the fact that the instanton vacuum theory, unlike the phenomenological approaches to the fractional quantum Hall effect based on Chern Simons gauge theory, [68] can be used to derive from first principles the complete *Luttinger liquid* theory of edge excitations in disordered abelian quantum Hall systems. [34] Along with the physics of the *edge* came the important general statement which says that the *fundamental features* of the quantum Hall effect should all be regarded as *super universal* features of the topological concept of an instanton vacuum, i.e. independent of the number of field components in the theory. [66]

Super universality includes not only the appearance of *massless chiral edge excitations* but also the existence of *gapless bulk excitations* at $\theta = \pi$ in general as well as the dynamic generation of *robust topological quantum numbers* that explain the *precision* and *observability* of the quantum Hall effect. [66] Moreover, the previously unrecognized concept of *super universality* provides the basic answer to the historical controversies on such fundamental issues as the *quantization of topological charge*, the exact significance of having *discrete topological sectors* in the theory, the precise meaning of *instantons* and *instanton gases*, [37, 38] the validity of the *replica method* etc. etc. One can now state that many of these historical problems arose because of a complete lack of any physical assessment of the theory, both in general and in more specific cases such as the exactly solvable large N expansion of the CP^{N-1} model.

2.1.2 The background field methodology

In 1987, Pruisken introduced a renormalization group scheme in replica field theory (non-linear sigma model) that was specifically designed for the purpose of extracting the non-perturbative features of the quantum Hall regime from the instanton angle θ. [20, 21] This procedure was motivated, to a large extend, by the Kubo formalism for the conductances which, in turn, has a natural translation in quantum field theory, namely the *background field methodology*.

Generally speaking, the background field procedure expresses the renormalization of the theory in terms of the *response* of the system to a change in the *boundary conditions*. It has turned out that this procedure has a quite general significance in asymptotically free field theory that is not limited to *replica limits* and condensed matter applications alone. It actually provides a general, conceptual framework for the understanding of the strong coupling aspects of the theory that otherwise remain inaccessible. For example, such non-perturbative features like *dynamic mass generation* are in one-to-one correspondence with the renormalized parameters of the theory since they are, by construction, a probe for the *sensitivity* of the system to a change in the *boundary conditions*.

The background field procedure has been particularly illuminating as far as the perturbative aspects of the renormalization group is concerned. First of all, it is the appropriate generalization of Thouless' ideas on localization, [69] indicating that the physical objectives in condensed matter theory and those in asymptotically free field theory are in many ways the same. Secondly, it provides certain technical advantages in actual computations and yields more relevant results. For example as we show in next chapter, it will lead to an exact solution (in the context of an ϵ expansion) of the the AC conductivity in the *mobility edge* problem or *metal-insulator* problem in $2 + \epsilon$ dimensions. The physical significance of these results is not limited, once more, to the theory in the *replica limit* alone. They teach us something quite general about the statistical mechanics of the Goldstone phase in low dimensions.

2.1.3 The strong coupling problem

With hindsight one can say that the background field procedure, [21] as it stood for a very long time, did not provide the complete conceptual framework that is necessary for general understanding of the quantum Hall effect or, for that matter, the instanton vacuum concept in quantum field theory. Unlike the conventional theory where the precise details of the "edge" do not play any significant role, in the presence of the instanton parameter θ the choice in the boundary conditions suddenly becomes an all important conceptual issue that is directly related to the definition of a fundamental quantity in the theory, the *Hall conductance*.

The physical significance of *boundary conditions* in this problem has been an annoying and long standing puzzle that has fundamentally complicated the development of a microscopic theory of the quantum Hall effect. [1] In most places in the literature this problem has been ignored altogether. [37, 38, 67] In several other cases, however, it has led to a mishandling of the theory. [70]

The discovery of *super universality* in non-linear sigma models [46] has provided the physical clarity that previously was lacking. The existence of *massless chiral edge excitations*, well known in studies of quantum Hall systems, implies that the instanton vacuum concept generally supports distinctly different modes of excitation, those describing the *bulk* of the system and those associated with the *edge*. It has turned out that each of these modes has a fundamentally different topological significance, and a completely different behavior under the action of the renormalization group.

The existence of massless chiral edge excitations forces one to develop a general understanding of the instanton vacuum concept that is in many ways very different from the conventional ideas and expectations in the field. It turns out that most of the physics of the problem emerges by asking how the two dimensional *bulk* modes and one dimensional *edge* modes can be separated and studied individually. At the same time, a distinction ought to be made between the physical *observables* that are defined by the *bulk* of the system and those that are associated with the *edge*.

Effective action for the edge modes

The remarkable thing about the problem with *edge* modes is that it *automatically* provides all the fundamental quantities and topological concepts that are necessary to describe and understand the low energy dynamics of the system. Much of the resolution to the problem resides in the fact that the theory can generally be written in terms of *bulk* field variables that are embedded in a *background* of the topologically different *edge* field configurations. This permits one to formulate an *effective action* for the *edge* modes, obtained by formally eliminating all the *bulk* degrees of freedom from the theory. [66, 46] It now turns out that the *effective action* procedure for the *edge* field variables proceeds along exactly the same lines as the *background field* methodology [21] that was previously introduced for entirely different physical reasons! This remarkable coincidence has a deep physical significance and far reaching physical consequences. In fact, the many different aspects of the problem (Kubo formulae, renormalization, edge currents etc.) as well as the various disconnected pieces of the puzzle (boundary conditions, quantization of topological charge, quantum Hall effect etc.) now become simultaneously important. They all come together as fundamental and distinctly different aspects of a single new concept in the problem that has emerged from the instanton vacuum itself, the effective action for the massless chiral edge excitations. [66, 46]

The quantum Hall effect

The effective action for massless edge excitations has direct consequences for the strong coupling behavior of the theory that previously remained concealed. It essentially tells us how the instanton vacuum *dynamically generates* the aforementioned *super universal* features of the quantum Hall effect, in the limit of large distances.

The large N expansion of the CP^{N-1} model can be used as an illuminating and exactly solvable example that sets the stage for the *super universality* concept in asymptotically free field theory. [46] The most significant quantities of the theory are the renormalization group β functions for the *response* parameters σ_{xx} and σ_{xy} that appear in the effective action for massless chiral edge excitations, (see also Fig. 2.1)

$$\frac{d\sigma_{xx}}{d\ln\mu} = \beta_\sigma(\sigma_{xx}, \sigma_{xy}), \tag{2.1.1}$$

$$\frac{d\sigma_{xy}}{d\ln\mu} = \beta_\theta(\sigma_{xx}, \sigma_{xy}). \tag{2.1.2}$$

Here, the parameters σ_{xx} and σ_{xy} are precisely analogous to the Kubo formulae for *longitudinal conductance* and *Hall conductance* in quantum Hall systems. They stand for the (inverse) *coupling constant* and $\theta/2\pi$ respectively, both of which appear as *running* parameters in quantum field theory.

The infrared *stable fixed points* in Fig. 2.1, located at integer values of σ_{xy}, indicate that the Hall conductance is *robustly quantized* with corrections that are exponentially small in the size of the system. [46] The *unstable fixed points* at half-integer values of σ_{xy} or $\theta = \pi$ indicate that the large N system develops a *gapless phase* or a *continuously divergent correlation length* ξ with a critical exponent ν equal to 1/2, [46]

$$\xi \propto |\theta - \pi|^{-1/2}. \tag{2.1.3}$$

2.1.4 The large N expansion

One of the most impressive features of the large N expansion is that it is exactly solvable for *all* values of θ. This is unlike the $O(3)$ non-linear sigma model, for example, which is known to be integrable for $\theta = 0$ and π only and the exact information that can be extracted is rather limited. [71] Nevertheless, both cases appear as outstanding limiting examples in the more general context of *replica field theory* or, equivalently, the Grassmannian $U(m+n)/U(m) \times U(n)$ non-linear sigma model. This Grassmannian manifold is a generalization of the CP^{N-1} manifold that describes, as is well known, the Anderson localization problem in strong magnetic fields. [19]

It is extremely important to know, however, that *none* of the *super universal* features of the instanton vacuum where previously known to exist, neither in the historical papers on the large N expansion [37, 38, 67, 36] nor in the intensively

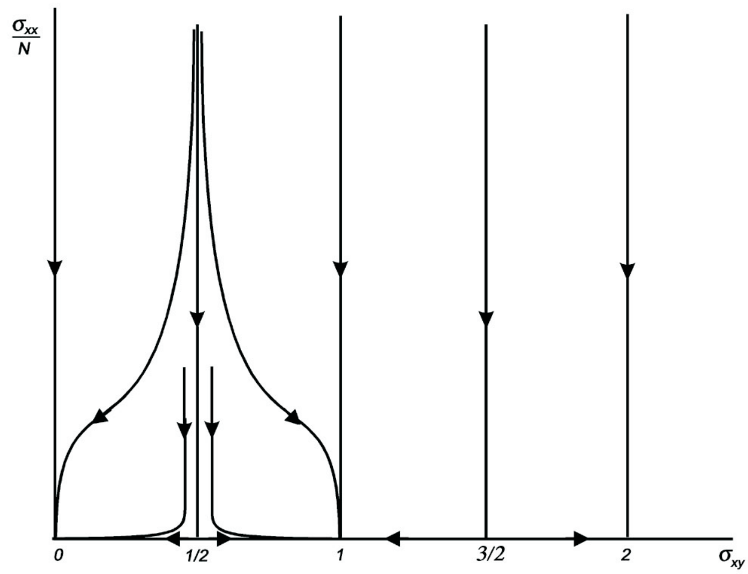


Figure 2.1: Large N renormalization group flow diagram for σ_{xx}/N and the Hall conductance $\sigma_{xy} = \theta/2\pi$. The arrows indicate the direction toward the infrared.

studied $O(3)$ case. [71] In fact, the large N expansion, as it now stands, is in many ways an onslaught on the many incorrect ideas and expectations in the field that are based on the historical papers on the subject. [37, 38, 36] These historical papers are not only in conflict with the basic features of the quantum Hall effect, but also present a fundamentally incorrect albeit misleading picture of the instanton vacuum concept as a whole.

Gapless excitations regained

One of the most important results of the large N expansion, the aforementioned *diverging correlation length* at $\theta = \pi$, has historically been overlooked. This is one of the main reasons why it is often assumed incorrectly that the excitations of the Grassmannian $U(m+n)/U(m) \times U(n)$ non-linear sigma model with $m, n \gtrsim 1$ always display a *gap*, also at $\theta = \pi$.

Notice that general arguments, based on 't Hooft's duality idea, [72] have already indicated that the theory at $\theta = \pi$ is likely to be different. The matter has important physical consequences because the lack of any *gapless* excitations in the problem (or, for that matter, the lack of *super universality* in non-linear sigma models) would seriously complicate the possibility of establishing a microscopic theory of the quantum Hall effect that is based on general topological principles.

It now has turned out that the large N expansion is one of the very rare examples where 't Hooft's idea of using *twisted boundary conditions* [72] can be worked out in great detail, thus providing an *explicit demonstration* of the existence of gapless excitations at $\theta = \pi$. [46] Besides all this, the large N expansion can also be used to demonstrate the general *nature* of the transition at $\theta = \pi$ which is otherwise is much harder to establish. For example, *complete scaling functions* have been obtained that set the stage for the *transitions* between adjacent *quantum Hall plateaus.* [46] In addition to this, *exact* expressions have been derived for the *distribution functions* of the mesoscopic conductance fluctuations in the problem. [66] These fluctuations render anomalously large (broadly distributed) as θ approaches π, a well known phenomenon in the theory of disordered metals. These results clearly indicate that the instanton vacuum generally displays richly complex physics that cannot be tapped if one is merely interested in the numerical value of the critical exponents alone.

The large N expansion is itself a good example of this latter statement. For example, the historical results on the large N expansion already indicated that the vacuum free energy with varying θ displays a *cusp* at $\theta = \pi$, i.e. a first order phase transition. This by itself is sufficient to establish the existence of a scaling exponent $\nu = 1/d$ with $d = 2$ denoting dimension of the system. [42] However, super universality as a whole remains invisible as long as one is satisfied with the merely heuristic arguments that historically have spanned the subject. [37, 38, 67] The discovery of a new aspect of the theory, the *massless chiral edge excitations*, was clearly necessary before the appropriate questions could be asked and *super universality* be finally established.

Given the new results on the large N expansion of the CP^{N-1} model, it may no longer be a complete surprise to know that the instanton vacuum at $\theta = \pi$ is *generically gapless*, independent of the number of field components in the theory. Since all members of the Grassmannian $U(m+n)/U(m) \times U(n)$ manifold are *topologically equivalent*, have important features in common such as *asymptotic freedom*, *instantons*, *massless chiral edge excitations* etc., it is imperative that the same basic phenomena are being displayed, independent of m and n. This includes of course the theory of actual interest, obtained by putting $m = n = 0$ (*replica limit*).

Unlike *super universality*, the details of the critical singularities at $\theta = \pi$ (*critical indices*) may in principle be different for different values of m and n. The situation is in this respect analogous to what happens to the classical Heisenberg ferromagnet in $2 + \epsilon$ spatial dimensions. Like in two dimensions, the basic physics is essentially the same for any value of m and n. The quantum critical behavior, however, strongly varies with a varying number of field components in the theory, each value of m and n representing a different *universality* class.

Instantons regained

Besides the *strong coupling* aspects of the instanton vacuum, the *effective action for massless chiral edge excitations* also provides a fundamentally new outlook on the *weak coupling* features of the theory that cannot be obtained in any different way. Topological excitations (*instantons*) have made a spectacularly novel entree, in the renormalization behavior of theory, especially after they have been totally mishandled and abused in the historical papers on the large N expansion. [37, 38, 67]

Within the recently established renormalization theory [46] of the CP^{N-1} model with large N (see Fig. 2.1), *instantons* emerge as non-perturbative topological objects

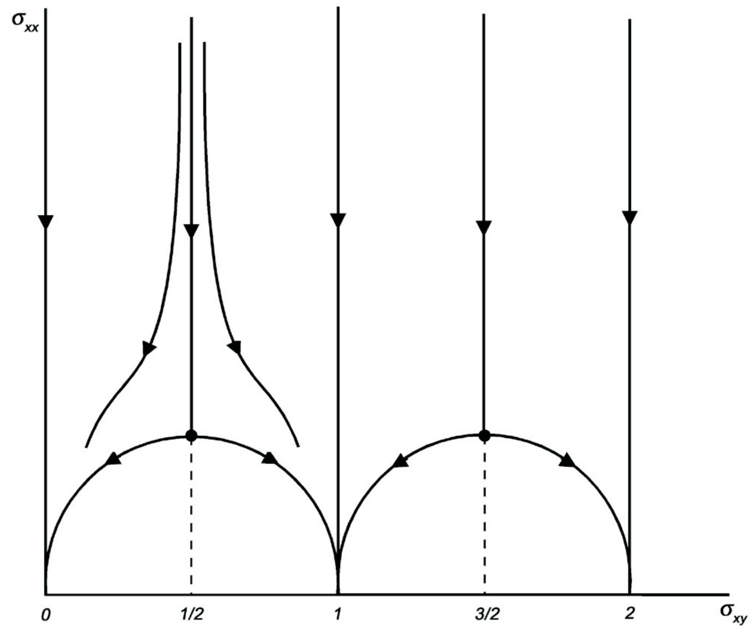


Figure 2.2: Renormalization group flow diagram for the conductances. The arrows indicate the scaling toward the infrared.

that facilitate the *cross-over* between the *Goldstone phase* at weak coupling or short distances ($\sigma_{xx} \gg 1$), and the super universal *strong coupling phase* of the instanton vacuum ($\sigma_{xx} \ll 1$) that generally appears in the limit of much larger distances only. A detailed knowledge of *instanton effects* is generally important since it provides a fundamentally new concept that the theory of ordinary perturbative expansions could never give, namely θ *renormalization* or, equivalently, the renormalization of the Hall conductance σ_{xy}.

The concept of θ renormalization originally arose in a series of detailed papers on instantons, based on the *background* field methodology, that were primarily aimed at a microscopic understanding of the quantum Hall effect. [20, 21] Until to date these pioneering papers have provided most of our insights into the singularity structure of the Grassmannian $U(m+n)/U(m)\times U(n)$ theory at $\theta = \pi$, in particular the case where the number of field components is 'small', $0 \leqslant m, n \lesssim 1$. Under these circumstances the instanton vacuum at half-integer values of σ_{xy} (or $\theta = \pi$) develops a critical fixed point with a *finite* value of σ_{xx} of order unity (see Fig. 2.2). This indicates that transition at $\theta = \pi$ becomes a true *second order* quantum phase transition with a non-trivial critical index ν that changes continuously with varying values of m and n in the range $0 \leqslant m, n \lesssim 1$. This situation is distinctly different from the overwhelming majority of Grassmannian non-linear sigma models with $m, n \gtrsim 1$ for which the scaling diagram is likely to be the same as the one found in the large N expansion (see Fig. 2.1). In that case one expects a *first order* phase transition but

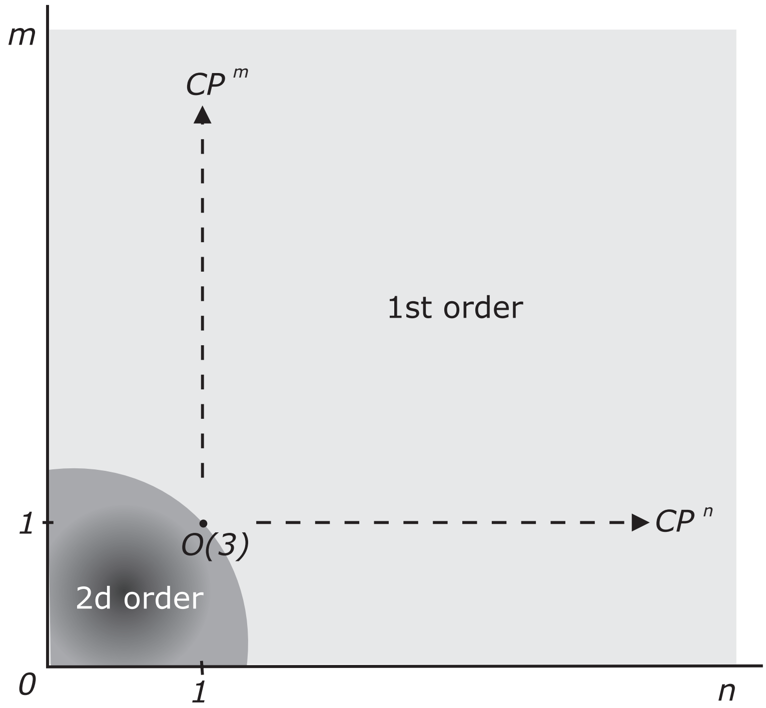


Figure 2.3: Nature of the transition at $\theta = \pi$ for different values of m and n.

with a *diverging* correlation length and a *fixed* exponent $\nu = 1/2$.

The instanton vacuum with $m = n = 1$ (the $SU(2)/U(1)$ or O(3) model) is in many ways special. This case is likely to be on the interface between a *large N-like* scaling diagram (see Fig. 2.1) with a *first order* transition at $\theta = \pi$, and an *instanton driven* scaling diagram (see Fig. 2.2) where the transition is of *second order*. The expected m and n dependence of quantum criticality is illustrated in Fig. 2.3 which is the main topic of the present chapter.

2.1.5 Quantum phase transitions

As is well known, the instanton angle θ in replica field theory was originally discovered in an attempt to resolve the fundamental difficulties of the scaling theory of Anderson localization in dealing with the quantum Hall effect. [18] However, it was not until the first experiments [23] on the *plateau transitions* had been conducted that the prediction of *quantum criticality* in the quantum Hall systems [22] became a well recognized and extensively pursued research objective in physics.

Quantum phase transitions in disordered systems are in many respects quite unusual, from ordinary critical phenomena point of view. For example, such unconventional phenomena like *multifractality* of the density fluctuations are known to appear as a peculiar aspect of the theory in the *replica limit* $m = n = 0$. [73] These subtle aspects of disordered systems primarily arose from the non-linear sigma model approach to the Anderson localization problem (*mobility edge problem*) in $2+\epsilon$ spatial dimensions. [74, 75]

The instanton vacuum theory of the quantum Hall effect essentially predicts that the *plateau transitions* in the two dimensional electron gas behave in all respects like the *metal-insulator transition* in $2+\epsilon$ dimensions. The reduced dimensionality of the quantum Hall system offers a rare opportunity to perform numerical work on the *mobility edge problem* and extract accurate results on quantum critical behavior. By now there exists an impressive stock of numerical data on the *critical indices* of the plateau transitions, including the correlation or *localization* length exponent (ν), [76, 77, 78, 79, 80, 81] the *multifractal* $f(\alpha)$ spectrum [82, 83, 84, 85, 86, 87] and even the leading *irrelevant* exponent (y_σ) in the problem. [88, 89]

Quantitative assessments

It is important to know that the *laboratory* experiments and later the *numerical* simulations on the *plateau transitions* in the quantum Hall regime have primarily been guided and motivated by the renormalization group ideas that were originally obtained on the basis of the θ parameter replica field theory as well as instanton calculus. [22, 20, 21] In addition to this, the more recent discovery of *super universality* in non-linear sigma models, along with the completely revised insights in the large N expansion, has elucidated the much sought after *strong coupling* features of the instanton vacuum, notably the *quantum Hall effect* itself, that previously remained concealed. [46] Both these strong coupling features and the renormalization group results based on instantons have put the theory of an instanton vacuum in a novel physical perspective. Together they provide the complete conceptual framework that is necessary for a detailed understanding of the quantum Hall effect as well as the θ dependence in the Grassmannian $U(m+n)/U(m)\times U(n)$ non-linear sigma model, for all non-negative values of m and n.

In Ref. [46] it has been already presented rough outlines on how the theory manages to interpolate between a *large N-like* scaling diagram for large values of m, n (Fig. 2.1) and an *instanton - driven* renormalization behavior for small m, n as indicated in Fig. 2.2. At present we take the theory several steps further and extend the instanton methodology in several ways. Our main objective is to make detailed predictions on the *quantum critical* behavior of the theory at $\theta = \pi$ with varying values of m and n. We benefit from the fact that this quantum critical behavior is bounded by the theory in the replica limit ($m = n = 0$) for which the aforementioned numerical data are available, and the distinctly different $O(3)$ non-linear sigma model ($m = n = 1$) for which the critical indices are known exactly. A detailed comparison between our general results and those known for specific examples should therefore provide a stringent and interesting test of the fundamental significance of instantons in the problem.

Our most important results are listed in Table 2.3 where we compare the critical exponents of the theory with $m = n = 0$ with those obtained from numerical simulations on the electron gas. These results clearly demonstrate the validity of a general statement made in Ref. [66, 46] which says that the fundamental significance of the instanton gas is primarily found in the renormalization behavior of the theory or, equivalently, the effective action for chiral edge excitations. This leads to a conceptual understanding of the non-perturbative aspects of the theory that cannot be obtained in any different manner.

Outline of this chapter

We start out in Section 2.2 with an introduction to the formalism, a brief summary of the effective action procedure for massless chiral edge excitations as well as a few comments explaining the super universal features of the instanton vacuum.

The bulk of this chapter mainly follows the formalism that was introduced in the original papers on instantons. [20, 21] However, the main focus at present is on several important aspects of the theory that previously remained unresolved. The first aspect concerns the ambiguity in the numerical factors that arises in the computation of the instanton determinant. These numerical factors are vitally important since they eventually enter into the renormalization group β functions and, hence, determine the critical fixed point properties of the theory. In Section 2.4 we show how the theory of observable parameters can be used to actually resolve this problem. The final expressions for the β functions that we obtain are *universal* in the sense that they are independent of the specific regularization scheme that is being used.

Secondly, the procedure so far did not include the effect of *mass terms* in the theory and, hence, the multifractal aspects of the quantum phase transition have not yet been investigated. The technical difficulties associated with mass terms are quite notorious, however. These have historically resulted in the construction of highly non-trivial extensions of the methodology such as working with *constrained instantons*. [45] We briefly introduce the subject matter in Section 2.3 and illustrate the main ideas by means of explicit examples.

It has sofar not been obvious, however, whether the concept of a *constrained instanton* is any useful in the development a quantum theory. In Ref. [31] it was pointed out that mass terms in the theory generally involve a different *metric tensor* than the one that naturally appears in the harmonic oscillator problem, i.e. their geometrical properties are incompatible. These and other complications are avoided in the methodology of *spatially varying masses* which is based on the various tricks introduced by 't Hooft in evaluating instanton determinants. [44] In Section 2.4 we elaborate further on the methodology of Ref. [31] and show that the idea of *spatially varying masses* generally facilitates explicit computations and yields more relevant results.

In Section 2.8 we present the detailed predictions of quantum criticality as a function of m and n and make a comparison with the results known from other sources. We end this chapter with a conclusion, Section 2.9.

2.2 Formalism

2.2.1 Non-linear sigma model

First we recall the non-linear sigma model defined on the Grassmann manifold $G/H = SU(m+n)/S(U(m) \times U(n))$ and in the presence of the θ term. We prefer to work in the context of the quantum Hall effect since this provides a clear *physical* platform for discussing the fundamental aspects of the theory that have previously remained unrecognized. It is easy enough, however, to make contact with the conventions and notations of more familiar models in quantum field theory and statistical mechanics such as the $O(3)$ formalism, obtained by putting $m = n = 1$, and the CP^{N-1} model which is obtained by taking $m = N - 1$ and $n = 1$.

The theory involves matrix field variables $Q(\mathbf{r})$ of size $(m+n)\times(m+n)$ that obey the non-linear constraint

$$Q^2(\mathbf{r}) = 1_{m+n}. \tag{2.2.1}$$

A convenient representation in terms of ordinary unitary matrices $T(\mathbf{r})$ is obtained by writing

$$Q(\mathbf{r}) = T^{-1}(\mathbf{r})\Lambda T(\mathbf{r}), \qquad \Lambda = \begin{pmatrix} 1_m & 0 \\ 0 & -1_n \end{pmatrix}. \tag{2.2.2}$$

The action describing the low energy dynamics of the two dimensional electron system subject to a static, perpendicular magnetic field is given by [19]

$$S = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 + \frac{\sigma_{xy}}{8}\int d\mathbf{r}\,\mathrm{tr}\,\varepsilon_{ab}Q\nabla_a Q\nabla_b Q + \pi\omega\rho_0\int d\mathbf{r}\,\mathrm{tr}\,\Lambda Q. \tag{2.2.3}$$

Here the quantities σ_{xx} and σ_{xy} represent the *meanfield* values for the *longitudinal* and *Hall* conductances respectively. The ρ_0 denotes the *density* of electronic levels in the bulk of the system, ω is the *external frequency* and $\varepsilon_{ab} = -\varepsilon_{ba}$ is the antisymmetric tensor.

2.2.2 Boundary conditions

The second term in Eq. (2.2.3) defines the topological invariant $\mathcal{C}[Q]$ which can also be expressed as a one dimensional integral over the edge of the system, [20]

$$\mathcal{C}[Q] = \frac{1}{16\pi i}\int d\mathbf{r}\,\mathrm{tr}\,\varepsilon_{ab}Q\nabla_a Q\nabla_b Q = \frac{1}{4\pi i}\oint dx\,\mathrm{tr}\,T\partial_x T^{-1}\Lambda. \tag{2.2.4}$$

Here, $\mathcal{C}[Q]$ is integer valued provided the field variable Q equals a constant matrix at the edge. It formally describes the mapping of the Grassmann manifold onto the plane following the homotopy theory result

$$\pi_2\,(G/H) = \pi_1\,(H) = \mathbb{Z}. \tag{2.2.5}$$

To study the σ_{xy} dependence of the theory it is in many ways natural to put $\omega = 0$ in Eq. (2.2.3) and let the infrared of the theory be regulated by the finite system size L. The conventional way of defining the θ vacuum is as follows

$$Z = \int_{\partial V} \mathcal{D}Q \exp\left[-\frac{1}{g}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 + i\theta\,\mathcal{C}[Q]\right], \tag{2.2.6}$$

where the subscript ∂V indicates that the functional integral has to be performed with $Q(\mathbf{r})$ kept fixed and constant at the boundary, say $Q = \Lambda$. Under these circumstances the topological charge $\mathcal{C}[Q]$ is strictly integer quantized. The parameter θ in Eq. (2.2.6) is equal to $2\pi\sigma_{xy}$ modulo 2π and the coupling constant g is identified as $8/\sigma_{xx}$,

$$\theta = 2\pi\sigma_{xy} \bmod(2\pi), \qquad g = \frac{8}{\sigma_{xx}}. \tag{2.2.7}$$

The theory of Eq. (2.2.6), as it stands, is one of the rare examples of an asymptotically free field theory with a vanishing mass gap. What has remarkably emerged over the years is that the theory at $\theta = \pi$ develops a *gapless* phase that generally can be associated with the *transitions* between adjacent quantum Hall plateaus. This is unlike the theory with $\theta = 0, 2\pi$ where the low energy excitations are expected to display a *mass gap*. The significance of the theory in terms of quantum Hall physics becomes all the more obvious if one recognizes that the mean field parameter σ_{xy} for strong magnetic fields is precisely equal to the filling fraction (ν_f) of the disordered Landau bands

$$\sigma_{xy} = \nu_f. \tag{2.2.8}$$

This means that the plateau transitions occur at *half-integer* filling fractions $\nu_f = k + 1/2$ with integer k. On the other hand, $\theta = 0, 2\pi$ corresponds to *integer* filling fractions $\nu_f = k$ which generally describe the center of the quantum Hall plateaus.

As was already mentioned in the introduction, the physical objectives of the quantum Hall effect have been - from early onward - in dramatic conflict with the ideas and expectations with which the θ parameter in quantum field theory was originally perceived. Such fundamental aspects like the existence of *robust topological quantum numbers*, for example, have previously been unrecognized. This is just one of the reasons why the quantum Hall effect primarily serves as an outstanding laboratory where the controversies in quantum field theory can be explored and investigated in detail.

Massless chiral edge excitations

It has turned out that the theory of Eq. (2.2.6) is not yet the complete story. By *fixing* the boundary conditions in this problem (or by discarding the *edge* all together) one essentially leaves out fundamental pieces of physics, the *massless chiral edge excitations*, that eventually will put the strong coupling problem of an instanton vacuum in a novel perspective.

To see how the physics of the edge enters into the problem we consider the case where the Fermi energy of the electron gas is located in an energy gap (*Landau gap*) between adjacent Landau bands. This is represented by Eq. (2.2.3) by putting $\sigma_{xx} = \rho_0 = 0$ and $\sigma_{xy} = k$, i.e. the *meanfield* value of the Hall conductance is an integer k (in units of e^2/h) and precisely equal to the number of completely filled Landau bands in the system. Eq. (2.2.3) can now be written as follows

$$\begin{aligned} S_{\text{edge}}[Q] &= 2\pi i k\, \mathcal{C}[Q] + \pi\omega\rho_{\text{edge}} \oint dx\, \operatorname{tr} Q\Lambda \\ &= \oint dx\, \operatorname{tr}\left[\frac{k}{2} T\partial_x T^{-1}\Lambda + \pi\omega\rho_{\text{edge}} Q\Lambda\right]. \end{aligned} \tag{2.2.9}$$

We have added a symmetry breaking term proportional to ρ_{edge} indicating that although the Fermi energy is located in the Landau gap, there still exists a finite density of edge levels (ρ_{edge}) that can carry the Hall current. Eq. (2.2.9) is exactly solvable and describes long ranged (*critical*) correlations along the edge of the system. Some important examples of edge correlations are given by [34]

$$\begin{aligned} \langle Q \rangle_{\text{edge}} &= \Lambda, \\ \langle Q^{\alpha\beta}_{+-}(x) Q^{\beta\alpha}_{-+}(x') \rangle_{\text{edge}} &= 4\vartheta(x-x') \exp\left[-\frac{4\pi\rho_{\text{edge}}\omega}{k}(x-x') \right]. \end{aligned} \tag{2.2.10}$$

Here the expectation is with respect to the one dimensional theory of Eq. (2.2.9). These results are the same for all m and n, indicating that the massless chiral edge excitations are a generic feature of the instanton vacuum concept.

Eqs (2.2.6) and (2.2.9) indicate that field configurations with an *integral* and *fractional* topological charge $\mathcal{C}[Q]$ describe fundamentally different physics and have fundamentally different properties. In the following Sections we show in a step by step fashion how Eq. (2.2.9) generally appears as the *fixed point action* of the strong coupling phase.

2.2.3 Effective action for the edge

We specialize from now onward to systems with an edge. The main problem next is to see how in general we can separate the *bulk* pieces of the action from those associated with the *edge*. The resolution of this problem provides fundamental information on the low energy dynamics of the system that is intimately related to the Kubo formalism of the conductances. At the end this leads to the much sought after Thouless' criterion for the existence of robust topological quantum numbers that explain the precision and observability of the quantum Hall effect.

Bulk and edge field variables

At the edge an arbitrary matrix Q can be written as

$$Q = t^{-1}_{\text{edge}} Q_0 t_{\text{edge}} \tag{2.2.11}$$

where $Q_0 = T_0^{-1} \Lambda T_0$ with T_0 being an arbitrary $U(m) \times U(n)$ invariant gauge at the edge, i.e. Q_0 satisfies the *spherical boundary conditions*

$$Q_0 \Big|_{\text{edge}} = \Lambda. \tag{2.2.12}$$

The matrix field t_{edge} describes the *fluctuations* about these special spherical boundary conditions. Next we continue the matrix field t_{edge} from the edge into the bulk according to the following prescription. For a given matrix field t_{edge}, we take the matrix t which obeys the classical equations of motion in the bulk and the following boundary conditions

$$t \Big|_{\text{edge}} = t_{\text{edge}}. \tag{2.2.13}$$

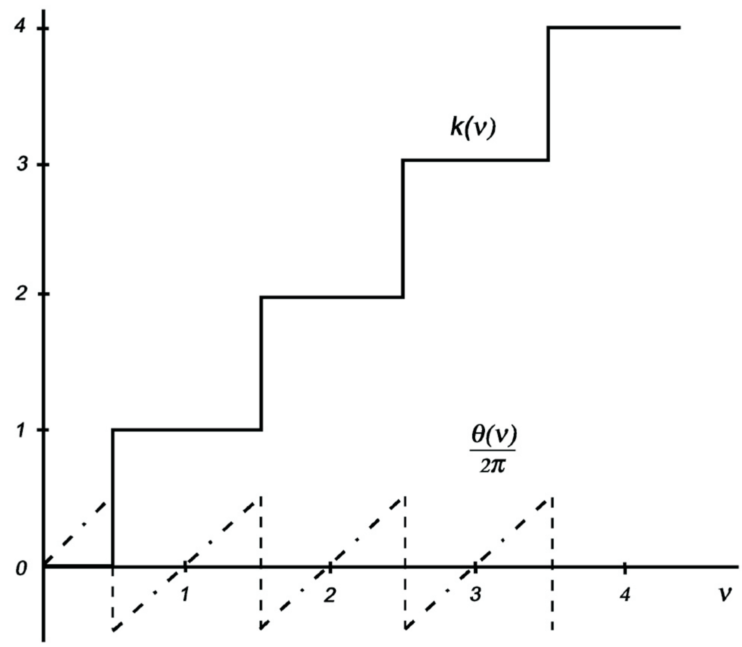


Figure 2.4: The quantity $\sigma_{xy} = \nu_f$ is the sum of a *quantized* edge part $k(\nu_f)$ and an *unquantized* bulk part $\theta(\nu_f)$.

Since the matrix t is fixed for a given t_{edge}, it is convenient to introduce a change of variables

$$Q = t^{-1} Q_0 t. \tag{2.2.14}$$

It is easy to see that the *topological charge* $\mathcal{C}[Q]$ can be written as the sum of two separate pieces

$$\mathcal{C}[Q] = \mathcal{C}[t^{-1} Q_0 t] = \mathcal{C}[Q_0] + \mathcal{C}[q], \qquad q = t^{-1} \Lambda t. \tag{2.2.15}$$

Here, the first part $\mathcal{C}[Q_0]$ is *integer* valued whereas the second part $\mathcal{C}[q]$ describes a *fractional* topological charge,

$$\mathcal{C}[Q_0] = k, \qquad -\frac{1}{2} < \mathcal{C}[q] \leqslant \frac{1}{2}. \tag{2.2.16}$$

It is easy to see that the action for the edge, Eq. (2.2.9), only contains the field variable t or q which we shall refer to as the *edge* fields in the problem. Similarly, the Q_0 are the only matrix field variables that enter into the usual definition of the θ vacuum, Eq. (2.2.6). We will refer to them as the *bulk* field variables.

Bulk and edge observables

Next, from Eqs (2.2.6) and (2.2.9) we infer that the bare or meanfield Hall conductance σ_{xy} should in general be split into a *bulk* piece $\theta(\nu_f)$ and a distinctly different *edge* piece $k(\nu_f)$ as follows (see also Fig. 2.4)

$$\sigma_{xy} = \nu_f = \frac{\theta(\nu_f)}{2\pi} + k(\nu_f). \tag{2.2.17}$$

Here, $k(\nu_f)$ is an integer whereas $\theta(\nu_f)$ is restricted to be in the interval

$$-\pi < \theta(\nu_f) \leqslant \pi. \tag{2.2.18}$$

The new symbol ν_f indicates that the meanfield parameter σ_{xy} is the same as the *filling fraction* of the Landau bands. Eq. (2.2.17) can also be obtained in a more natural fashion and it actually appears as a basic ingredient in the microscopic derivation of the theory. [90]

We use the results to rewrite the topological piece of the action as follows

$$2\pi i\sigma_{xy}\mathcal{C}[Q] = i(\theta(\nu_f)+2\pi k(\nu_f))(\mathcal{C}[q]+\mathcal{C}[Q_0]) = 2\pi ik(\nu_f)\mathcal{C}[q]+i\theta(\nu_f)\mathcal{C}[Q]. \tag{2.2.19}$$

In the second equation we have left out the term that gives rise to unimportant phase factor. Finally we arrive at the following form of the action

$$S = S_{\text{bulk}}[t^{-1}Q_0t] + S_{\text{edge}}[q], \tag{2.2.20}$$

where

$$S_{\text{bulk}}[Q] = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\text{tr}(\nabla Q)^2 + i\theta(\nu_f)\mathcal{C}[Q] \tag{2.2.21}$$

$$S_{\text{edge}}[q] = 2\pi ik(\nu_f)\,\mathcal{C}[q]. \tag{2.2.22}$$

Response formulae

After these preliminaries we come to the most important part of this Section, namely the definition of the *effective action for chiral edge modes* S_{eff}. This is obtained by formally eliminating the *bulk* field variable Q_0 from the theory

$$S_{\text{eff}}[q] = S_{\text{edge}}[q] + S'[q], \qquad \exp S'[q] = \int_{\partial V} \mathcal{D}[Q_0]\exp S_{\text{bulk}}[t^{-1}Q_0t]. \tag{2.2.23}$$

The subscript ∂V reminds us of the fact that the functional integral has to be performed with a fixed value $Q_0 = \Lambda$ at the edge of the system. It is instructive to write the $S_{\text{bulk}}[t^{-1}Q_0t]$ as follows

$$S_{\text{bulk}}[t^{-1}Q_0t] = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\text{tr}(\nabla Q_0 + [\mathbf{A}_t, Q_0])^2 + i\theta(\nu_f)\mathcal{C}[Q_0] + i\theta(\nu_f)\mathcal{C}[q], \tag{2.2.24}$$

where we have introduced $\mathbf{A}_t = t\nabla t^{-1}$. Thence, the $S'[q]$ can be formally obtained as an infinite series in powers of the "vector potential" $\mathbf{A}_t$

$$S'[q] = i\theta(\nu_f)\mathcal{C}[q] + \ln\left\{1 + \sum_{s=0}^{\infty}\frac{(-1)^s\sigma_{xx}^s}{2^{3s}s!}\left\langle\left[\int d\mathbf{r}\,\text{tr}\left([\mathbf{A}_t, Q_0]^2 + 4\mathbf{A}_tQ_0\nabla Q_0\right)\right]^s\right\rangle\right\}. \tag{2.2.25}$$

Here the average is with the respect to the action $S_{\rm bulk}[Q_0]$. From symmetry considerations alone it is easily established that $S'[q]$ must be of the general form

$$S'[q] = -\frac{\sigma'_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla q)^2 + i\theta'\mathcal{C}[q]. \tag{2.2.26}$$

We have omitted the higher dimensional terms which generally describe the properties of the electron gas at *mesoscopic* scales. [91] The most important feature of this result is that the quantities $\sigma'_{xx} = \sigma'_{xx}(L)$ and $\theta' = \theta'(L)$ can be identified with the Kubo formulae for the *longitudinal* and *Hall* conductances respectively (L denoting the linear dimension of the system). Notice that these quantities are by definition a measure for the *response* of the bulk of the system to an infinitesimal change in the boundary conditions (on the matrix field variable Q_0).

At the same time we can regulate the infrared of the system in a different manner by introducing $U(m)\times U(n)$ invariant mass terms in Eqs (2.2.21)-(2.2.22)

$$S_{\rm edge}[q] \to S_{\rm edge}[q] + \pi\omega_{\rm edge}\rho_{\rm edge}\oint dx\,\mathrm{tr}\,\Lambda q, \tag{2.2.27}$$

$$S_{\rm bulk}[t^{-1}Q_0 t] \to S_{\rm bulk}[t^{-1}Q_0 t] + \pi\omega_0\rho_0\int d\mathbf{r}\,\mathrm{tr}\,\Lambda Q_0. \tag{2.2.28}$$

The different symbols $\omega_{\rm edge}$ and ω_0 indicate that the frequency ω plays a different role for the *bulk* fields and *edge* fields respectively. Notice that the response parameters σ'_{xx} and σ'_{xy} in Eq. (2.2.26), for L large enough, now depend on frequency ω_0 rather than L.

$$\sigma'_{xx} \to \sigma'_{xx}(\omega_0), \qquad \sigma'_{xy} \to \sigma'_{xy}(\omega_0) = k(\nu_f) + \frac{\theta'(\omega_0)}{2\pi}. \tag{2.2.29}$$

Background field methodology

Let us next go back to the background field methodology and notice the subtle differences with the effective action procedure as considered here. In this methodology we consider a slowly varying but fixed background matrix field b that is applied directly to the original theory of Eq. (2.2.3),

$$\begin{aligned}\exp S_{\rm eff}[b^{-1}\Lambda b] &= \int \mathcal{D}Q\exp\left(S_\sigma[b^{-1}Qb] + \pi\omega_0\rho_0\int d\mathbf{r}\,\mathrm{tr}\,Q\Lambda\right),\\ S_\sigma[Q] &= -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 + 2\pi i\sigma_{xy}\,\mathcal{C}[Q].\end{aligned} \tag{2.2.30}$$

Here, $S_{\rm eff}$ can again be written in the following general form

$$S_{\rm eff}[q_b] = -\frac{\sigma'_{xx}(\omega_0)}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla q_b)^2 + 2\pi i\sigma'_{xy}(\omega_0)\mathcal{C}[q_b], \tag{2.2.31}$$

where now $q_b = b^{-1}\Lambda b$. Notice that an obvious difference with the situation before is that the functional integral in Eq. (2.2.30) is now performed for an arbitrary matrix field Q whereas in Eq. (2.2.23) the Q_0 is always restricted to have $Q_0 = \Lambda$ at the edge. However, as long as one works with a finite frequency ω_0 the boundary conditions on the Q_0 field are immaterial and the response parameters in Eq. (2.2.29) and those in Eq. (2.2.31) should be identically the same. As an important check on these statements we consider $\sigma_{xx} = 0$ and $\sigma_{xy} = k$, i.e. S_σ equals the action for the edge S_{edge}. In this case Eq. (2.2.31) is obtained as follows

$$\begin{aligned} \exp S_{\text{eff}}[q_b] &= \int \mathcal{D}Q \exp\left[\oint dx \left(\frac{k}{2} \operatorname{tr} Tb\partial_x b^{-1} T^{-1} \Lambda + \pi\omega\rho_{\text{edge}} \operatorname{tr} Q\Lambda\right)\right] \\ &= \left(\int \mathcal{D}Q \exp S_{\text{edge}}[Q]\right) \exp\left[\frac{k}{2}\oint dx \operatorname{tr} b\partial_x b^{-1} \langle Q\rangle_{\text{edge}}\right]. \quad (2.2.32) \end{aligned}$$

Using Eq. (2.2.10) we can write, discarding constants,

$$S_{\text{eff}}[q_b] = 2\pi ik\ \mathcal{C}[q_b] \quad (2.2.33)$$

Comparing Eqs (2.2.31) and (2.2.33) we see that $\sigma'_{xx} = 0$ and $\sigma'_{xy} = k$ as it should be. Notice that we obtain essentially the same result if instead of integrating we fix the $Q = \Lambda$ at the edge.

Eq. (2.2.32) clearly demonstrates why critical edge correlations should be regarded as a fundamental aspect of the instanton vacuum concept. If, for example, S_{edge} were to display gapped excitations at the edge then we certainly would have $\langle Q\rangle_{\text{edge}} = 0$ in Eq. (2.2.32) and instead of Eq. (2.2.33) we would have had a vanishing Hall conductance!

This, then, would be serious conflict with the quantum Hall effect which says that the quantization of the Hall conductance is in fact a *robust* phenomenon.

Thouless criterion

We next show that a *Thouless criterion* [69] for the quantum Hall effect can be obtained directly, as a corollary of the aforementioned effective action procedure. For this purpose we notice that if the system develops a mass gap or a finite correlation length ξ in the bulk, then the theory of S_{bulk}, Eq. (2.2.21), should be insensitive to any changes in the boundary conditions, provided the system size L is large enough. Under these circumstances the response quantities $\sigma'_{xx}(L)$ and $\theta'(L)$ in Eq. (2.2.26) should vanish. In terms of the conductances we can write

$$\sigma'_{xx}(L) = \mathcal{O}\left(e^{-L/\xi}\right), \qquad \sigma'_{xy}(L) = k(\nu_f) + \frac{\theta'(L)}{2\pi} = k(\nu_f) + \mathcal{O}\left(e^{-L/\xi}\right). \quad (2.2.34)$$

This important result indicates that the *quantum Hall effect* is in fact a *universal, strong coupling* feature of the θ vacuum, independent of the number of field components m and n. The *fixed point action* of the quantum Hall state is generally given by the one dimensional action [34]

$$S_{\text{eff}}[q] = \oint dx \operatorname{tr}\left[\frac{k(\nu_f)}{2} t\partial_x t^{-1}\Lambda + \pi\omega\rho_{\text{edge}} q\Lambda\right], \quad (2.2.35)$$

which is none other than the aforementioned action for *massless chiral edge excitations.*

In summary we can say that the *background field* methodology that was previously introduced for renormalization group purposes alone, now gets a new appearance in the theory and a fundamentally different meaning in term of the *effective action* for massless edge excitations. This effective action procedure emerges from the theory itself and, unlike the background field methodology, it provides the much sought after *Thouless criterion* which associates the exact quantization of the Hall conductance with the insensitivity of the bulk of the system to changes in the boundary conditions.

Conductance fluctuations and level crossing

The definition of the effective action, Eq. (2.2.23), implies that the exact expression for S_{eff} is invariant under a change in the matrix field $t \to Ut$, i.e. the replacement

$$S_{\text{eff}}[t^{-1}\Lambda t] \to S_{\text{eff}}[t^{-1}U^{-1}\Lambda Ut] \tag{2.2.36}$$

should leave the theory invariant. Here, the $U = U(\mathbf{r})$ represents a unitary matrix field with an *integer* valued topological charge, i.e. $U(\mathbf{r})$ reduces to an arbitrary $U(m) \times U(n)$ gauge at the edge of the system.

Notice that the expressions for S_{eff}, Eqs (2.2.23) and (2.2.26), in general violate the invariance under Eq. (2.2.36). This is so because we have expressed $S'[q]$ to lowest order in a series expansion in powers of the derivatives acting on the q field. If the response parameters $\sigma'_{xx}(L)$ and $\theta'(L)$ are finite then the invariance under Eq. (2.2.36) is generally *broken* at each and every order in the derivative expansion. The invariance is truly *recovered* only after the complete series is taken into account, to infinite order. This situation typically describes a *gapless* phase in the theory. The infinite series (or at least an infinite subset of it) can in general be rearranged and expressed in terms of *probability distributions* for the response parameters $\sigma'_{xx}(L)$ and $\theta'(L)$. Quantum critical points, for example, seem to generically display *broadly* distributed response parameters. The large N expansion has provided, once more, a lucid and exact example of these statements. [46]

On the other hand, if the theory displays a *mass gap* then each term in the series for $S'[q]$ should vanish, i.e. the response parameters are all exponentially small in the system size. It is therefore possible to employ Eq. (2.2.36) as an alternative criterion for the quantum Hall effect, namely by demanding that the theory be invariant under Eq. (2.2.36) *order by order* in the derivative expansion. This criterion immediately demands that σ'_{xx} and θ' in Eq. (2.2.26) be zero and the effective action be given by Eq. (2.2.35). Equations (2.2.35) and (2.2.36) imply

$$S_{\text{eff}}[t^{-1}U^{-1}\Lambda Ut] = S_{\text{eff}}[t^{-1}\Lambda t] + 2\pi i k(\nu_f)\mathcal{C}[U^{-1}\Lambda U]. \tag{2.2.37}$$

Although the matrix U merely gives rise to a phase factor that can be dropped, physically it corresponds to an integer number n_e of (edge) electrons, equal to $k(\nu_f)\mathcal{C}[U^{-1}\Lambda U]$, that have crossed the Fermi level. Eq. (2.2.37) is therefore synonymous for the statement which says that the quantization of the Hall conductance $\sigma_{xy} = k(\nu_f)\left[e^2/h\right]$ is related to the quantization of flux $\Phi = \mathcal{C}[U^{-1}\Lambda U]\,[h/e]$ and quantization of charge $q_e = n_e\,[e]$ according to

$$q_e = \sigma_{xy}\Phi. \tag{2.2.38}$$

2.2.4 Physical observables

The response parameters σ'_{xx} and θ' as well as other observables, associated with the mass terms in the non-linear sigma model, are in many ways the most significant physical quantities in the theory that contain complete information on the low energy dynamics of the instanton vacuum. These quantities can all be expressed in terms of correlation functions of the *bulk* field variables Q_0 alone. In this Section we give a summary of the *physical observables* in the theory which then completes the theory of massless chiral edge excitations.

To start we introduce the following theory for the *bulk* matrix field variables (dropping the subscript "0" on the Q_0 from now onward)

$$Z = \int\limits_{\partial V} \mathcal{D}Q \exp\left(S_\sigma[Q] + S_h[Q] + S_a[Q] + S_s[Q]\right). \tag{2.2.39}$$

The subscript ∂V indicates, as before, that the Q is kept fixed at $Q = \Lambda$ at the edge of the system. S_σ stands for

$$S_\sigma[Q] = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 + \frac{\theta}{16\pi}\int d\mathbf{r}\,\mathrm{tr}\,\varepsilon_{ab}Q\nabla_a Q\nabla_b Q. \tag{2.2.40}$$

The quantities S_h, S_a and S_s are the $U(m)\times U(n)$ invariant mass terms that will be specified below.

Kubo formula

Explicit expressions for response quantities σ'_{xx} and θ' can be derived following the analysis of Ref. [21]. The following $U(m)\times U(n)$ invariant results have been obtained

$$\sigma'_{xx} = \sigma_{xx} + \frac{\sigma_{xx}^2}{16mnL^2}\int d\mathbf{r}d\mathbf{r}'\,\mathrm{tr}\langle Q(\mathbf{r})\nabla Q(\mathbf{r})Q(\mathbf{r}')\nabla Q(\mathbf{r}')\rangle, \tag{2.2.41}$$

$$\theta' = \theta - \frac{(m+n)\pi}{4mnL^2}\sigma_{xx}\int d\mathbf{r}\,\mathrm{tr}\langle \Lambda Q\varepsilon_{ab}r_a\partial_b Q\rangle + \frac{\pi\sigma_{xx}^2}{8mnL^2}\int d\mathbf{r}d\mathbf{r}'\,\mathrm{tr}\,\varepsilon_{ab}\langle Q(\mathbf{r})\nabla_a Q(\mathbf{r})Q(\mathbf{r}')\nabla_b Q(\mathbf{r}')\Lambda\rangle. \tag{2.2.42}$$

Here and from now onward the expectations are defined with respect to the theory of Eq. (2.2.39).

Mass terms

We shall be interested in traceless $U(m)\times U(n)$ invariant operators that are linear in Q (O_h) and bilinear in Q ($O_{s,a}$),

$$S_h[Q] = z_h\int d\mathbf{r}\,O_h[Q], \qquad S_{a,s}[Q] = z_{a,s}\int d\mathbf{r}\,O_{a,s}[Q]. \tag{2.2.43}$$

Here,

$$O_h[Q] = \mathrm{tr}\left[\Lambda - \frac{m-n}{m+n}1_{m+n}\right]Q. \tag{2.2.44}$$

The bilinear operators generally involve a *symmetric* and an *antisymmetric* combination [75]

$$O_{s,a}[Q] = \sum_{p,q}^{\alpha,\beta} K_{s,a}^{pq} \left[Q_{pp}^{\alpha\alpha} Q_{qq}^{\beta\beta} \pm Q_{pq}^{\alpha\beta} Q_{qp}^{\beta\alpha} \right], \tag{2.2.45}$$

where $K_{s,a}$ is given as a 2×2 matrix

$$K_{s,a} = \begin{pmatrix} -\dfrac{m}{n \pm 1} & 1 \\ 1 & -\dfrac{n}{m \pm 1} \end{pmatrix}. \tag{2.2.46}$$

These quantities permit us to define physical observables z'_h and $z'_{s,a}$ that are associated with the z_h and $z_{s,a}$ fields respectively. Specifically,

$$z'_h = z_h \frac{\langle O_h[Q] \rangle}{O_h[\Lambda]}, \quad z'_s = z_s \frac{\langle O_s[Q] \rangle}{O_s[\Lambda]}, \quad z'_a = z_a \frac{\langle O_a[Q] \rangle}{O_a[\Lambda]}. \tag{2.2.47}$$

The ratio on the right hand side merely indicates that the expectation value of the operators is normalized with respect to the classical value.

Observable and renormalized theories in $2+\epsilon$ dimensions

Brézin et.al. [92] originally showed that the non-linear sigma model in $2+\epsilon$ dimensions generally involves a renormalization of the coupling constant or σ_{xx} and one renormalization associated with each of the operators O_i. Denoting the bare parameters of the theory by $\sigma_{xx} = 1/g_0$ and z_i^0 then the relation between the *bare* theory and *renormalized* theory g and z_i is given by

$$\sigma_{xx} = \frac{1}{g_0} = \frac{1}{g} \mu^\epsilon Z(g), \qquad z_i^0 = z_i Z_i(g) \tag{2.2.48}$$

with μ an arbitrary momentum scale. The functions $Z(g)$ and $Z_i(g)$ are usually fixed by the requirement that the theory be finite in ϵ. According to the minimum subtraction scheme, for example, one employs the Z and Z_i for the purpose of absorbing the pole terms in ϵ and nothing but the pole terms. However, it is well known that in order to be consistent with the infrared behavior of the theory the terms that are finite in ϵ can play an important role. Cross-over problems, for example, are treated incorrectly within the minimum subtraction scheme and usually involve a very specific choice of the functions Z that includes terms that are finite in ϵ. In this Section we show that the arbitrariness in the renormalization group is in general avoided if one employs the renormalizations Z and Z_i for the purpose of identifying the *renormalized* and *observable* theories. For simplicity we shall present the results for the theory in the presence of the operator O_h only. It is convenient to introduce a change of variables. Write

$$\sigma'_{xx} = \frac{1}{g'}, \qquad \sigma_{xx} = \frac{1}{g_0}, \tag{2.2.49}$$

and

$$z'_h = \frac{(h')^2}{g'}, \qquad z_h^0 = \frac{h_0^2}{g_0}. \tag{2.2.50}$$

Table 2.1: Coefficients of two-loop computation

G/H	a	A_1	A_2	B
$\frac{SO(m+n)}{S(O(m)\times O(n))}$	1	$m+n-2$	$2mn-m-n$	$4-2m-2n$
$\frac{SU(m+n)}{S(U(m)\times U(n))}$	2	$m+n$	$2(mn+1)$	0
$\frac{SP(m+n)}{SP(m)\times SP(n)}$	4	$m+n+1$	$2mn+\frac{m+n}{2}$	$m+n+1$

Since the g', h' fields have the same meaning for the various Grassmannian manifolds listed in Table 2.1 we shall from now onward work within the general G/H non-linear sigma model. Starting from the action

$$S = \frac{a}{2} S_\sigma[Q] + S_h[Q], \tag{2.2.51}$$

then a straightforward computation to order ϵ^2 of the observable quantities σ'_{xx} (see Eq. (2.2.41)) and z'_h (see Eq. (2.2.47)) yields the following results

$$\frac{1}{g'} = \frac{1}{g_0}\left(1 + A_1 \frac{g_0 h_0^\epsilon}{\epsilon} + \frac{1}{2}(A_2 - B)\frac{g_0^2 h_0^{2\epsilon}}{\epsilon} - \frac{1}{2} A_2 C g_0^2 h_0^{2\epsilon}\right), \tag{2.2.52}$$

$$\frac{(h')^2}{g'} = \frac{h_0^2}{g_0}\left(1 + \left(A_1 - \frac{B}{A_1}\right)\frac{g_0 h_0^\epsilon}{\epsilon} - \left(A_1 - \frac{B}{A_1}\right)\frac{B}{A_1}\frac{g_0^2 h_0^{2\epsilon}}{\epsilon^2}\left(1+\frac{\epsilon}{2}\right)\right). \tag{2.2.53}$$

Here, the coefficients A_1, A_2 and B are listed in Table 2.1 and C is a numerical constant. A factor $2\Gamma(1-\epsilon/2)(4\pi)^{-1-\epsilon/2}$ has been absorbed in a redefinition of the g_0 and g'.

The results of Eqs (2.2.52) and (2.2.53) have originally been used in Ref. [93] for the purpose of expressing the observable parameters in $2+\epsilon$ dimensions in terms of the *equations of state*. Here we shall point out a slightly different interpretation of these results which is obtained by recognizing that the h' field actually plays the role of *momentum scale* that is associated with the *observable* quantities g' or σ'_{xx} and z'_h. This observation permits one to identify the *observable* and *renormalized* theories in the following manner. First we employ Eqs (2.2.52) and (2.2.53) and eliminate the h_0 field in the dimensionless combination $g_0 h_0^\epsilon$ in favor of the induced momentum scale h'. The results can be written as

$$\frac{1}{g'} = \frac{1}{g_0}\left(1 + A_1 \frac{g_0 (h')^\epsilon}{\epsilon} + \frac{1}{2} A_2 \frac{g_0^2 (h')^{2\epsilon}}{\epsilon} - \frac{1}{2} A_2 C g_0^2 (h')^{2\epsilon}\right), \tag{2.2.54}$$

$$\frac{(h')^2}{g'} = \frac{h_0^2}{g_0}\left(1 + \left(A_1 - \frac{B}{A_1}\right)\frac{g_0 (h')^\epsilon}{\epsilon} - \left(A_1 - \frac{B}{A_1}\right)\frac{B}{A_1}\frac{g_0^2 (h')^{2\epsilon}}{\epsilon^2}\right). \tag{2.2.55}$$

As a second step we make use of Eq. (2.2.54) and eliminate the bare parameter g_0 in the combination $g_0(h')^\epsilon$ in favor of the g'. Introducing the dimensionless quantity

$$\bar{g} = g'(h')^\epsilon, \tag{2.2.56}$$

then we obtain from Eqs (2.2.54) and (2.2.55)

$$\frac{1}{g_0} = \frac{(h')^\epsilon}{\bar{g}} Z(\bar{g}), \tag{2.2.57}$$

$$\frac{h_0^2}{g_0} = \frac{(h')^{2+\epsilon}}{\bar{g}} Z_h(\bar{g}), \tag{2.2.58}$$

where to order $\bar{g}^2$ the Z and Z_h are given by

$$Z(\bar{g}) = 1 - A_1 \frac{\bar{g}}{\epsilon} - A_2(1-\epsilon C)\frac{\bar{g}^2}{2\epsilon}, \tag{2.2.59}$$

$$Z_h(\bar{g}) = 1 - \left(A_1 - \frac{B}{A_1}\right)\frac{\bar{g}}{\epsilon}. \tag{2.2.60}$$

Equations (2.2.57) and (2.2.58) provide a natural definition of the quantities $Z(g)$ and $Z_h(g)$ that appear in the expressions of the *renormalized* theory, Eqs (2.2.48). By fixing the renormalizations Z and Z_h according to Eqs (2.2.59) and (2.2.60) we obtain renormalization group β and γ functions in the usual manner

$$\beta(g) = \frac{dg}{d\ln\mu} = \frac{\epsilon g}{1 - g\dfrac{d\ln Z}{dg}} = \epsilon g - A_1 g^2 - g^3 A_2(1-\epsilon C), \tag{2.2.61}$$

$$\gamma_h(g) = -\frac{d\ln z_h}{d\ln\mu} = \beta(g)\frac{d}{dg}\ln Z_h(g) = -\left(A_1 - \frac{B}{A_1}\right) g + O(g^3). \tag{2.2.62}$$

Moreover, the choice of Eqs (2.2.57) and (2.2.58) implies that the observable theories at different momentum scales h' and h respectively can in general be expressed in terms of the β and γ functions according to

$$\bar{g} = g(h') = g(h) + \int_h^{h'} \frac{d\mu}{\mu}\beta(g) \tag{2.2.63}$$

$$z'_h = z_h(h') = z_h(h) - \int_h^{h'} \frac{d\mu}{\mu}\gamma_h(g) z_h(\mu). \tag{2.2.64}$$

The skeptical reader might want to explicitly verify the fact that the results of Eqs (2.2.63) and (2.2.64) are consistent with the original definition of the observable theory, Eqs (2.2.54) and (2.2.55). Starting from Eq. (2.2.54), for example, one proceeds by inserting $g_0 = h^{-\epsilon} g Z^{-1}(g)$ where g is now defined for momentum scale h. This leads to the following expression

$$\bar{g} = g(h') = g(h) + \mathcal{I}\left(g(h), \frac{h'}{h}\right), \tag{2.2.65}$$

where

$$\mathcal{I}\left(g(h), \frac{h'}{h}\right) = g\left[\left(\frac{h'}{h}\right)^{\epsilon} - 1\right] - A_1 \frac{g^2}{\epsilon^2}\left\{\epsilon + A_1 g\left[\left(\frac{h'}{h}\right)^{\epsilon} - 1\right]\right\}\left(\frac{h'}{h}\right)^{\epsilon} \times \left[\left(\frac{h'}{h}\right)^{\epsilon} - 1\right] - A_2 \frac{g^3}{2\epsilon}(1 - \epsilon C)\left(\frac{h'}{h}\right)^{\epsilon}\left[\left(\frac{h'}{h}\right)^{2\epsilon} - 1\right]. \quad (2.2.66)$$

Simple algebra next shows that to the appropriate order in g the following identity holds

$$\frac{d}{d\ln h}\mathcal{I}\left(g(h), \frac{h'}{h}\right) = -\beta(g(h)). \quad (2.2.67)$$

This means that Eqs (2.2.63) and (2.2.65) are indeed identical expressions. Similarly, by starting from Eq. (2.2.55), i.e.

$$z_h' = z_h^0\left(1 + \left(A_1 - \frac{B}{A_1}\right)\frac{g_0(h')^{\epsilon}}{\epsilon} - \left(A_1 - \frac{B}{A_1}\right)\frac{B}{A_1}\frac{g_0^2(h')^{2\epsilon}}{\epsilon^2}\right), \quad (2.2.68)$$

one obtains the following result for the observable quantity z_h'

$$z_h' = z_h(h') = z_h(h) + \mathcal{J}\left(g(h), \frac{h'}{h}\right) \quad (2.2.69)$$

where

$$\mathcal{J}\left(g(h), \frac{h'}{h}\right) = z_h(h)\left(A_1 - \frac{B}{A_1}\right)\frac{g}{\epsilon}\left[1 - \frac{B}{A_1}\frac{g}{\epsilon}\left(\frac{h'}{h}\right)^{\epsilon}\right]\left[\left(\frac{h'}{h}\right)^{\epsilon} - 1\right]. \quad (2.2.70)$$

Differentiating with respect to $\ln h$ leads to the following result

$$\frac{d}{d\ln h}\mathcal{J}\left(g(h), \frac{h'}{h}\right) = z_h(h)\gamma_h(g(h)) \quad (2.2.71)$$

which means that Eqs (2.2.64) and (2.2.69) are identically the same as well.

β and γ functions

The *observable parameters* of the previous sections facilitate a renormalization group study that can be extended to include the non-perturbative effects of instantons. We shall first recapitulate some of the results obtained from ordinary perturbative expansions. Let μ' denote the momentum scale associated with the observable theory then the quantities $\sigma_{xx}' = \sigma_{xx}(\mu')$, $z_i' = z_i(\mu')$ can be expressed in terms of the renormalization group β and γ functions according to (see previous section)

$$\sigma_{xx}' = \sigma_{xx}(\mu') = \sigma_{xx}(\mu_0) + \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu}\, \beta_\sigma(\sigma_{xx}) \quad (2.2.72)$$

$$z_i' = z_i(\mu') = z_i(\mu_0) - \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu}\, \gamma_i(\sigma_{xx}) z_i(\mu), \quad (2.2.73)$$

where [92]

$$\beta_\sigma(\sigma_{xx}) = \frac{m+n}{2\pi} + \frac{mn+1}{2\pi^2\sigma_{xx}} + \mathcal{O}(\sigma_{xx}^{-2}) \tag{2.2.74}$$

$$\gamma_h(\sigma_{xx}) = -\frac{m+n}{2\pi\sigma_{xx}} + \mathcal{O}(\sigma_{xx}^{-2}) \tag{2.2.75}$$

and [75]

$$\gamma_s(\sigma_{xx}) = -\frac{m+n+2}{2\pi\sigma_{xx}} + \mathcal{O}(\sigma_{xx}^{-2}) \tag{2.2.76}$$

$$\gamma_a(\sigma_{xx}) = -\frac{m+n-2}{2\pi\sigma_{xx}} + \mathcal{O}(\sigma_{xx}^{-2}). \tag{2.2.77}$$

The effects of instantons have been studied in great details in Refs [18, 19] where the idea of the θ renormalization was introduced. The main objective of the present chapter is to extend the results of Eqs (2.2.74)-(2.2.77) to include the effect of instantons. Recall that the γ_i functions are of very special physical interest. For example, the quantity γ_h should vanish in the limit $m, n \to 0$ indicating that the density of levels of the electron gas is in general unrenormalized. At the same time one expects the anomalous dimension γ_a to become positive as $m, n \to 0$ since it physically describes the singular behavior of the (inverse) participation ratio of the electronic levels. These statements serve as an important physical constraint that one in general should impose upon the theory. [75]

2.3 Instantons

2.3.1 Introduction

In the absence of symmetry breaking terms, the existence of finite action solutions (instantons) follows from the Schwartz inequality

$$\operatorname{tr}(\nabla_x Q \pm iQ\nabla_y Q)^2 \geqslant 0, \tag{2.3.1}$$

which implies

$$\frac{1}{8}\int d\mathbf{r}\, \operatorname{tr}(\nabla Q)^2 \geqslant 2\pi|\mathcal{C}[Q]|. \tag{2.3.2}$$

Matrix field configurations that fulfill inequality (2.3.2) as an equality are called *instantons*. The classical action becomes

$$S_\sigma^{\text{inst}} = -2\pi\sigma_{xx}|\mathcal{C}[Q]| + i\theta\mathcal{C}[Q]. \tag{2.3.3}$$

In this thesis we consider single instantons only with a topological charge $\mathcal{C}[Q] = \pm 1$. A convenient representation of the single instanton solution is given by [20, 21]

$$Q_{\text{inst}}(\mathbf{r}) = T^{-1}\Lambda_{\text{inst}}(\mathbf{r})T, \qquad \Lambda_{\text{inst}}(\mathbf{r}) = \Lambda + \rho(\mathbf{r}). \tag{2.3.4}$$

Here, the matrix $\rho_{pq}^{\alpha\beta}(\mathbf{r})$ has four non-zero matrix elements only

$$\rho_{11}^{11} = -\rho_{-1-1}^{11} = -\frac{2\lambda^2}{|z-z_0|^2+\lambda^2}, \qquad \rho_{1-1}^{11} = \bar{\rho}_{-11}^{11} = \frac{2\lambda(z-z_0)}{|z-z_0|^2+\lambda^2}, \tag{2.3.5}$$

with $z = x + iy$. The quantity z_0 describes the *position* of the instanton and λ is the scale size. These parameters, along with the global unitary rotation $T \in U(m+n)$, describe the manifold of the single instanton. The *anti-instanton* solution with topological charge $\mathcal{C}[Q] = -1$ is simply obtained by complex conjugation.

Next, to discuss the mass terms in the theory, we substitute Eq. (2.3.4) into Eqs (2.2.44) and (2.2.45). Putting $T = 1_{m+n}$ for the moment then one can split the result for the operators O_i into a topologically trivial part and an instanton part as follows

$$O_i[Q_{\text{inst}}] = O_i[\Lambda_{\text{inst}}] = O_i[\Lambda] + O_i^{\text{inst}}(\mathbf{r}), \tag{2.3.6}$$

where

$$O_h[\Lambda] = \frac{4mn}{m+n}, \quad O_a[\Lambda] = -4mn, \quad O_s[\Lambda] = -4mn \tag{2.3.7}$$

and

$$\begin{aligned} O_h^{\text{inst}}(\mathbf{r}) = 2\rho_{11}^{11}(\mathbf{r}), \qquad O_a^{\text{inst}}(\mathbf{r}) = -4(m+n-1)\rho_{11}^{11}(\mathbf{r}) \\ O_s^{\text{inst}}(\mathbf{r}) = -4(m+n+1)\rho_{11}^{11}(\mathbf{r})\left(1+\frac{2\rho_{11}^{11}(\mathbf{r})}{m+n+2}\right). \end{aligned} \tag{2.3.8}$$

Similarly we can write the free energy as the sum of two parts

$$\mathcal{F}^{\text{class}} = \mathcal{F}_0^{\text{class}} + \mathcal{F}_{\text{inst}}^{\text{class}}, \tag{2.3.9}$$

where $\mathcal{F}_0^{\text{class}}$ denotes the contribution of the trivial vacuum with topological charge equal to zero ($Q = \Lambda$) and $\mathcal{F}_{\text{inst}}^{\text{class}}$ is the instanton part

$$\mathcal{F}_0^{\text{class}} = z_h \int d\mathbf{r} O_h[\Lambda] + z_s \int d\mathbf{r} O_s[\Lambda] + z_a \int d\mathbf{r} O_a[\Lambda] \tag{2.3.10}$$

$$\begin{aligned} \mathcal{F}_{\text{inst}}^{\text{class}} = \int\limits_{\text{inst}} \exp\Big[&- 2\pi\sigma_{xx} \pm i\theta + z_h \int d\mathbf{r} O_h^{\text{inst}}(\mathbf{r}) \\ &+ z_s \int d\mathbf{r} O_s^{\text{inst}}(\mathbf{r}) + z_a \int d\mathbf{r} O_a^{\text{inst}}(\mathbf{r})\Big]. \end{aligned} \tag{2.3.11}$$

The subscript "inst" on the integral sign indicates that the integral is in general to be performed over the manifold of instanton parameters. Notice, however, that the global matrix T is no longer a part of the instanton manifold except for the subgroup $U(m) \times U(n)$ only that leaves the action invariant. In the presence of the mass terms we therefore have, instead of Eq. (4.3.3),

$$Q_{\text{inst}}(\mathbf{r}) = W^{-1}\Lambda_{\text{inst}}(\mathbf{r})W = \Lambda + W^{-1}\rho(\mathbf{r})W \tag{2.3.12}$$

with $W \in U(m) \times U(n)$. On the other hand, the spatial integrals in the exponential of Eq. (2.3.11) still display a logarithmic divergence in the size of the system. To deal with these and other complications we shall in this thesis follow the methodology as outlined in Ref. [31]. Before embarking on the quantum theory, however, we shall first address the idea of working with *constrained instantons.* [45]

2.3.2 Constrained instantons

It is well known that the discrete topological sectors do not in general have stable classical minima for finite values of $z_i = 0$. It is nevertheless possible construct matrix field configurations that minimize the action under certain fixed constraints. In this Section we are interested in finite action field configurations Q that smoothly turn into the instanton solution of Eq. (2.3.12) in the limit where the symmetry breaking fields z_i all go to zero. Although the methodology of this thesis avoids the idea of *constrained instantons* all together, the problem nevertheless arises in the discussion of a special aspect of the theory, the replica limit (Section 2.7.5). To simplify the analysis we will consider the theory in the presence of the S_h term only.

Explicit solution

To obtain finite action configurations for finite values of z_h it is in many ways natural to start from the original solution of Eq. (2.3.12) and minimize the action with respect to a *spatially varying* scale size $\lambda(\mathbf{r})$ rather than a spatially independent parameter λ. Write

$$\lambda^2 \to \lambda^2(\mathbf{r}) = \lambda^2 f\left(x, \tilde{h}^2\right), \tag{2.3.13}$$

where $f(x, \tilde{h}^2)$ is a dimensionless function of the dimensionless quantities $x = r^2/\lambda^2$ and $\tilde{h}^2 = 4z_h\lambda^2/\sigma_{xx}$ respectively. The strategy is to find an optimal function f with the following constraints

$$f(x,0) = 1 \tag{2.3.14}$$
$$f(0,\tilde{h}^2) = 1 \tag{2.3.15}$$
$$f(x\to\infty,\tilde{h}^2>0) = 0. \tag{2.3.16}$$

The first of these equations ensures that in the absence of mass terms ($z_h = \tilde{h}^2 = 0$) we regain the original instanton solution. The second and third ensure that the classical action is finite for finite values of z_h or $\tilde{h}^2$.

It is easy to see that for all functions $f(x, \tilde{h}^2)$ satisfying Eqs (2.3.14) - (2.3.16) the topological charge equals unity, i.e.

$$\mathcal{C}[Q] = \int_0^\infty dx \frac{f - x\partial_x f}{(x+f)^2} = 1. \tag{2.3.17}$$

Next, the expression for the action becomes (discarding the constant terms in the definition of S_h)

$$S_\sigma + S_h = -\pi\sigma_{xx}\int_0^\infty dx \left\{ \frac{f}{(x+f)^2}\left[1 + \left(1 - \frac{x\partial_x f}{f}\right)^2\right] + \tilde{h}^2 \frac{f}{x+f} \right\}. \tag{2.3.18}$$

The function $f_0(x, \tilde{h}^2)$ that optimizes the action satisfies the following equation

$$-x[2\partial_x^2 f_0 - f_0(\partial_x f_0)^2] + 2\frac{f_0\partial_x f_0}{x+f_0}[x\partial_x f_0 - 2f_0] + \tilde{h}^2 f_0^2 = 0. \tag{2.3.19}$$

We are generally interested in the limit $\tilde{h}^2 \ll 1$ only. Eq. (2.3.19) cannot be solved analytically. The asymptotic behavior of $f_0(x, \tilde{h}^2)$ in the limit of large and small values of x is obtained as follows. For $x \gg 1$, we find

$$f_0(x, \tilde{h}^2) \propto \tilde{h}^2 x K_1^2(\tilde{h}\sqrt{x}) = \begin{cases} \frac{\pi \tilde{h}}{2}\sqrt{x}\exp(-2\tilde{h}\sqrt{x}), & x \gg \tilde{h}^{-2}, \\ 1 + \tilde{h}^2 x \ln \frac{\tilde{h}\sqrt{x}}{2}, & \tilde{h}^{-2} \gg x \gg 1. \end{cases} \tag{2.3.20}$$

Here, $K_1(z)$ stands for the modified Bessel function. In the regime $x \ll 1$ we obtain

$$f_0(x, \tilde{h}^2) \approx 1 + 4\tilde{h}^2 x, \qquad x \ll 1. \tag{2.3.21}$$

Notice that these asymptotic results can also be written more simply as follows

$$f_0(x, \tilde{h}^2) \approx \begin{cases} \frac{\pi}{2} hr \exp(-2hr), & r \gg h^{-1} \\ 1 + (hr)^2 \ln \frac{hr}{2}, & h^{-1} \gg r \gg \lambda \\ 1 + 4(hr)^2, & r \ll \lambda. \end{cases} \tag{2.3.22}$$

In the intermediate regimes $r \approx \lambda$ and $r \approx h^{-1}$ we have obtained the solution numerically. In Fig. 2.5 we plot the results in terms of the matrix element $\rho_{11}^{11}(r)$, Eq. (2.3.5),

$$\rho_{11}^{11}(r) = -\frac{2\lambda}{r^2 + \lambda^2} \to -\frac{2f}{x + f}. \tag{2.3.23}$$

Comparing the result with $\tilde{h}^2 = 0.3$ with the original instanton result, $f = 1$, we see that the main difference is in the asymptotic behavior with large r where the ρ_{11}^{11} for the "constrained instanton" vanishes exponentially ($\propto \exp(-2hr)$), rather than algebraically ($\propto 1/r^2$).

Finite action

We conclude that in the presence of the mass term S_h the requirement of finite action forces the instanton scale size to become spatially dependent in general such that the contributions from large distances $r \gg h^{-1}$ are strongly suppressed in the spatial integrals. On the other hand, from Eqs (2.3.20) and (2.3.21) one can also see that for small scale sizes λ such that

$$\lambda \ll \lambda_h \propto h^{-1}, \tag{2.3.24}$$

the dominant contribution comes from the original, unconstrained instanton solution of Eq. (2.3.4).

In order to see the effect of the constrained instanton on the analytic form of the action we first consider a simpler example of a trial function $f_{\mathrm{tr}}(x)$ of the type

$$f_{\mathrm{tr}}(x) = \exp(-ax). \tag{2.3.25}$$

Here we use a as a variational parameter. The classical action becomes

$$S_\sigma + S_h = -2\pi\sigma_{xx}\Big[1 + \frac{a}{2} + \frac{\tilde{h}^2}{2}\Big(\ln\frac{1}{a} - \gamma\Big)\Big] + \mathcal{O}(\tilde{h}^4), \tag{2.3.26}$$

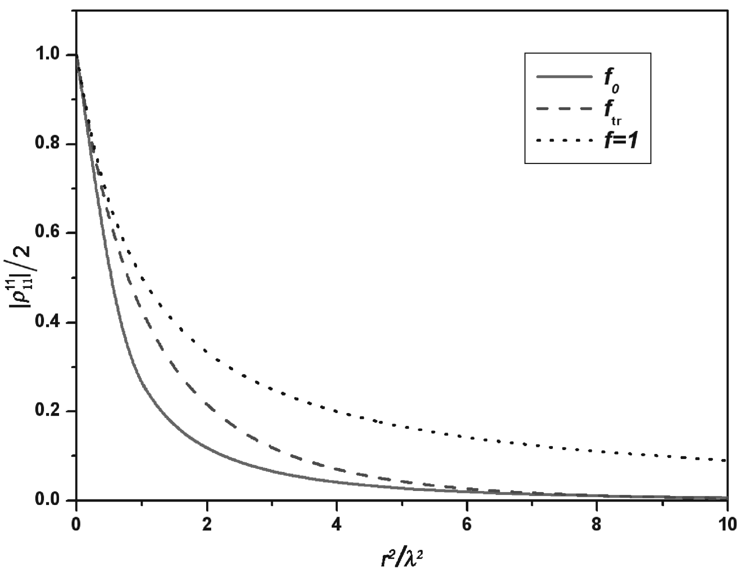


Figure 2.5: The matrix element $|\rho_{11}^{11}|/2$ as the function of r^2/λ^2 for $\tilde{h}^2 = 0.3$.

where the constant $\gamma \approx 0.577$ stands for the Euler constant. The optimal value of a is given by

$$a = \tilde{h}^2, \qquad f_{\rm tr}(x) = \exp(-\tilde{h}^2 x). \tag{2.3.27}$$

Hence we find for Eq. (2.3.26)

$$S_\sigma + S_h = -2\pi\sigma_{xx}\Big[1 + \frac{\tilde{h}^2}{2}\ln\frac{1.53}{\tilde{h}^2}\Big]. \tag{2.3.28}$$

This simple result obtained from the trial function (2.3.27) has the same asymptotic form as one determined by the much more complicated optimal function $f_0(x)$ in the limit $\tilde{h}^2 \to 0$, except that the numerical constant 1.53 is replaced by 0.85.

Eq. (2.3.28) indicates that that in the limit $h \to 0$ the finite action of the constrained instanton smoothly goes over into the finite action of the unconstrained instanton. Furthermore, it is easy to see that the final result of Eq. (2.3.28) has precisely the features that one would normally associate with mass terms. Consider for example the action of the unconstrained instanton with a fixed scale size λ. By putting the system inside a large circle of radius R we obtain the following result

$$S_\sigma + S_h = -2\pi\sigma_{xx}\Big[1 - \frac{\lambda^2}{R^2} + \frac{\tilde{h}^2}{2}\ln\frac{R^2}{\lambda^2}\Big]. \tag{2.3.29}$$

By comparing Eq. (2.3.28) and Eq. (2.3.29) we conclude that the following effective size R_h is induced by the h field

$$R_h = \frac{1.47}{h}. \tag{2.3.30}$$

The effect of mass terms can therefore be summarized as follows. First, the integral over scale sizes λ (see Eq. (2.3.11)) effectively proceeds over the interval $\lambda \lesssim h^{-1}$. Secondly, the spatial integrals in the theory are cut off in infrared and the effective sample size equals R_h. The two different scales that are being induced by S_h, h^{-1} and R_h, are of the same order of magnitude.

2.4 Quantum theory

One of the problems with the idea of constrained instantons is that it does not facilitate an expansion of the theory about $z_i = 0$ in any obvious fashion. This is unlike the method of *spatially varying masses* which is based on the results obtained in Ref. [20] and [21]. To start we first recapitulate the formalism of the theory without mass terms in Section 2.4.1. In Section 2.4.2 we introduce the idea of *spatially varying masses*. Finally, in Section 2.4.3 we present the complete action of the small oscillator problem that will be used in the remainder of this chapter.

2.4.1 Preliminaries

To obtain the most general matrix field variable Q with topological charge equal to unity we first rewrite the instanton solution Λ_{inst} in Eqs (2.3.4) and (2.3.5) as a unitary rotation R about the trivial vacuum Λ

$$\Lambda_{\text{inst}} = R^{-1} \Lambda R \tag{2.4.1}$$

where

$$R = \begin{pmatrix} \delta^{\alpha\beta} + (\bar{e}_1 - 1)\delta^{\alpha 1}\delta^{\beta 1} & e_0 \delta^{\alpha 1}\delta^{\beta 1} \\ -e_0 \delta^{\alpha 1}\delta^{\beta 1} & \delta^{\alpha\beta} + (e_1 - 1)\delta^{\alpha 1}\delta^{\beta 1} \end{pmatrix}. \tag{2.4.2}$$

The quantities e_0 and e_1 are defined as

$$e_0 = \frac{\lambda}{\sqrt{|z - z_0|^2 + \lambda^2}}, \qquad e_1 = \frac{z - z_0}{\sqrt{|z - z_0|^2 + \lambda^2}}. \tag{2.4.3}$$

For illustration we have written the full matrix $R^{\alpha\beta}_{pp'}$ in Fig. 2.6. It is a simple matter to next generalize these expressions and the result is

$$Q = T_0^{-1} R^{-1} q \, R \, T_0. \tag{2.4.4}$$

Here, T_0 denotes a global $U(m+n)$ rotation. The matrix q with $q^2 = 1_{m+n}$ represents the small fluctuations about the one instanton. Write

$$q = w + \Lambda \sqrt{1_{m+n} - w^2} \tag{2.4.5}$$

with

$$w = \begin{pmatrix} 0 & v \\ v^\dagger & 0 \end{pmatrix} \tag{2.4.6}$$

then the matrix q can formally be written as a series expansion in powers of the $m \times n$ complex matrices v, $v^\dagger$ which are taken as the independent field variables in the problem.

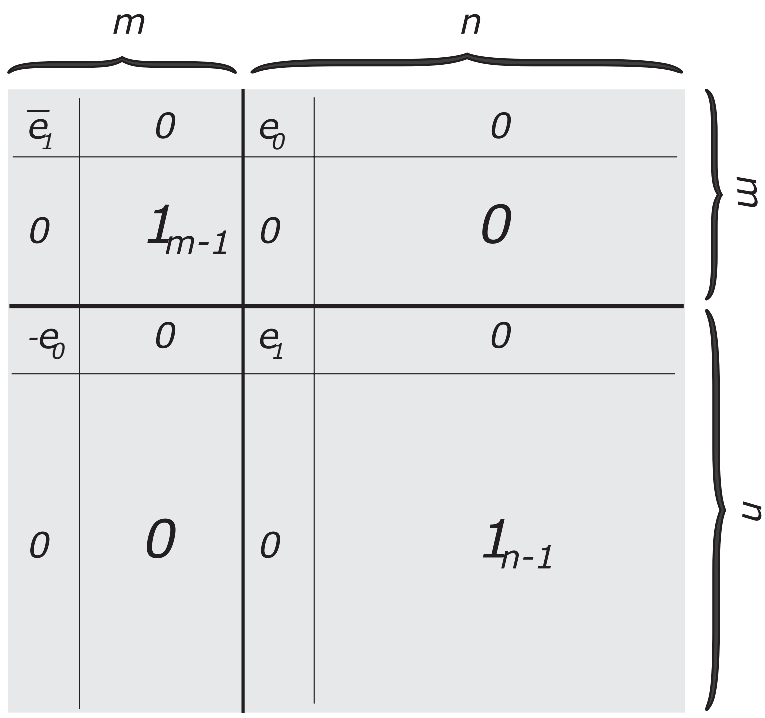


Figure 2.6: The matrix R.

Stereographic projection

Eq. (2.4.4) lends itself to an exact analysis of the small oscillator problem. First we recall the results obtained for the theory without mass terms, [20, 21]

$$\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 = \frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}[\nabla + \mathbf{A}, q]^2, \tag{2.4.7}$$

where the matrix $\mathbf{A}$ contains the instanton degrees of freedom

$$\mathbf{A} = RT_0\nabla T_0^{-1}R^{-1} = R\nabla R^{-1}. \tag{2.4.8}$$

By expanding the q in Eq. (2.4.7) to quadratic order in the quantum fluctuations v, $v^\dagger$ we obtain the following results

$$\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}[\nabla + \mathbf{A}, q]^2 = \frac{\sigma_{xx}}{4}\int d\mathbf{r}\mu^2(\mathbf{r})\Bigg[v^{11}O^{(2)}v^{\dagger 11} + \sum_{\alpha=2}^{m} v^{\alpha 1}O^{(1)}v^{\dagger 1\alpha} + \sum_{\beta=2}^{n} v^{1\beta}O^{(1)}v^{\dagger\beta 1} + \sum_{\alpha=2}^{m}\sum_{\beta=2}^{n} v^{\alpha\beta}O^{(0)}v^{\dagger\beta\alpha}\Bigg]. \tag{2.4.9}$$

The three different operators $O^{(a)}$ are given as

$$O^{(a)} = \frac{(r^2+\lambda^2)^2}{4\lambda^2}\left[\nabla_b + \frac{ia\varepsilon_{bc}r_c}{r^2+\lambda^2}\right]^2 + \frac{a}{2}. \tag{2.4.10}$$

The introduction of a measure $\mu^2(\mathbf{r})$ for the spatial integration in Eq. (2.4.9),

$$\mu(\mathbf{r}) = \frac{2\lambda}{r^2+\lambda^2}, \tag{2.4.11}$$

indicates that the quantum fluctuation problem is naturally defined on a sphere with radius λ. It is convenient to employ the stereographic projection

$$\eta = \frac{r^2-\lambda^2}{r^2+\lambda^2}, \qquad -1 < \eta < 1 \tag{2.4.12}$$

$$\theta = \tan^{-1}\frac{y}{x}, \qquad 0 \leqslant \theta < 2\pi. \tag{2.4.13}$$

In terms of η, θ the integration can be written as

$$\int d\mathbf{r}\mu^2(\mathbf{r}) = \int d\eta d\theta. \tag{2.4.14}$$

Moreover,

$$e_0 = \sqrt{\frac{1-\eta}{2}}, \qquad e_1 = \sqrt{\frac{1+\eta}{2}}e^{i\theta}, \tag{2.4.15}$$

and the operators become

$$O^{(a)} = \frac{\partial}{\partial\eta}\left[(1-\eta^2)\frac{\partial}{\partial\eta}\right] + \frac{1}{1-\eta^2}\frac{\partial^2}{\partial^2\theta} - \frac{ia}{1-\eta}\frac{\partial}{\partial\theta} - \frac{a^2}{4}\frac{1+\eta}{1-\eta} + \frac{a}{2}, \tag{2.4.16}$$

with $a = 0, 1, 2$. Finally, using Eq. (2.4.9) we can count the total number of fields $v^{\alpha\beta}$ on which each of the operators $O^{(a)}$ act. The results are listed in Table 2.2.

Table 2.2: Counting the number of zero modes

Operator	The number of fields $v^{\alpha\beta}$ involved	Degeneracy
$O^{(0)}$	$(m-1)(n-1)$	1
$O^{(1)}$	$(m-1)+(n-1)$	2
$O^{(2)}$	1	3

Energy spectrum

We are interested in the eigenvalue problem

$$O^{(a)}\Phi^{(a)}(\eta,\theta) = E^{(a)}\Phi^{(a)}(\eta,\theta), \tag{2.4.17}$$

where the set of eigenfunctions $\Phi^{(a)}$ are taken to be orthonormal with respect to the scalar product

$$(\bar{\Phi}_1^{(a)}, \Phi_2^{(a)}) = \int d\eta d\theta\, \bar{\Phi}_1^{(a)}(\eta,\theta)\Phi_2^{(a)}(\eta,\theta). \tag{2.4.18}$$

The Hilbert space of square integrable eigenfunctions is given in terms of Jacobi polynomials,

$$P_n^{\alpha,\beta}(\eta) = \frac{(-1)^n}{2^n n!}(1-\eta)^{-\alpha}(1+\eta)^{-\beta}\frac{d^n}{d\eta^n}(1-\eta)^{n+\alpha}(1+\eta)^{n+\beta} \tag{2.4.19}$$

Introducing the quantum number J to denote the discrete energy levels

$$\begin{array}{rclrcl} E_J^{(0)} &=& J(J+1), & J &=& 0,1,\cdots \\ E_J^{(1)} &=& (J-1)(J+1), & J &=& 1,2,\cdots \\ E_J^{(2)} &=& (J-1)(J+2), & J &=& 1,2,\cdots \end{array} \tag{2.4.20}$$

then the eigenfunctions are labelled by (J,M) and can be written as follows

$$\begin{array}{rclrcl} \Phi_{J,M}^{(0)} &=& C_{J,M}^{(0)} e^{iM\theta}\sqrt{(1-\eta^2)^M}P_{J-M}^{M,M}(\eta), & M &=& -J,\cdots,J \\ \Phi_{J,M}^{(1)} &=& C_{J,M}^{(1)} e^{iM\theta}\sqrt{(1-\eta^2)^M}\sqrt{1-\eta}P_{J-M-1}^{M+1,M}(\eta), & M &=& -J,\cdots,J-1 \\ \Phi_{J,M}^{(2)} &=& C_{J,M}^{(2)} e^{iM\theta}\sqrt{(1-\eta^2)^M}(1-\eta)P_{J-M-1}^{M+2,M}(\eta), & M &=& -J-1,\cdots,J-1 \end{array} \tag{2.4.21}$$

where the normalization constants equal

$$\begin{aligned} C^{(0)}_{J,M} &= \frac{\sqrt{\Gamma(J-M+1)\Gamma(J+M+1)(2J+1)}}{2^{M+1}\sqrt{\pi}\Gamma(J+1)}, \\ C^{(1)}_{J,M} &= \frac{\sqrt{\Gamma(J-M)\Gamma(J+M+1)}}{2^{M+1}\sqrt{\pi}\Gamma(J)}, \\ C^{(2)}_{J,M} &= \frac{\sqrt{\Gamma(J-M)\Gamma(J+M+2)(2J+1)}}{2^{M+2}\sqrt{\pi}\Gamma(J)\sqrt{J(J+1)}}. \end{aligned} \tag{2.4.22}$$

Zero modes

From Eq. (2.4.20) we see that the operators $O^{(a)}$ have the following zero modes

$$\begin{aligned} O^{(0)} &\Longrightarrow \Phi^{(0)}_{0,0} = 1, \\ O^{(1)} &\Longrightarrow \Phi^{(1)}_{1,-1} = \frac{1}{\sqrt{2\pi}}\bar{e}_1, \quad \Phi^{(1)}_{1,0} = \frac{1}{\sqrt{2\pi}}e_0, \\ O^{(2)} &\Longrightarrow \Phi^{(2)}_{1,-2} = \sqrt{\frac{3}{4\pi}}\bar{e}_1^2, \quad \Phi^{(2)}_{1,-1} = \sqrt{\frac{3}{2\pi}}e_0\bar{e}_1, \quad \Phi^{(2)}_{1,0} = \sqrt{\frac{3}{4\pi}}e_0^2. \end{aligned} \tag{2.4.23}$$

The number of the zero modes of each $O^{(a)}$ is listed in Table 2.2. The total we find $2(mn+m+n)$ zero modes in the problem. Next, it is straight forward to express these zero modes in terms of the instanton degrees of freedom contained in the matrices R and T_0 of Eq. (2.4.4). For this purpose we write the instanton solution as follows

$$Q_{\rm inst}(\xi_i) = U^{-1}(\xi_i)\Lambda U(\xi_i). \tag{2.4.24}$$

Here, $U = RT_0$ and the ξ_i stand for the position z_0 of the instanton, the scale size λ and the generators of T_0. The effect of an infinitesimal change in the instanton parameters $\xi_i \to \xi_i + \varepsilon_i$ on the $Q_{\rm inst}$ can be written in the form of Eq. (2.4.4) as follows

$$Q_{\rm inst}(\xi_i+\varepsilon_i) = U^{-1}(\xi_i)q_\varepsilon U(\xi_i), \tag{2.4.25}$$

where

$$q_\varepsilon \approx \Lambda - \varepsilon_i\left[U\partial_{\varepsilon_i}U^{-1},\Lambda\right]. \tag{2.4.26}$$

Notice that

$$-\varepsilon_i\left[U\partial_{\varepsilon_i}U^{-1},\Lambda\right] = 2\varepsilon_i\begin{pmatrix} 0 & \left[U\partial_{\varepsilon_i}U^{-1}\right]^{\alpha\beta}_{1,-1} \\ -\left[U\partial_{\varepsilon_i}U^{-1}\right]^{\alpha\beta}_{-1,1} & 0\end{pmatrix}. \tag{2.4.27}$$

By comparing this expression with Eq. (2.4.4) we see that the small changes ε_i tangential to the instanton manifold can be cast in the form of the quantum fluctuations v, $v^\dagger$ according to

$$v^{\alpha\beta} = 2\varepsilon_i\left[RT_0\partial_{\varepsilon_i}T_0^{-1}R^{-1}\right]^{\alpha\beta}_{1,-1}, \qquad v^{\dagger\alpha\beta} = -2\varepsilon_i\left[RT_0\partial_{\varepsilon_i}T_0^{-1}R^{-1}\right]^{\alpha\beta}_{-1,1}. \tag{2.4.28}$$

Next we wish to work out these expressions explicitly. Let t denote an infinitesimal $U(m+n)$ rotation

$$T_0 = 1_{m+n} + i\,t, \tag{2.4.29}$$

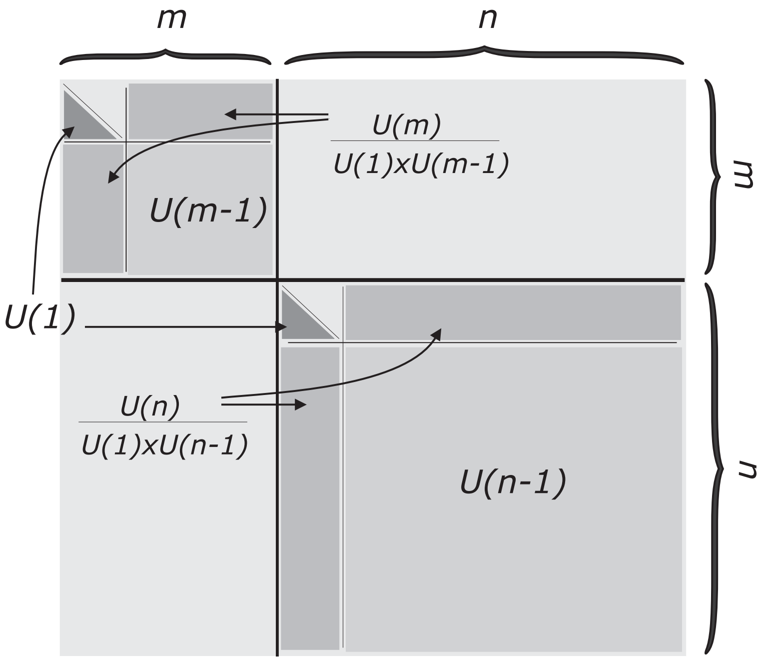


Figure 2.7: Symmetry breaking.

and $\delta\lambda$, δz_0 infinitesimal changes in the scale size and position respectively,

$$R(\lambda+\delta\lambda, z_0+\delta z_0) = R(\lambda, z_0)\left[1_{m+n} + \delta\lambda\, R^{-1}\partial_\lambda R + \delta z_0\, R^{-1}\partial_{z_0} R\right]. \qquad (2.4.30)$$

The zero frequency modes can be expressed in terms of the instanton parameters t, $\delta\lambda$ and δz_0 and the eigenfunctions $\Phi^{(a)}_{JM}$ according to

$$\begin{aligned} v^{\alpha\beta} &= 2it^{\alpha\beta}_{1,-1}\Phi^{(0)}_{0,0} \\ \begin{bmatrix} v^{\alpha 1} \\ v^{1\beta} \end{bmatrix} &= 2\sqrt{2\pi}i\begin{pmatrix} t^{\alpha 1}_{1,-1} & -t^{\alpha 1}_{1,1} \\ t^{1\beta}_{1,-1} & t^{1\beta}_{-1,-1} \end{pmatrix}\begin{bmatrix} \Phi^{(1)}_{1,-1} \\ \Phi^{(1)}_{1,0} \end{bmatrix} \\ v^{11} &= 4\sqrt{\frac{\pi}{3}}\left[it^{11}_{-1,1}, \left(it^{11}_{-1,-1} - it^{11}_{1,1} - \frac{\delta\lambda}{\lambda}\right), -it^{11}_{1,-1} + \frac{\delta\bar{z}_0}{\lambda}\right]\begin{bmatrix} \Phi^{(2)}_{1,-2} \\ \Phi^{(2)}_{1,-1} \\ \Phi^{(2)}_{1,0} \end{bmatrix} \end{aligned} \qquad (2.4.31)$$

From this one can see that besides the scale size λ and position z_0 the instanton manifold is spanned by the $t^{\alpha\beta}_{1,-1}$ and $t^{\alpha\beta}_{-1,1}$ which are the generators of $U(m+n)/U(m)\times U(n)$. The $t^{1\alpha}_{1,1}$ and $t^{\alpha 1}_{1,1}$ with $\alpha > 1$ are the generators of $U(n)/U(n-1)\times U(1)$ and the $t^{1\alpha}_{-1,-1}$ and $t^{\alpha 1}_{-1,-1}$ those of $U(m)/U(m-1)\times U(1)$. Finally, $t^{11}_{1,1} - t^{11}_{-1,-1}$ is the $U(1)$ generator describing the rotation of the $O(3)$ instanton about the z axis. In total we find $2(mn+m+n)$ zero modes as it should be. The hierarchy of symmetry breaking by the one-instanton solution is illustrated in Fig. 2.7.

2.4.2 Spatially varying masses

We have seen that the quantum fluctuations about the instanton acquire the metric of a *sphere*, Eq. (2.4.11). This, however, complicates the problem of mass terms which are naturally written in *flat* space. To deal with this problem we modify the definition of the mass terms and introduce a spatially varying momentum scale $\mu(\mathbf{r})$ as follows

$$z_i \to z_i\, \mu^2(\mathbf{r}), \tag{2.4.32}$$

such that the action now becomes *finite* and can be written as

$$S_i = z_i \int d\mathbf{r}\, O_i[Q] \to z_i \int d\mathbf{r}\, \mu^2(\mathbf{r}) O_i[Q]. \tag{2.4.33}$$

Several comments are in order. First of all, we expect that the introduction of a spatially varying momentum scale $\mu(\mathbf{r})$ does not alter the singularity structure of the theory at short distances. We shall therefore proceed and first develop a full quantum theory for the modified mass terms in Sections 2.5 and 2.6. Secondly, we postpone the problem of *curved* versus *flat* space all the way until the end of the computation in Section 2.7 where we elaborate on the tricks developed by 't Hooft. [44]

As we shall discuss in detail in the remainder of this chapter, the validity of the procedure with *spatially varying masses* relies entirely on the statement which says that the quantum theory of the modified instanton problem displays exactly the same ultraviolet singularities as those obtained in ordinary perturbative expansions. In fact, we shall greatly benefit from our introduction of *observable parameters* since it can be used to explicitly verify this statement.

Since the action is now finite one can go ahead and formally expand the theory about $z_i = 0$. To see how this works let us first consider the operator O_h. Using Eq. (2.4.4) we can write for S_h

$$S_h = z_h \int d\mathbf{r}\, \mu^2(\mathbf{r})\, \mathrm{tr}\, A_h q, \tag{2.4.34}$$

where

$$A_h = R\, T_0 \left(\Lambda - \frac{m-n}{m+n} 1_{m+n} \right) T_0^{-1} R^{-1}. \tag{2.4.35}$$

Now we can formally proceed by evaluating the expectation $\langle q \rangle$ with respect to the theory of the previous Section, Eq. (2.4.9). It is important to keep in mind, however, that the global unitary matrix T_0 is eventually restricted to run over the subgroup $U(m) \times U(n)$ only. Since it is in many ways simpler to carry out the quantum fluctuations about the theory with $T_0 = 1_{m+n}$, we shall in what follows specialize to this simpler case. We will come back to the more general situation in Section 2.6 where we show that the final results are in fact independent of the specific choice made for the matrix T_0.

2.4.3 Action for the quantum fluctuations

Keeping the remarks of the previous Section in mind we obtain the complete action as the sum of a classical part S^{inst} and a quantum part δS as follows

$$S = S^{\text{inst}} + \delta S, \tag{2.4.36}$$

where

$$S^{\rm inst} = -2\pi\sigma_{xx} + i\,\theta + S_h^{\rm inst} + S_a^{\rm inst} + S_s^{\rm inst} \qquad (2.4.37)$$

and

$$\delta S = \delta S_\sigma + \delta S_h + \delta S_a + \delta S_s. \qquad (2.4.38)$$

Here, $S_i^{\rm inst}$ stands for the classical action of the modified mass terms $O_i^{\rm inst}$, Eq. (2.3.6), and is given by

$$\begin{aligned} S_h^{\rm inst} &= z_h \int d\mathbf{r}\,\mu^2(\mathbf{r}) O_h^{\rm inst}(\mathbf{r}) = -8\pi z_h, \\ S_a^{\rm inst} &= z_a \int d\mathbf{r}\,\mu^2(\mathbf{r}) O_a^{\rm inst}(\mathbf{r}) = 16\pi(m+n-1) z_a, \\ S_s^{\rm inst} &= z_s \int d\mathbf{r}\,\mu^2(\mathbf{r}) O_s^{\rm inst}(\mathbf{r}) = 16\pi(m+n+1) z_s \left[1 - \frac{8}{3(m+n+2)}\right]. \end{aligned} \qquad (2.4.39)$$

Next, the results for δS_σ and δS_i can be written up to quadratic order in v, $v^\dagger$ as follows

$$\begin{aligned} \delta S_\sigma = -\frac{\sigma_{xx}}{4} \int_{\eta\theta} \Bigg[& \sum_{\alpha=2}^{m}\sum_{\beta=2}^{n} v^{\alpha\beta} O^{(0)} v^{\dagger\beta\alpha} + \sum_{\alpha=2}^{m} v^{\alpha 1} O^{(1)} v^{\dagger 1\alpha} + \sum_{\beta=2}^{n} v^{1\beta} O^{(1)} v^{\dagger\beta 1} \\ & + v^{11} O^{(2)} v^{\dagger 11} \Bigg] \end{aligned} \qquad (2.4.40)$$

where

$$\int_{\eta\theta} \equiv \int d\eta d\theta. \qquad (2.4.41)$$

The operators $O^{(a)}$ with $a = 0, 1, 2$ are given by Eq. (2.4.16). Furthermore,

$$\begin{aligned} \delta S_h = -z_h \int_{\eta\theta} \Bigg[& \sum_{\alpha=2}^{m}\sum_{\beta=2}^{n} v^{\alpha\beta} v^{\dagger\beta\alpha} + (1-e_0^2)\left(\sum_{\alpha=2}^{m} v^{\alpha 1} v^{\dagger 1\alpha} + \sum_{\beta=2}^{n} v^{1\beta} v^{\dagger\beta 1}\right) \\ & + (1-2e_0^2) v^{11} v^{\dagger 11} + 2(e_0 e_1 v^{11} + e_0 \bar{e}_1 v^{\dagger 11}) \Bigg], \end{aligned} \qquad (2.4.42)$$

$$\begin{aligned} \delta S_a = 2(m+n-1) z_a \int_{\eta\theta} \Bigg[& \frac{m+n-2-4e_0^2}{m+n-2} \sum_{\alpha=2}^{m}\sum_{\beta=2}^{n} v^{\alpha\beta} v^{\dagger\beta\alpha} + (1-e_0^2) \\ & \times \Big(\sum_{\alpha=2}^{m} v^{\alpha 1} v^{\dagger 1\alpha} + \sum_{\beta=2}^{n} v^{1\beta} v^{\dagger\beta 1}\Big) + (1-2e_0^2) v^{11} v^{\dagger 11} \\ & + 2(e_0 e_1 v^{11} + e_0 \bar{e}_1 v^{\dagger 11}) \Bigg], \end{aligned} \qquad (2.4.43)$$

$$\delta S_s = 2\frac{m+n+1}{m+n+2}z_s\int_{\eta\theta}\Bigg[(m+n+2-4e_0^2)\sum_{\alpha=2}^{m}\sum_{\beta=2}^{n}v^{\alpha\beta}v^{\dagger\beta\alpha}+(1-e_0^2)$$
$$\times(m+n+2-8e_0^2)\left(\sum_{\alpha=2}^{m}v^{\alpha 1}v^{\dagger 1\alpha}+\sum_{\beta=2}^{n}v^{1\beta}v^{\dagger\beta 1}\right)$$
$$+\Big[(1-2e_0^2)(m+n+2-8e_0^2)-8e_0^2|e_1|^2\Big]v^{11}v^{\dagger 11}$$
$$-4(e_0^2e_1^2v^{11}v^{11}+e_0^2\bar{e}_1^2v^{\dagger 11}v^{\dagger 11})$$
$$+2(m+n+2-8e_0^2)(e_0e_1v^{11}+e_0\bar{e}_1v^{\dagger 11})\Bigg]. \quad (2.4.44)$$

Notice that the terms linear in v, $v^\dagger$ in Eqs (2.4.42)-(2.4.44) can be written as

$$\int_{\eta\theta}\left[e_0e_1v^{11}+e_0\bar{e}_1v^{\dagger 11}\right]\propto\int_{\eta\theta}\left[\Phi^{(2)}_{1,-1}v^{11}+\bar{\Phi}^{(2)}_{1,-1}v^{\dagger 11}\right]\propto\frac{\delta\lambda}{\lambda}. \quad (2.4.45)$$

This means that the fluctuations tangential to the instanton parameter λ are the only unstable fluctuations in the problem. However, the linear fluctuations are not of any special interest to us and we proceed by formally evaluating the quantum fluctuations to first order in the fields z_i only. The expansion is therefore with respect to the theory with δS_σ alone and this has been analyzed in detail in Ref.[20, 21].

Finally, we also need the action S^0 for the quantum fluctuations about the trivial vacuum. The result is given by

$$S^{(0)}=\delta S^{(0)}_\sigma+\delta S^{(0)}_h+\delta S^{(0)}_a+\delta S^{(0)}_s \quad (2.4.46)$$

where

$$\delta S^{(0)}_\sigma=-\frac{\sigma_{xx}}{4}\int_{\eta\theta}\sum_{\alpha=1}^{m}\sum_{\beta=1}^{n}v^{\alpha\beta}O^{(0)}v^{\dagger\beta\alpha},\qquad \delta S^{(0)}_h=-z_h\int_{\eta\theta}\sum_{\alpha=1}^{m}\sum_{\beta=1}^{n}v^{\alpha\beta}v^{\dagger\beta\alpha},$$
$$\delta S^{(0)}_{a,s}=2(m+n\mp 1)z_{a,s}\int_{\eta\theta}\sum_{\alpha=1}^{m}\sum_{\beta=1}^{n}v^{\alpha\beta}v^{\dagger\beta\alpha}. \quad (2.4.47)$$

2.5 Pauli-Villars regulators

Recall that after integration over the quantum fluctuations one is left with two sources of divergences. First, there are the ultraviolet divergences which eventually lead to the renormalization of the coupling constant or σ_{xx}. At present we wish to extend the analysis to include the renormalization of the the z_i fields. The ultraviolet can be dealt with in a standard manner, employing Pauli-Villars regulator fields with masses $\mathcal{M}_f$ ($f=0,1,\cdots,K$) and with an alternating metric $\hat{e}_f$. [94] We assume $\hat{e}_0=1$,

$\mathcal{M}_0 = 0$ and large masses $\mathcal{M}_f \gg 1$ for $f > 1$. The following constraints are imposed

$$\sum_{f=0}^{K} \hat{e}_f \mathcal{M}_f^k = 0, \qquad 0 \leqslant k < K, \qquad \sum_{f=1}^{K} \hat{e}_f \ln \mathcal{M}_f = -\ln \mathcal{M}.$$

The regularized theory is then defined as

$$\delta S_{\text{reg}} = \delta S_0 + \sum_{f=1}^{K} \hat{e}_f \delta S_f. \tag{2.5.1}$$

Here action δS_f is the same as the action δS except that the kinetic operators $O^{(a)}$ are all replaced by $O^{(a)} + \mathcal{M}_f^2$.

Our task is to evaluate Eq. (2.5.1) to first order in the fields z_i. This means that δS_0 still naively diverges due to the zero modes of the operators $O^{(a)}$. These zero modes are handled separately in Section 2.6, by employing the collective coordinate formalism introduced in Ref. [20]. The regularized theory δS_{reg} is therefore defined by omitting the contributions of all the zero modes in δS_0.

To simplify the notation we shall work with $m = n$ in the subsequent Sections. The final answer will be expressed in terms of m and n, however.

2.5.1 Explicit computations

To simplify the notation we will first collect the results obtained after a naive integration over the field variables $v, v^\dagger$. These are easily extended to include the alternating metric and the Pauli-Villars masses which will be main topic of the next Section. Consider the ratio

$$\begin{aligned} \frac{Z_{\text{inst}}}{Z_0} = \frac{\int \mathcal{D}[v, v^\dagger] \exp S}{\int \mathcal{D}[v, v^\dagger] \exp S_0} = \exp\Big[& - 2\pi\sigma_{xx} \pm i\theta + D \\ & + S_h^{\text{inst}} + \Delta S_h \\ & + S_a^{\text{inst}} + \Delta S_a \\ & + S_s^{\text{inst}} + \Delta S_s \Big]. \end{aligned} \tag{2.5.2}$$

Here, the quantum corrections D, ΔS_h, ΔS_a and ΔS_s can be expressed in terms of the propagators

$$\mathcal{G}_a = \frac{1}{O^{(a)}} = \sum_{JM} \frac{|JM\rangle_{(a)(a)}\langle JM|}{E_J^{(a)}}, \qquad a = 0, 1, 2. \tag{2.5.3}$$

The results can be written as follows

$$
\begin{aligned}
D &= \mathrm{tr}\Big[2(n-1)\Big(\ln\mathcal{G}_1 - \ln\mathcal{G}_0\Big) - \Big(\ln\mathcal{G}_2 - \ln\mathcal{G}_0\Big)\Big], &(2.5.4)\\
\Delta S_h &= -\frac{4z_h}{\sigma_{xx}}\,\mathrm{tr}\Big[2(n-1)[(1-e_0^2)\mathcal{G}_1 - \mathcal{G}_0] + [(1-2e_0^2)\mathcal{G}_2 - \mathcal{G}_0]\Big], &(2.5.5)\\
\Delta S_a &= \frac{8(2n-1)z_a}{\sigma_{xx}}\,\mathrm{tr}\Big[(n-1)[(n-1-2e_0^2)\mathcal{G}_0 - (n-1)\mathcal{G}_0] + 2(n-1)\\
&\qquad \times[(1-e_0^2)\mathcal{G}_1 - \mathcal{G}_0] + [(1-2e_0^2)\mathcal{G}_2 - \mathcal{G}_0]\Big], &(2.5.6)\\
\Delta S_s &= \frac{8(2n+1)z_s}{(n+1)\sigma_{xx}}\,\mathrm{tr}\Big[(n-1)^2[(n+1-2e_0^2)\mathcal{G}_0 - (n+1)\mathcal{G}_0]\\
&\qquad +2(n-1)[(1-e_0^2)(n+1-4e_0^2)\mathcal{G}_1 - (n+1)\mathcal{G}_0]\\
&\qquad +\Big([(1-2e_0^2)(n+1-4e_0^2) - 4e_0^2(1-e_0^2)]\mathcal{G}_2\\
&\qquad -(n+1)\mathcal{G}_0\Big)\Big]. &(2.5.7)
\end{aligned}
$$

In these expressions the trace is taken with respect to the complete set of eigenfunctions of the operators $O^{(a)}$. To evaluate these expressions we need the help of the following identities (see Appendix 2.A)

$$\sum_{M=-J-a+k_a}^{J-k_a} {}_{(a)}\langle J,M|e_0^2|M,J\rangle_{(a)} = \frac{1}{2}(2J+1+a-2k_a), \qquad (2.5.8)$$

$$\sum_{M=-J-a+k_a}^{J-k_a} {}_{(a)}\langle J,M|e_0^4|M,J\rangle_{(a)} = \frac{1}{3}(2J+1+a-2k_a), \qquad (2.5.9)$$

where $k_a = 0,1,1$ for $a = 0,1,2$ respectively. After elementary algebra we obtain

$$
\begin{aligned}
D &= -\quad 2(n-1)D^{(1)} - D^{(2)}, &(2.5.10)\\
\Delta S_h &= -\quad z_h\frac{4}{\sigma_{xx}}\Big[(n-1)Y^{(1)} - (2n-1)Y^{(0)}\Big], &(2.5.11)\\
\Delta S_a &= \quad z_a\frac{8(2n-1)}{\sigma_{xx}}\Big[(n-1)Y^{(1)} - (3n-2)Y^{(0)}\Big], &(2.5.12)\\
\Delta S_s &= \quad z_s\frac{8(2n+1)(3n-1)}{3(n+1)\sigma_{xx}}\Big[(n-1)Y^{(1)} - 3nY^{(0)}\Big]. &(2.5.13)
\end{aligned}
$$

We have introduced the following quantities

$$
\begin{aligned}
D^{(r)} &= \sum_{J=1}^{\infty}(2J+r-1)\ln E_J^{(r)} - \sum_{J=0}^{\infty}(2J+1)\ln E_J^{(0)}, \quad r=1,2 &(2.5.14)\\
Y^{(s)} &= \sum_{J=s}^{\infty}\frac{2J+1-s}{E_J^{(s)}}, \qquad s=0,1. &(2.5.15)
\end{aligned}
$$

2.5.2 Regularized expressions

To obtain the regularized theory one has to include the alternating metric e_f and add the masses $\mathcal{M}_f$ to the energies $E_J^{(a)}$ in the expressions for $D^{(r)}$ and $Y^{(s)}$. To start, let us define the function

$$\Phi^{(\Lambda)}(p) = \sum_{J=p}^{\Lambda} 2J \ln(J^2 - p^2). \tag{2.5.16}$$

According to Eq. (2.5.1), the regularized function $\Phi_{\text{reg}}^{(\Lambda)}(p)$ is given by

$$\Phi_{\text{reg}}^{(\Lambda)}(p) = \sum_{J=p+1}^{\Lambda} 2J \ln(J^2 - p^2) + \sum_{f=1}^{K} e_f \sum_{J=p}^{\Lambda} 2J \ln(J^2 - p^2 + \mathcal{M}_f^2), \tag{2.5.17}$$

where we assume that the cut-off Λ is much larger than $\mathcal{M}_f$. In the presence of a large mass $\mathcal{M}_f$ we may consider the logarithm to be a slowly varying function of the discrete variable J. We may therefore approximate the summation by using the Euler-Maclaurin formula

$$\sum_{J=p+1}^{\Lambda} g(J) = \int_p^{\Lambda} dx\, g(x) + \frac{1}{2} g(x)\Big|_p^{\Lambda} + \frac{1}{12} g'(x)\Big|_p^{\Lambda}. \tag{2.5.18}$$

After some algebra we find that Eq. (2.5.17) can be written as follows [20]

$$\begin{aligned} \Phi_{\text{reg}}^{(\Lambda)}(p) = & - 2\Lambda(\Lambda+1)\ln\Lambda + \Lambda^2 - \frac{\ln e\Lambda}{3} + 4\sum_{J=1}^{\Lambda} J \ln J \\ & + \frac{1-6p}{3} \ln \mathcal{M} + 2p^2 - 2\sum_{J=1}^{2p} (J-p)\ln J. \end{aligned} \tag{2.5.19}$$

The regularized expression for $D^{(r)}$ can be obtained as

$$D_{\text{reg}}^{(r)} = \lim_{\Lambda\to\infty} \left[\Phi_{\text{reg}}^{(\Lambda)}\left(\frac{1+r}{2}\right) - \Phi_{\text{reg}}^{(\Lambda)}\left(\frac{1}{2}\right) \right]. \tag{2.5.20}$$

From this we obtain the final results

$$D_{\text{reg}}^{(1)} = -\ln\mathcal{M} + \frac{3}{2} - 2\ln 2, \qquad D_{\text{reg}}^{(2)} = -2\ln\mathcal{M} + 4 - 3\ln 3 - \ln 2. \tag{2.5.21}$$

Next, we introduce another function

$$Y^{(\Lambda)}(p) = \sum_{J=p}^{\Lambda} \frac{2J}{J^2 - p^2}. \tag{2.5.22}$$

According to Eq. (2.5.1), the regularized function $Y_{\text{reg}}^{(\Lambda)}(p)$ is given by

$$Y_{\text{reg}}^{(\Lambda)}(p) = \sum_{f=1}^{K} e_f \sum_{J=p}^{\Lambda} \frac{2J}{J^2 - p^2 + \mathcal{M}_f^2} + \sum_{J=p+1}^{\Lambda} \frac{2J}{J^2 - p^2}, \tag{2.5.23}$$

where as before we assume that the cut-off $\Lambda \gg \mathcal{M}_f$. By using a similar procedure as discussed above we now find

$$Y^{(\Lambda)}_{\rm reg}(p) = 2\ln\mathcal{M} + 2\gamma - \sum_{J=1}^{2p} \frac{1}{J} + \mathcal{O}(\Lambda^{-1}). \tag{2.5.24}$$

The regularized expressions for $Y^{(s)}$ can be written as

$$Y^{(s)}_{\rm reg} = \lim_{\Lambda\to\infty} Y^{(\Lambda)}_{\rm reg}\left(\frac{1+s}{2}\right), \tag{2.5.25}$$

such that we finally obtain

$$\begin{aligned} Y^{(0)}_{\rm reg} &= 2\ln\mathcal{M} + 2\gamma - 1, \\ Y^{(1)}_{\rm reg} &= 2\ln\mathcal{M} + 2\gamma - \frac{3}{2}, \\ Y^{(2)}_{\rm reg} &= 2\ln\mathcal{M} + 2\gamma - \frac{11}{6}. \end{aligned} \tag{2.5.26}$$

We therefore have the following results for the quantum corrections

$$D^{\rm reg} = 2n\ln\mathcal{M} + n(4\ln 2 - 3) - 1 + 3\ln\frac{3}{2}, \tag{2.5.27}$$

$$\Delta S^{\rm reg}_h = z_h \frac{8n}{\sigma_{xx}}\left[\ln\mathcal{M}e^{\gamma-1/2} - \frac{2n-1}{4n}\right], \tag{2.5.28}$$

$$\Delta S^{\rm reg}_a = -z_a \frac{16(2n-1)^2}{\sigma_{xx}}\left[\ln\mathcal{M}e^{\gamma-1/2} - \frac{3n-2}{2(2n-1)}\right], \tag{2.5.29}$$

$$\Delta S^{\rm reg}_s = -z_s \frac{16(2n+1)^2(3n-1)}{3\sigma_{xx}(n+1)}\left[\ln\mathcal{M}e^{\gamma-1/2} - \frac{3n}{2(2n+1)}\right]. \tag{2.5.30}$$

Apart from the logarithmic singularity in $\mathcal{M}$, the numerical constants in the expression for $D^{\rm reg}$ are going to play an important role in what follows. This is unlike the expressions for $\Delta S^{\rm reg}_i$ where the second term in the brackets should actually be considered as higher order terms in an expansion in powers of $1/\sigma_{xx}$. We collect the various terms together and obtain the following result for the instanton contribution to the free energy

$$\begin{aligned} \ln\left[\frac{Z_{\rm inst}}{Z_0}\right]^{\rm reg} = &- 1 + 3\ln\frac{3}{2} - (m+n)\left(\gamma + \frac{3}{2} - 2\ln 2\right) \pm i\,\theta && (2.5.31) \\ &- 2\pi\sigma_{xx}\left[1 - \frac{m+n}{2\pi\sigma_{xx}}\ln\mathcal{M}e^{\gamma}\right] && (2.5.32) \\ &- 8\pi z_h\left[1 - \frac{m+n}{2\pi\sigma_{xx}}\ln\mathcal{M}e^{\gamma-1/2}\right] && (2.5.33) \\ &+ 16\pi(m+n-1)z_a\left[1 - \frac{m+n-1}{\pi\sigma_{xx}}\ln\mathcal{M}e^{\gamma-1/2}\right] && (2.5.34) \\ &+ \frac{16\pi(m+n+1)(3m+3n-2)}{3(m+n+2)}z_s\left[1 - \frac{m+n+1}{\pi\sigma_{xx}}\ln\mathcal{M}e^{\gamma-1/2}\right]. && (2.5.35) \end{aligned}$$

2.5.3 Observable theory in Pauli-Villars regularization

The important feature of these last expressions, as we shall see next, is that the quantum corrections to the parameters σ_{xx}, z_h, z_a and z_s are all identically the same as those obtained from the perturbative expansions of the observable parameters σ'_{xx}, z'_h, z'_a and z'_s introduced in Section 2.2.4. Notice that we have already evaluated this theory in *dimensional regularization* in Section 2.2.4. The problem, however, is that the different regularization schemes (*dimensional* versus *Pauli-Villars*) are not related to one another in any obvious fashion. Unlike dimensional regularization, for example, it is far from trivial to see how the general form of the observable parameters, Eqs (2.2.72) and (2.2.73), can be obtained from the theory in Pauli-Villars regularization.

In Appendix 2.B we give the details of the computation using Pauli-Villars regulators. Denoting the results for σ'_{xx} and z'_i by $\sigma_{xx}(\mathcal{M})$ and $z_i(\mathcal{M})$ respectively,

$$\sigma'_{xx} = \sigma_{xx}(\mathcal{M}), \quad z'_i = z_i(\mathcal{M}) \tag{2.5.36}$$

then we have

$$\sigma_{xx}(\mathcal{M}) = \sigma_{xx}\left[1 - \frac{m+n}{2\pi\sigma_{xx}}\ln \mathcal{M}e^{\gamma}\right], \tag{2.5.37}$$

$$z_i(\mathcal{M}) = z_i\left[1 + \frac{\gamma_i^{(0)}}{\sigma_{xx}}\ln \mathcal{M}e^{\gamma-1/2}\right]. \tag{2.5.38}$$

The coefficients $\gamma_i^{(0)}$ are given by

$$\gamma_h^{(0)} = -\frac{m+n}{2\pi}, \qquad \gamma_a^{(0)} = -\frac{m+n-1}{\pi}, \qquad \gamma_s^{(0)} = -\frac{m+n+1}{\pi}. \tag{2.5.39}$$

Since so much of what follows is based on the results obtained in this Section and the previous one, it is worthwhile to first present a summary of the various issues that are involved. First of all, it is important to emphasize that our results for observable parameter σ'_{xx}, Eq. (2.5.37), resolve an ambiguity that is well known to exist, in the instanton analysis of scale invariant theories. Given Eq. (2.5.37) one uniquely fixes the quantity $\sigma_{xx}(\mathcal{M})$ (Eq. (2.5.32)) and the constant term of order unity (Eq. (2.5.31)) that is otherwise left undetermined. This result becomes particularly significant when we address the non-perturbative aspects of the renormalization group β functions in Section 2.7. Secondly, the results for z_i in the observable theory, Eq. (2.5.38), explicitly shows that the idea of spatially varying masses does not alter the ultraviolet singularity structure of the instanton theory, i.e. Eqs (2.5.33) - (2.5.35). A deeper understanding of this problem is provided by the computations in Appendix 2.B where we show that Pauli-Villars regularization retains translational invariance in the sense that the expectation of local operators like $\langle O_i[Q(r)]\rangle$ is independent of r. This aspect of the problem is especially meaningful when dealing with the problem of electron-electron interactions. As is well known, the presence of mass terms generally alters the renormalization of the theory at short distances in this case, i.e. the renormalization group β functions. [32] Finally, on the basis of the theory of observable parameters Eqs (2.5.37) and (2.5.38) we may summarize the results of our instanton computation, Eqs (2.5.31)-(2.5.35), as follows

$$\begin{aligned}\ln\left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} &= - \quad 1+3\ln\frac{3}{2}-(m+n)\left(\gamma+\frac{3}{2}-2\ln 2\right) && (2.5.40)\\ &- \quad 2\pi\sigma_{xx}(\mathcal{M})\pm i\theta(\nu_f) && (2.5.41)\\ &+ \quad z_h(\mathcal{M})\int d\mathbf{r}\mu^2(\mathbf{r})O_h^{\text{inst}}(\mathbf{r}) && (2.5.42)\\ &+ \quad z_s(\mathcal{M})\int d\mathbf{r}\mu^2(\mathbf{r})O_s^{\text{inst}}(\mathbf{r}) && (2.5.43)\\ &+ \quad z_a(\mathcal{M})\int d\mathbf{r}\mu^2(\mathbf{r})O_a^{\text{inst}}(\mathbf{r}). && (2.5.44)\end{aligned}$$

Here, the quantities $O_i^{\text{inst}}(\mathbf{r})$ are the classical expressions given by Eq. (2.3.8). On the other hand, the parameters $\sigma_{xx}(\mathcal{M})$ and $z_i(\mathcal{M})$ are precisely those obtained from the observable theory.

2.6 Instanton manifold

In this Section we first recapitulate the integration over the zero frequency modes following Refs [20] and [21]. In the second part of this Section we address the zero modes describing the $U(m+n)/U(m)\times U(n)$ rotation of the instanton that we sofar have discarded.

2.6.1 Zero frequency modes

The complete expression for Z_{inst}/Z_0 can be written as follows

$$\frac{Z_{\text{inst}}}{Z_0}=\frac{\int\mathcal{D}[Q_{\text{inst}}]\,[Z_{\text{inst}}]^{\text{reg}}}{\int\mathcal{D}[Q_0]\,[Z_0]^{\text{reg}}}. \tag{2.6.1}$$

Here, Q_{inst} denotes the manifold of the instanton parameters as is illustrated in Fig. 2.7

$$\begin{aligned}\int\mathcal{D}[Q_{\text{inst}}] &= A_{\text{inst}}\int d\mathbf{r}_0\int\frac{d\lambda}{\lambda^3}\int\mathcal{D}[U(1)]\\ &\times \int\mathcal{D}\left[\frac{U(m)}{U(1)\times U(m-1)}\right]\int\mathcal{D}\left[\frac{U(n)}{U(1)\times U(n-1)}\right]\\ &\times \int\mathcal{D}\left[\frac{U(m+n)}{U(m)\times U(n)}\right]. && (2.6.2)\end{aligned}$$

The Q_0 represents the zero modes associated with the trivial vacuum

$$\int\mathcal{D}[Q_0]=A_0\int\mathcal{D}\left[\frac{U(m+n)}{U(m)\times U(n)}\right]. \tag{2.6.3}$$

The numerical factors A_{inst} and A_0 are given by

$$\begin{aligned} A_{\text{inst}} &= \langle e_0^4\rangle\langle|e_1|^4\rangle\langle e_0^2|e_1|^2\rangle \left(\langle e_0^2\rangle\langle|e_1|^2\rangle\right)^{m+n-2}\langle 1\rangle^{(m-1)(n-1)}\pi^{-mn-m-n} \\ A_0 &= \langle 1\rangle^{mn}\pi^{-mn} \end{aligned} \quad (2.6.4)$$

where the average $\langle\cdots\rangle$ is with respect to the surface of a sphere

$$\langle f\rangle = \sigma_{xx}\int_{-1}^{1} d\eta \int_0^{2\pi} d\theta f(\eta,\theta). \quad (2.6.5)$$

Notice that in the absence of symmetry breaking terms the integration over $U(m+n)/U(m)\times U(n)$ drops out in the ratio. We shall first discard this integration in the final answer which is then followed by a justification in Section 2.6.2. With the help of the identity [20]

$$\int \mathcal{D}\left[\frac{U(k)}{U(1)\times U(k-1)}\right] = \frac{\pi^{k-1}}{\Gamma(k)}, \quad (2.6.6)$$

we can write the complete result as follows

$$\frac{Z_{\text{inst}}}{Z_0} = \frac{mn}{2}D_{mn}\int d\mathbf{r}_0\int\frac{d\lambda}{\lambda^3}(2\pi\sigma_{xx})^{n+m}\exp S'_{\text{inst}} \quad (2.6.7)$$

where

$$S'_{\text{inst}} = -2\pi\sigma_{xx}(\mathcal{M})\pm i\theta + z_h(\mathcal{M})\int_{\eta\theta} O_h^{\text{inst}} + z_s(\mathcal{M})\int_{\eta\theta} O_s^{\text{inst}} + z_a(\mathcal{M})\int_{\eta\theta} O_a^{\text{inst}}. \quad (2.6.8)$$

The numerical constant D_{mn} is given by

$$D_{mn} = \frac{4}{\pi e}e^{-(m+n)(\gamma+3/2-\ln 2)}\Gamma^{-1}(1+m)\Gamma^{-1}(1+n). \quad (2.6.9)$$

2.6.2 The $U(m+n)/U(m)\times U(n)$ zero modes

To justify the result of Eq. (2.6.7) we next consider the full expression for Z_{inst}/Z_0 that includes the $U(m+n)/U(m)\times U(n)$ rotational degrees of freedom. For simplicity we limit ourselves to the theory in the presence of the z_h field only. We now have

$$\frac{Z_{\text{inst}}}{Z_0} = \frac{\int\mathcal{D}[Q_{\text{inst}}]\exp\{-2\pi\sigma_{xx}\pm i\,\theta + z_h\int_{\eta\theta}\operatorname{tr}\Lambda_h T_0^{-1}\left[R^{-1}\langle q\rangle R\right]^{\text{reg}} T_0\}}{\int\mathcal{D}[Q_0]\exp\{z_h\int_{\eta\theta}\operatorname{tr}\Lambda_h t_0^{-1}\left[\langle q\rangle_0\right]^{\text{reg}} t_0\}}. \quad (2.6.10)$$

Here, we have defined

$$\Lambda_h = \Lambda - \frac{m-n}{m+n}1_{m+n}. \quad (2.6.11)$$

The expectation $\langle\cdots\rangle$ is with respect to the theory of δS_σ, Eq. (2.4.40), whereas $\langle\cdots\rangle_0$ refers to the quantum theory of the trivial vacuum which is obtained from δS_σ

by replacing all operators $O^{(a)}$ by $O^{(0)}$. Let us write the rotational degrees of freedom $T_0 \in U(m+n)$ as follows

$$T_0 = t_0 W, \qquad Q_0 = t_0^{-1} \Lambda t_0. \tag{2.6.12}$$

The quantity t_0 or Q_0 runs over the manifold $U(m+n)/U(m) \times U(n)$ and W stands for the remaining degrees of freedom $U(m)/U(m-1) \times U(1)$, $U(n)/U(n-1) \times U(1)$ and $U(1)$ respectively. We can write

$$\int \mathcal{D}[Q_{\text{inst}}] = A_{\text{inst}} \int d\mathbf{r}_0 \frac{d\lambda}{\lambda^3} \mathcal{D}[Q_0]\mathcal{D}[W] \tag{2.6.13}$$

The quantity $\left[R^{-1}\langle q\rangle R\right]^{\text{reg}}$, unlike $[\langle q\rangle_0]^{\text{reg}}$ in Eq. (2.6.10), is not invariant under $U(m) \times U(n)$ rotations. We therefore perform the integration over W explicitly as follows

$$\left\langle e^{z_h \int_{\eta\theta} \operatorname{tr} \Lambda_h t_0^{-1} W^{-1} [R^{-1}\langle q\rangle R]^{\text{reg}} W t_0} \right\rangle_W = e^{z_h \left\langle \int_{\eta\theta} \operatorname{tr} \Lambda_h t_0^{-1} W^{-1} [R^{-1}\langle q\rangle R]^{\text{reg}} W t_0 \right\rangle_W + \mathcal{O}(z_h^2)} \tag{2.6.14}$$

Here,

$$\langle X[W]\rangle_W = \frac{\int \mathcal{D}[W] X[W]}{\int \mathcal{D}[W]}. \tag{2.6.15}$$

The matrix $[R^{-1}\langle q\rangle R]^{\text{reg}}$ is given as

$$[R^{-1}\langle q\rangle R]^{\text{reg}} = R^{-1}\Lambda R - \frac{2}{\sigma_{xx}}\Lambda G + \frac{1}{\sigma_{xx}}(R^{-1}\Lambda R - \Lambda)(G_1^{(1)} + G_{-1}^{(1)}), \tag{2.6.16}$$

where the matrix G is diagonal in the replica and retarded-advanced spaces

$$G_{pp'}^{\alpha\beta} = \delta_{pp'}^{\alpha\beta} G_p^{(\alpha)} \tag{2.6.17}$$

and the diagonal elements G_p^α are defined as

$$G_1^{(\alpha)} = \begin{cases} \dfrac{m-1}{O^{(1)}} + \dfrac{1}{O^{(2)}}, & \alpha = 1, \\ \dfrac{m-1}{O^{(0)}} + \dfrac{1}{O^{(1)}}, & \alpha = 2, \cdots n. \end{cases} \tag{2.6.18}$$

$$G_{-1}^{(\alpha)} = \begin{cases} \dfrac{n-1}{O^{(1)}} + \dfrac{1}{O^{(2)}}, & \alpha = 1, \\ \dfrac{n-1}{O^{(0)}} + \dfrac{1}{O^{(1)}}, & \alpha = 2, \cdots m. \end{cases} \tag{2.6.19}$$

Integration over W restores the $U(m)\times U(n)$ invariance and we obtain

$$
\begin{aligned}
z_h & \left\langle \int_{\eta\theta} \operatorname{tr}\Lambda_h t_0^{-1} W^{-1} [R^{-1}\langle q\rangle R]^{\mathrm{reg}} W t_0 \right\rangle_W \\
&= -\frac{z_h}{2mn}\int_{\eta\theta}\left(1-\frac{1}{\sigma_{xx}}\left[(m+n-2)Y^{(1)}_{\mathrm{reg}} + 2Y^{(2)}_{\mathrm{reg}}\right]\right)\operatorname{tr}\Lambda_h t_0^{-1}\Lambda X t_0 \\
&+ z_h\int_{\eta\theta}\operatorname{tr}\Lambda_h t_0^{-1}\Lambda\Big[1_{m+n} - \frac{1}{mn\sigma_{xx}}\Big(Y^{(2)}_{\mathrm{reg}} + (m+n-2)Y^{(1)}_{\mathrm{reg}} \\
&\qquad + (m-1)(n-1)Y^{(0)}_{\mathrm{reg}}\Big)X\Big]t_0, \qquad (2.6.20)
\end{aligned}
$$

where

$$
X = (m+n)1_{m+n} + (m-n)\Lambda. \qquad (2.6.21)
$$

By using the traceless nature of the matrix Λ_h, we find

$$
z_h\left\langle \int_{\eta\theta}\operatorname{tr}\Lambda_h t_0^{-1} W^{-1}[R^{-1}\langle q\rangle R]^{\mathrm{reg}} W t_0\right\rangle_W = z_h\int_{\eta\theta}\frac{O_h[Q_0]}{O_h[\Lambda]}\operatorname{tr}\Lambda_h\left[R^{-1}\langle q\rangle R\right]^{\mathrm{reg}} \qquad (2.6.22)
$$

The Q_0 is now the only rotational degree of freedom left in the terms with z_h and the result can therefore be written as follows

$$
\frac{Z_{\mathrm{inst}}}{Z^{(0)}} = \frac{mn}{2}D_{mn}\int d\mathbf{r}_0\int\frac{d\lambda}{\lambda^3}(2\pi\sigma_{xx})^{m+n}\exp\left(-2\pi\sigma_{xx}\pm i\,\theta + \delta S'_{\mathrm{inst}}\right), \qquad (2.6.23)
$$

where

$$
\exp\delta S'_{\mathrm{inst}} = \frac{\int\mathcal{D}[Q_0]\exp\left\{z_h\int_{\eta\theta}\frac{O_h[Q_0]}{O_h[\Lambda]}\operatorname{tr}\Lambda_h\left[R^{-1}\langle q\rangle R\right]^{\mathrm{reg}}\right\}}{\int\mathcal{D}[Q_0]\exp\left\{z_h\int_{\eta\theta}\frac{O_h[Q_0]}{O_h[\Lambda]}\operatorname{tr}\Lambda_h\left[\langle q\rangle_0\right]^{\mathrm{reg}}\right\}}. \qquad (2.6.24)
$$

We can write the result as an expectation with respect to the matrix field variable Q_0,

$$
\exp\delta S'_{\mathrm{inst}} = \left\langle \exp\left\{z_h\int_{\eta\theta}\frac{O_h[Q_0]}{O_h[\Lambda]}\operatorname{tr}\Lambda_h\left[R^{-1}\langle q\rangle R - \langle q\rangle_0\right]^{\mathrm{reg}}\right\}\right\rangle_{Q_0}, \qquad (2.6.25)
$$

where

$$\langle X[Q_0]\rangle_{Q_0} = \frac{\int \mathcal{D}[Q_0] \exp\left\{ z_h \int\limits_{\eta\theta} \frac{O_h[Q_0]}{O_h[\Lambda]} \operatorname{tr} \Lambda_h \left[\langle q\rangle_0\right]^{\text{reg}} \right\} X[Q_0]}{\int \mathcal{D}[Q_0] \exp\left\{ z_h \int\limits_{\eta\theta} \frac{O_h[Q_0]}{O_h[\Lambda]} \operatorname{tr} \Lambda_h \left[\langle q\rangle_0\right]^{\text{reg}} \right\}}. \tag{2.6.26}$$

Notice that by putting the classical value $\langle q\rangle = \langle q\rangle_0 = \Lambda$ in the expression $[\cdots]^{\text{reg}}$ in Eq. (2.6.25) we precisely obtain the quantity O_h^{inst}. We therefore obtain (see also Eq. (2.6.8))

$$\int\limits_{\eta\theta} z_h \operatorname{tr} \Lambda_h \left[R^{-1}\langle q\rangle R - \langle q\rangle_0\right]^{\text{reg}} = \int\limits_{\eta\theta} \left[z_h(\mathcal{M}) + \frac{5}{24\pi\sigma_{xx}} z_h \right] O_h^{\text{inst}}. \tag{2.6.27}$$

This is precisely the result that was obtained before, by fixing $Q_0 = \Lambda$ at the outset. By the same token we write

$$\int\limits_{\eta\theta} z_h \operatorname{tr} \Lambda_h \left[\langle q\rangle_0\right]^{\text{reg}} = \int\limits_{\eta\theta} z_h(\mathcal{M}) O_h(\Lambda) \tag{2.6.28}$$

We have already seen that the quantity that appears in Eq. (2.6.27) differs from $z_h(\mathcal{M})$ by a constant of order $1/\sigma_{xx}$ which is not of interest to us. The final expression for $\delta S'_{\text{inst}}$ can now be written in a more transparent fashion as follows

$$\exp \delta S'_{\text{inst}} = \left\langle \exp\left\{ \frac{O_h[Q_0]}{O_h[\Lambda]} \int\limits_{\eta\theta} z_h(\mathcal{M}) O_h^{\text{inst}} \right\} \right\rangle_{Q_O}, \tag{2.6.29}$$

where instead of Eq. (2.6.26) we now write

$$\langle X[Q_0]\rangle_{Q_0} = \frac{\int \mathcal{D}[Q_0] \exp\{\int_{\eta\theta} z_h(\mathcal{M}) O_h[Q_0]\} X[Q_0]}{\int \mathcal{D}[Q_0] \exp\{\int_{\eta\theta} z_h(\mathcal{M}) O_h[Q_0]\}}. \tag{2.6.30}$$

In summary we can say that as long as one works with mass terms in *curved* space, the rotational degrees of freedom are non-trivial and the integration over the global matrix field Q_0 has to be performed in accordance with Eq. (2.6.30). However, we are ultimately interested in the theory in *flat* space which means that the integral over the unit sphere $\int d\eta d\theta$ in Eq. (2.6.30) is going to be replaced by the integral over the entire plane in flat space, $\int d\mathbf{r}$. This then fixes the matrix variable Q_0 in Eqs (2.6.29) and (2.6.30) to its classical value $Q_0 = \Lambda$. The final results are therefore the same as those that are obtained by putting $Q_0 = \Lambda$ at the outset of the problem.

2.7 Transformation from curved space to flat space

In this Section we embark on the various steps that are needed in order express the final answer in quantities that are defined in flat space. As a first step we have undo the transformation $z_i \to z_i \mu^2(\mathbf{r})$ that was introduced in Section 2.4.2 (see Eq. (2.4.32)). This means that the integrals over η, θ in the expression for S'_{inst}, Eq. (2.6.8), have to be replaced as follows

$$\int d\eta d\theta O_i^{\text{inst}} = \int d\mathbf{r}\mu^2(\mathbf{r}) O_i^{\text{inst}}(\mathbf{r}) \quad \to \quad \int d\mathbf{r} O_i^{\text{inst}}(\mathbf{r}). \tag{2.7.1}$$

The complete expression for the instanton contribution to the free energy is therefore the same as Eq. (2.6.7) but with S'_{inst} now given by

$$S'_{\text{inst}} \quad \to \quad - \quad 2\pi\sigma_{xx}(\mathcal{M}) \pm i\theta \tag{2.7.2}$$

$$+ \int' d\mathbf{r} z_h(\mathcal{M}) O_h^{\text{inst}}(\mathbf{r}) \tag{2.7.3}$$

$$+ \int' d\mathbf{r} z_s(\mathcal{M}) O_s^{\text{inst}}(\mathbf{r}) \tag{2.7.4}$$

$$+ \int' d\mathbf{r} z_a(\mathcal{M}) O_a^{\text{inst}}(\mathbf{r}). \tag{2.7.5}$$

The "prime" on the integral signs reminds us of the fact that the mass terms still formally display a logarithmic divergence in the infrared. However, from the discussion on *constrained* instantons we know that a finite value of z_i generally induces an infrared cut-off on both the spatial integrals $\int d\mathbf{r} O_i^{\text{inst}}(\mathbf{r})$ and the integral over scale sizes λ in the theory. Keeping this in mind, we can proceed and evaluate the expressions for the *physical observables* of the theory, introduced in Section 2.2.4.

2.7.1 Physical observables

According to definitions in Section 2.2.4 we obtain the following results for the parameters σ'_{xx} and θ' (see also Refs [20] and [21])

$$\sigma'_{xx} = \sigma_{xx}(\mathcal{M}) - D_{mn} \int' \frac{d\lambda}{\lambda} (2\pi\sigma_{xx})^{m+n+2} e^{-2\pi\sigma_{xx}(\mathcal{M})} \cos\theta \tag{2.7.6}$$

$$\frac{\theta'}{2\pi} = \frac{\theta}{2\pi} - D_{mn} \int' \frac{d\lambda}{\lambda} (2\pi\sigma_{xx})^{m+n+2} e^{-2\pi\sigma_{xx}(\mathcal{M})} \sin\theta. \tag{2.7.7}$$

Similarly we obtain the z'_i parameters as follows

$$z'_i = z_i(\mathcal{M}) + D_{mn} \int' \frac{d\lambda}{\lambda} (2\pi\sigma_{xx})^{m+n} e^{-2\pi\sigma_{xx}(\mathcal{M})} \cos\theta \tag{2.7.8}$$

$$\times z_i(\mathcal{M}) \left(\frac{mn}{O_i[\Lambda]} \int' \frac{d\mathbf{r}}{\lambda^2} O_i^{\text{inst}}(\mathbf{r}) \right),$$

where $i = a, h$ and s. By using the results of Eqs (2.3.6) and (2.3.7), the expression simplifies somewhat and can be written in a more general fashion as follows

$$z_i' = z_i(\mathcal{M}) + \pi\gamma_i^{(0)} D_{mn} \int' \frac{d\lambda}{\lambda} (2\pi\sigma_{xx})^{m+n} e^{-2\pi\sigma_{xx}(\mathcal{M})} \cos\theta \times z_i(\mathcal{M}) \int' d\mathbf{r} \frac{\mu(\mathbf{r})}{\lambda}. \tag{2.7.9}$$

The important feature of the results of this Section is that the non-perturbative (instanton) contributions are all unambiguously expressed in terms of the perturbative quantities $\sigma_{xx}(\mathcal{M})$, θ and $z_i(\mathcal{M})$.

2.7.2 Transformation $\mu^2(\mathbf{r})\mathcal{M} \to \mu_0$

Next we wish to obtain the results in terms of a spatially *flat* momentum scale μ_0, rather than in the spatially varying quantity $\mu^2(\mathbf{r})\mathcal{M}$ which appears in the Pauli-Villars regularization scheme. For this purpose we introduce the following renormalization group counter terms

$$\begin{aligned} \sigma_{xx}(\mathcal{M}) \to \sigma_{xx}(\mathcal{M}) &\left[1 + \frac{m+n}{2\pi\sigma_{xx}} \ln \frac{\mu(\mathbf{r})\mathcal{M}}{\mu_0}\right] \\ = \sigma_{xx} &\left[1 - \frac{m+n}{2\pi\sigma_{xx}} \ln \frac{\mu_0}{\mu(\mathbf{r})} e^{\gamma}\right] = \sigma_{xx}(\mu(\mathbf{r})), \end{aligned} \tag{2.7.10}$$

$$\begin{aligned} z_i(\mathcal{M}) \to \ z_i(\mathcal{M}) &\left[1 + \frac{\gamma_i^{(0)}}{\sigma_{xx}} \ln \frac{\mu(\mathbf{r})\mathcal{M}}{\mu_0}\right] \\ = \ z_i &\left[1 - \frac{\gamma_i^{(0)}}{\sigma_{xx}} \ln \frac{\mu_0}{\mu(\mathbf{r})} e^{\gamma - 1/2}\right] = z_i(\mu(\mathbf{r})). \end{aligned} \tag{2.7.11}$$

The expression for S'_{inst} now becomes

$$S'_{\text{inst}} \quad \to \quad - \int d\mathbf{r}\sigma_{xx}(\mu(\mathbf{r}))\, \text{tr}(\nabla Q_{\text{inst}}(\mathbf{r}))^2 \pm i\theta \tag{2.7.12}$$

$$+ \int' d\mathbf{r} z_h(\mu(\mathbf{r})) O_h^{\text{inst}}(\mathbf{r}) \tag{2.7.13}$$

$$+ \int' d\mathbf{r} z_s(\mu(\mathbf{r})) O_s^{\text{inst}}(\mathbf{r}) \tag{2.7.14}$$

$$+ \int' d\mathbf{r} z_a(\mu(\mathbf{r})) O_a^{\text{inst}}(\mathbf{r}). \tag{2.7.15}$$

2.7.3 The β functions

Let us first evaluate Eq. (2.7.12) which can be written as

$$\int d\mathbf{r}\, \sigma_{xx}(\mu(\mathbf{r}))\, \text{tr}(\nabla Q_{\text{inst}}(\mathbf{r}))^2 = \int d\mathbf{r} \mu^2(\mathbf{r}) \sigma_{xx}(\mu(\mathbf{r})) = 2\pi\sigma_{xx}(\zeta\lambda), \tag{2.7.16}$$

where

$$\sigma_{xx}(\zeta\lambda) = \sigma_{xx} - \frac{m+n}{2\pi} \ln \zeta\lambda\mu_0 e^{\gamma}, \qquad \zeta = e^2/4. \tag{2.7.17}$$

Notice that the expression for $\sigma_{xx}(\zeta\lambda)$ can be simply obtained from $\sigma_{xx}(\mathcal{M})$ by replacing the Pauli-Villars mass $\mathcal{M}$ according to

$$\mathcal{M} \to \zeta\lambda\mu_0. \tag{2.7.18}$$

We next wish to express the quantity $\sigma'_{xx}(\mathcal{M})$ in a similar fashion. Write

$$\begin{aligned} \sigma'_{xx}(\mathcal{M}) &\to \sigma'_{xx}(\mu'(\mathbf{r})), & (2.7.19)\\ \sigma'_{xx}(\zeta\lambda') &= \frac{1}{2\pi}\int d\mathbf{r}(\mu'(\mathbf{r}))^2 \sigma'_{xx}(\mu'(\mathbf{r})). & (2.7.20)\end{aligned}$$

One can think of the $\mu'(\mathbf{r}) = 2\lambda'/(r^2+\lambda'^2)$ as being a background instanton with a large scale size λ'. The expressions for σ'_{xx} and θ' in flat space can now be written as follows

$$\begin{aligned} \sigma'_{xx}(\zeta\lambda') &= \sigma_{xx}(\zeta\lambda') - D_{mn}\int' \frac{d\lambda}{\lambda}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta, & (2.7.21)\\ \frac{\theta'(\zeta\lambda')}{2\pi} &= \frac{\theta}{2\pi} - D_{mn}\int' \frac{d\lambda}{\lambda}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}(\zeta\lambda)}\sin\theta. & (2.7.22)\end{aligned}$$

In words, the scale size λ' has identically the same meaning for the perturbative and instanton contributions. Notice that $\sigma_{xx}(\zeta\lambda')$ is the same as Eq. (2.7.17) with λ replaced by λ'. Next, introducing an arbitrary scale size λ_0 we can write the perturbative expression $\sigma_{xx}(\zeta\lambda_0)$ as follows

$$\sigma_{xx}(\zeta\lambda') = \sigma_{xx}(\zeta\lambda_0) - \frac{m+n}{2\pi}\ln\frac{\lambda'}{\lambda_0} = \sigma_{xx}(\zeta\lambda_0) - \frac{m+n}{2\pi}\int_{\lambda_0}' \frac{d\lambda}{\lambda}. \tag{2.7.23}$$

On the basis of these results one obtains the following complete expressions for the quantities σ'_{xx} and θ'

$$\begin{aligned} \sigma'_{xx} &= \sigma_{xx}(\zeta\lambda_0) - \int_{\zeta\lambda_0}' \frac{d[\zeta\lambda]}{\zeta\lambda}\left[\frac{m+n}{2\pi} + D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta\right],\\ \frac{\theta'}{2\pi} &= \frac{\theta(\zeta\lambda_0)}{2\pi} - \int_{\zeta\lambda_0}' \frac{d[\zeta\lambda]}{\zeta\lambda} D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}(\zeta\lambda)}\sin\theta. & (2.7.24)\end{aligned}$$

Several remarks are in order. First of all, we have made use of the well known fact that the quantities σ_{xx} in the integral over scale sizes all acquire the same quantum corrections and can be replaced by $\sigma_{xx}(\zeta\lambda)$. Secondly, although the instanton contributions are finite in the ultraviolet, they have nevertheless dramatic consequences for the behavior of the system in the infrared. Equations (2.7.24) determine the renormalization group β functions as follows

$$\begin{aligned} \sigma'_{xx} &= \sigma_{xx}(\zeta\lambda_0) - \int_{\zeta\lambda_0}' \frac{d[\zeta\lambda]}{\zeta\lambda}\beta_\sigma(\sigma_{xx}(\zeta\lambda), \theta(\zeta\lambda)), & (2.7.25)\\ \frac{\theta'}{2\pi} &= \frac{\theta(\zeta\lambda_0)}{2\pi} - \int_{\zeta\lambda_0}' \frac{d[\zeta\lambda]}{\zeta\lambda}\beta_\theta(\sigma_{xx}(\zeta\lambda), \theta(\zeta\lambda)), & (2.7.26)\end{aligned}$$

where

$$\begin{aligned}\beta_\sigma(\sigma_{xx},\theta) &= -\frac{d\sigma_{xx}}{d\ln\lambda} = \frac{m+n}{2\pi} + D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}}\cos\theta,\\ \beta_\theta(\sigma_{xx},\theta) &= -\frac{d(\theta/2\pi)}{d\ln\lambda} = D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}}\sin\theta. \end{aligned}\qquad (2.7.27)$$

These final results which generalize those obtained earlier, on the basis of perturbative expansions (see Eq. (2.2.74)), are universal in the sense that they are independent of the particular regularization scheme that is being used to define the renormalized theory.

2.7.4 Negative anomalous dimension

Eqs (2.7.17), (2.7.18) and (2.7.20) provide a general prescription that should be used to translate the parameters $z_i(\mathcal{M})$ and $z_i'(\mathcal{M})$ into the corresponding quantities $z_i(\zeta\lambda)$ and $z_i'(\zeta\lambda')$ in *flat* space. Analogous to Eqs (2.7.16) and (2.7.20) we introduce the parameters z_i and z_i' associated with scale sizes λ and λ' respectively as follows

$$z_i(\zeta\lambda) = \frac{1}{2\pi}\int d\mathbf{r}\mu^2(\mathbf{r})z_i(\mu(\mathbf{r})), \qquad (2.7.28)$$

$$z_i'(\zeta\lambda') = \frac{1}{2\pi}\int d\mathbf{r}(\mu'(\mathbf{r}))^2 z_i'(\mu'(\mathbf{r})). \qquad (2.7.29)$$

Equation (2.7.28) implies that $z_i(\zeta\lambda)$ is related to $z_i(\mathcal{M})$ according to the prescription of Eq. (2.7.18),

$$z_i(\zeta\lambda) = z_i\left[1 - \frac{\gamma_i^{(0)}}{\sigma_{xx}}\ln\zeta\lambda\mu_0 e^{\gamma-1/2}\right]. \qquad (2.7.30)$$

It is important to emphasize that the final expressions for $z_i(\zeta\lambda)$ and $\sigma_{xx}(\zeta\lambda)$ are consistent with those obtained in dimensional regularization. Next we make use of Eqs (2.7.17), (2.7.28) and (2.7.29) and write the result for z_i', Eq. (2.7.9), as follows

$$z_i'(\zeta\lambda') = z_i(\zeta\lambda') + D_{mn}\int' \frac{d\lambda}{\lambda}(2\pi\sigma_{xx})^{m+n}e^{-2\pi\sigma_{xx}(\zeta\lambda)}A_i\cos\theta. \qquad (2.7.31)$$

The problem that remains is to find the appropriate expression for the quantity A_i which is defined as

$$A_i = \pi\gamma_i^{(0)}\int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda}z_i(\mu(\mathbf{r})). \qquad (2.7.32)$$

Amplitude A_i

To evaluate A_i further it is convenient to introduce the quantity $M_i(\mathbf{r})$ and write

$$A_i = -2\pi^2\gamma_i^{(0)}z_i(\mu(0))\int_{\mu(0)}^{\mu(L')} d[\ln\mu(\mathbf{r})]M_i(\mathbf{r}), \qquad M_i(\mathbf{r}) = \frac{z_i(\mu(\mathbf{r}))}{z_i(\mu(0))}. \qquad (2.7.33)$$

In the language of the Heisenberg ferromagnet $M_i(\mathbf{r})$ represents a spatially varying *spontaneous magnetization* which is measured relative to the center $|\mathbf{r}| = 0$ of the instanton. Notice that for small instanton sizes λ which are of interest to us, the associated momentum scale $\mu(\mathbf{r})$ strongly varies from *large* values $O(\lambda^{-1})$ at short distances ($|\mathbf{r}| \ll \lambda$) to *small* values $O(\lambda/|\mathbf{r}|^2)$ at very large distances ($|\mathbf{r}| \gg \lambda$). Since a continuous symmetry cannot be spontaneously broken in two dimensions [95] the results indicate that $M_i(\mathbf{r})$ generally vanishes for large $|\mathbf{r}|$. We therefore expect the amplitude A_i to remain finite as $L' \to \infty$. This is quite unlike the theory at a classical level where A_i diverges and one is forced to work with the idea of constrained instantons.

Notice that the theory in the replica limit $m = n = 0$ is in many ways special. In this case the anomalous dimension γ_i of mass terms can have an arbitrary sign which means that $M_a(\mathbf{r})$ can *diverge* as $|\mathbf{r}|$ increases. In what follows we shall first deal with the problem of ordinary negative anomalous dimensions, including $\gamma_i = 0$. This is then followed by an analysis of the special cases.

Details of computation

To simplify the discussion of the amplitude A_i we limit ourselves to the theory with $\theta = 0, \pi$ such that $\beta_\theta = 0$ and β_σ, γ_i are functions of σ_{xx} only,

$$\gamma_i(\sigma_{xx}, \theta) \to \gamma_i(\sigma_{xx}), \qquad \beta_\sigma(\sigma_{xx}, \theta) \to \beta_\sigma(\sigma_{xx}). \tag{2.7.34}$$

Write

$$M_i(\mathbf{r}) = \exp\Big\{-\int_{\ln\mu(0)}^{\ln\mu(\mathbf{r})} d[\ln\mu]\gamma_i\Big\}, \tag{2.7.35}$$

then the complete expression for A_i becomes

$$A_i = -2\pi^2\gamma_i^{(0)} z_i(\mu(0)) \int_{\ln\mu(0)}^{\ln\mu(L')} d[\ln\mu(\mathbf{r})] \exp\Big\{-\int_{\ln\mu(0)}^{\ln\mu(\mathbf{r})} d[\ln\mu]\gamma_i\Big\}. \tag{2.7.36}$$

As a next step we change the integrals over $\ln\mu$ into integrals over σ_{xx} and write

$$A_i = z_i(\mu(0))\mathcal{H}_i(\sigma_{xx}(\mu(0))), \tag{2.7.37}$$

where

$$\mathcal{H}_i = -2\pi^2\gamma_i^{(0)} \int_{\sigma_{xx}(\mu(0))}^{\sigma_{xx}(\mu(L'))} \frac{d\sigma_{xx}}{\beta_\sigma(\sigma_{xx})} \exp\Big\{-\int_{\sigma_{xx}(\mu(0))}^{\sigma_{xx}} \frac{d\sigma}{\beta_\sigma(\sigma)}\gamma_i(\sigma)\Big\}. \tag{2.7.38}$$

The meaning of this result becomes more transparent if we write it in differential form. Taking the derivative of $\mathcal{H}_i$ with respect to $\ln\lambda$ we find

$$\left(\beta_\sigma(\sigma_{xx}(\mu(0)))\frac{d}{d\sigma_{xx}(\mu(0))} - \gamma_i(\sigma_{xx}(\mu(0)))\right)\mathcal{H}_i = 2\pi^2\gamma_i^{(0)}\left(1 + M_i(L')\right). \tag{2.7.39}$$

Since in general we have $M_i(L') \to 0$ for $\gamma_i^{(0)} < 0$ we can safely put $L' = \infty$ from now onward. At the same time one can solve Eq. (2.7.39) in the weak coupling limit

$\lambda \to 0$ where $\mu(0)$, $\sigma_{xx}(\mu(0)) \to \infty$. Under these circumstances it suffices to insert for γ_i and β_σ the perturbative expressions

$$\gamma_i(\sigma_{xx}) = \frac{\gamma_i^{(0)}}{\sigma_{xx}} + O(\sigma_{xx}^{-2}), \tag{2.7.40}$$

$$\beta_\sigma(\sigma_{xx}) = \beta_0 + \frac{\beta_1}{\sigma_{xx}} + O(\sigma_{xx}^{-2}), \tag{2.7.41}$$

where

$$\beta_0 = \frac{m+n}{2\pi}, \qquad \beta_1 = \frac{mn+1}{2\pi^2}. \tag{2.7.42}$$

The differential equation (2.7.39) becomes

$$\left(\beta_0 + \frac{\beta_1}{\sigma_{xx}(\mu(0))}\right)\frac{d\mathcal{H}_i}{d\sigma_{xx}(\mu(0))} - \frac{\gamma_i^{(0)}}{\sigma_{xx}(\mu(0))}\mathcal{H}_i = 2\pi^2\gamma_i^{(0)}. \tag{2.7.43}$$

The special solution can be written as follows

$$\mathcal{H}_i^{(1)} = \frac{2\pi^2\gamma_i^{(0)}}{\beta_0 - \gamma_i^{(0)}}\left(\sigma_{xx}(\mu(0)) + \frac{\beta_1}{\gamma_i^{(0)}}\right) \tag{2.7.44}$$

indicating that $\mathcal{H}_i$ can be written as a series expansion in powers of $1/\sigma_{xx}(\mu(0))$. The special solution $\mathcal{H}_i^{(1)}$ does not generally vanish when $\gamma_i^{(0)} \to 0$, however. To obtain the solution with the appropriate boundary conditions we need to solve the homogeneous equation. The result is

$$\mathcal{H}_i^{(0)} = C\left(1 + \frac{\beta_0}{\beta_1}\sigma_{xx}(\mu(0))\right)^{\gamma_i^{(0)}/\beta_0}. \tag{2.7.45}$$

We obtain $\mathcal{H}_i = 0$ for $\gamma_i^{(0)} = 0$ provided we choose

$$C = -\frac{2\pi^2\gamma_i^{(0)}}{\beta_0 - \gamma_i^{(0)}}\frac{\beta_1}{\gamma_i^{(0)}}. \tag{2.7.46}$$

The desired result for $\mathcal{H}_i(\sigma_{xx}(\mu(0)))$ therefore becomes

$$\mathcal{H}_i = \frac{2\pi^2\gamma_i^{(0)}}{\beta_0 - \gamma_i^{(0)}}\left\{\sigma_{xx}(\mu(0)) + \frac{\beta_1}{\gamma_i^{(0)}}\left[1 - \left(1 + \frac{\beta_0}{\beta_1}\sigma_{xx}(\mu(0))\right)^{\gamma_i^{(0)}/\beta_0}\right]\right\}. \tag{2.7.47}$$

As a final step we next express $\sigma_{xx}(\mu(0))$ and $z_i(\mu(0))$ in terms of the flat space quantities $\sigma_{xx}(\zeta\lambda)$ and $z_i(\zeta\lambda)$ respectively. From the definitions of Eqs (2.7.10), (2.7.11), (2.7.17), and (2.7.30), we obtain the following relations

$$\sigma_{xx}(\mu(0)) = \sigma_{xx}(\zeta\lambda)\left[1 + \frac{\beta_0}{\sigma_{xx}(\zeta\lambda)}\ln 2\zeta\right],$$

$$z_i(\mu(0)) = z_i(\zeta\lambda)\left[1 + \frac{\gamma_i^{(0)}}{\sigma_{xx}(\zeta\lambda)}\ln 2\zeta\right]. \tag{2.7.48}$$

For our purposes the correction terms $O(\sigma_{xx}^{-1})$ are unimportant and it suffices to simply replace the $\sigma_{xx}(\mu(0))$ and $z_i(\mu(0))$ by $\sigma_{xx}(\zeta\lambda)$ and $z_i(\zeta\lambda)$ respectively in the final expression for z_i',

$$z_i'(\zeta\lambda') = z_i(\zeta\lambda') + D_{mn}\int' \frac{d\lambda}{\lambda}(2\pi\sigma_{xx})^{m+n} z_i(\zeta\lambda)\mathcal{H}_i(\sigma_{xx}(\zeta\lambda))e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta. \tag{2.7.49}$$

This result solves the problem stated at the outset which is to express the amplitude A_i, Eq. (2.7.32), in terms of the quantities $z_i(\zeta\lambda)$ and $\sigma_{xx}(\zeta\lambda)$, i.e.

$$A_i = z_i(\zeta\lambda)\mathcal{H}_i(\zeta\lambda) \tag{2.7.50}$$

with the function $\mathcal{H}_i$ given by Eq. (2.7.47).

γ_i function

Introducing an arbitrary renormalization point λ_0 as before one can write the perturbative expression for $z_i(\zeta\lambda')$, Eq. (2.7.30), as follows

$$z_i(\zeta\lambda') = z_i(\zeta\lambda_0) + \int'_{\zeta\lambda_0} \frac{d[\zeta\lambda]}{\zeta\lambda}\frac{\gamma_i^{(0)}}{\sigma_{xx}(\zeta\lambda)} z_i(\zeta\lambda), \tag{2.7.51}$$

then Eq. (2.7.49) can be written in terms of the γ_i function as follows

$$z_i'(\zeta\lambda') = z_i(\zeta\lambda_0) + \int'_{\zeta\lambda_0} \frac{d[\zeta\lambda]}{\zeta\lambda}\gamma_i(\sigma_{xx}(\zeta\lambda), \theta(\zeta\lambda)) z_i(\zeta\lambda), \tag{2.7.52}$$

where the complete expression for γ_i equals

$$\gamma_i(\sigma_{xx}, \theta) = \frac{\gamma_i^{(0)}}{\sigma_{xx}} + D_{mn}(2\pi\sigma_{xx})^{m+n}\mathcal{H}_i(\sigma_{xx})e^{-2\pi\sigma_{xx}}\cos\theta, \tag{2.7.53}$$

with

$$\mathcal{H}_i(\sigma_{xx}) = \frac{2\pi^2\gamma_i^{(0)}}{\beta_0 - \gamma_i^{(0)}}\left\{\sigma_{xx} + \frac{\beta_1}{\gamma_i^{(0)}}\left[1 - \left(1 + \frac{\beta_0}{\beta_1}\sigma_{xx}\right)^{\gamma_i^{(0)}/\beta_0}\right]\right\}. \tag{2.7.54}$$

This final expression generalizes the results obtained earlier, on the basis of perturbative expansions (see Eq. (2.2.75)). Moreover, it demonstrates that the replica limit can in general be taken, at least for all operators with $\gamma_i^{(0)} \leqslant 0$.

2.7.5 Free energy

For completeness we next discuss the part S_h of the free energy (2.7.4)

$$S_h = \int' d\mathbf{r} z_h(\mu(\mathbf{r}))O_h^{\text{inst}}(\mathbf{r}). \tag{2.7.55}$$

Notice that S_h can be expressed in terms of the amplitude A_h (see Eq. (2.7.50)),

$$S_h = \frac{2\lambda^2 A_h}{\pi\gamma_h^{(0)}} = 2\lambda^2 z_h(\zeta\lambda)\frac{\mathcal{H}_h(\zeta\lambda)}{\pi\gamma_h^{(0)}}. \tag{2.7.56}$$

Although this result is correct for positive values of m and n, it does not give the right result in the replica limit $m = n = 0$ where both $\gamma_h^{(0)}$ and $\mathcal{H}_i$ go to zero. To obtain the correct result for $\mathcal{H}_h(\zeta\lambda)/(\pi\gamma_h^{(0)})$ in this case we write Eq. (2.7.39) as follows

$$\left(\beta_\sigma(\sigma_{xx}(\mu(0)))\frac{d}{d\sigma_{xx}(\mu(0))} - \gamma_h(\sigma_{xx}(\mu(0)))\right)\left[\frac{\mathcal{H}_h(\zeta\lambda)}{\pi\gamma_h^{(0)}}\right] = -2\pi\left(1 + M_h(L')\right). \tag{2.7.57}$$

In the limit where $\gamma_h = 0$ we have $M_h(L') = 1$ and Eq.(2.7.57) becomes simply

$$\frac{d}{d\ln\lambda}\frac{\mathcal{H}_h(\zeta\lambda)}{\pi\gamma_h^{(0)}} = 4\pi. \tag{2.7.58}$$

The result is given by

$$\frac{\mathcal{H}_h(\zeta\lambda)}{\pi\gamma_h^{(0)}} = 2\pi\ln\tilde{h}^2 + \text{const} \tag{2.7.59}$$

and the expression for S_h becomes

$$S_h = 4\pi z_h\lambda^2\left(\ln\tilde{h}^2 + \text{const}\right). \tag{2.7.60}$$

This result can of course be obtained directly from Eq. (2.7.36) by substituting h^{-1} for L'. In this way we recover the result on the basis of constrained instantons, Section 2.3.2.

2.7.6 Positive anomalous dimension

In case the anomalous dimension of the z_a becomes positive ($\gamma_a^{(0)} > 0$) we have to follow a slightly different route. The quantity to consider in this case is $y_a = z_a^{-1}$ which has an ordinary negative anomalous dimension

$$y_a' = y_a(\mathcal{M}) - \pi\gamma_a^{(0)}D_{mn}\int' \frac{d\lambda}{\lambda}(2\pi\sigma_{xx})^{m+n}e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta\int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda}. \tag{2.7.61}$$

In fact, the analysis for the y_a field proceeds along exactly the same lines as written in the previous Section and the γ function is correctly given by Eq. (2.7.53). Since the γ_a functions of the y_a and z_a fields are identical except for a difference in the overall sign, one can trivially obtain the final result for the z_a field from the known expression for the y_a field. This leads to the following generalization of Eq. (2.7.53)

$$\gamma_i = \frac{\gamma_i^{(0)}}{\pi\sigma_{xx}} + D_{mn}(2\pi\sigma_{xx})^{m+n}\tilde{\mathcal{H}}_i(\sigma_{xx})e^{-2\pi\sigma_{xx}}\cos\theta, \tag{2.7.62}$$

where

$$\tilde{\mathcal{H}}_i = \frac{2\pi^2\gamma_i^{(0)}}{\beta_0 + |\gamma_i^{(0)}|}\left[\sigma_{xx} - \frac{\beta_1}{|\gamma_i^{(0)}|}\left\{1 - \left(1 + \frac{\beta_0}{\beta_1}\sigma_{xx}\right)^{-|\gamma_i^{(0)}|/\beta_0}\right\}\right]. \tag{2.7.63}$$

This final expression which holds for both positive and negative values of $\gamma_i^{(0)}$, including $\gamma_i^{(0)} = 0$, is one of the main results of the present chapter. In the remainder of this chapter we shall embark on the physical consequences of our results.

2.8 Summary of results

Since the theory with an ordinary integer number of field components $m, n \gtrsim 1$ is distinctly different from the one with $0 \leqslant m, n \lesssim 1$ we next discuss the two cases separately.

2.8.1 $m, n \gtrsim 1$

In this case the results for the β and γ functions are essentially the same for all values of m, n and consistent with the Mermin-Wagner-Coleman theorem which says that a continuous symmetry cannot be spontaneously broken in two dimensions. [95] The γ_i functions are all negative and the results can be written as

$$\begin{aligned} \beta_\sigma &= \frac{m+n}{2\pi} + \frac{mn+1}{2\pi^2\sigma_{xx}} + D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \\ \beta_\theta &= D_{mn}(2\pi\sigma_{xx})^{m+n+2}e^{-2\pi\sigma_{xx}}\sin 2\pi\sigma_{xy}, \end{aligned} \tag{2.8.1}$$

$$\begin{aligned} \gamma_{s,a} &= -\frac{m+n\pm 1}{\pi\sigma_{xx}} - \frac{2\pi(m+n\pm 1)}{3(m+n)\pm 2}D_{mn}(2\pi\sigma_{xx})^{m+n+1}e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \\ \gamma_h &= -\frac{m+n}{2\pi\sigma_{xx}} - \frac{\pi}{2}D_{mn}(2\pi\sigma_{xx})^{m+n+1}e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \end{aligned} \tag{2.8.2}$$

where we remind

$$D_{mn} = \frac{4}{\pi e}e^{-(m+n)(\gamma+3/2-\ln 2)}\Gamma^{-1}(1+m)\Gamma^{-1}(1+n). \tag{2.8.3}$$

We see that along the $\theta = 0$ line the instanton contributions $\propto e^{-2\pi\sigma_{xx}}$ generally reinforce the results obtained from ordinary perturbation theory. This means that the instanton contribution generally tends to make the β_σ function more *positive* and the γ_i functions more *negative*. Therefore, upon decreasing the momentum scale - or in the limit of large distances - the instantons *enhance* the flow of the system toward the *strong coupling* phase or *symmetric* phase.

Notice that for $\theta = \pi$ the perturbative and non-perturbative contributions carry an opposite sign indicating that the $\theta = \pi$ line displays infrared properties that are generally different from those along $\theta = 0$. The instanton contributions indicate that the renormalization group flow is generally controlled by the weak coupling fixed point located at $\sigma_{xx} = \infty$ and the strong coupling fixed points located at $\sigma_{xx} = 0$. The large

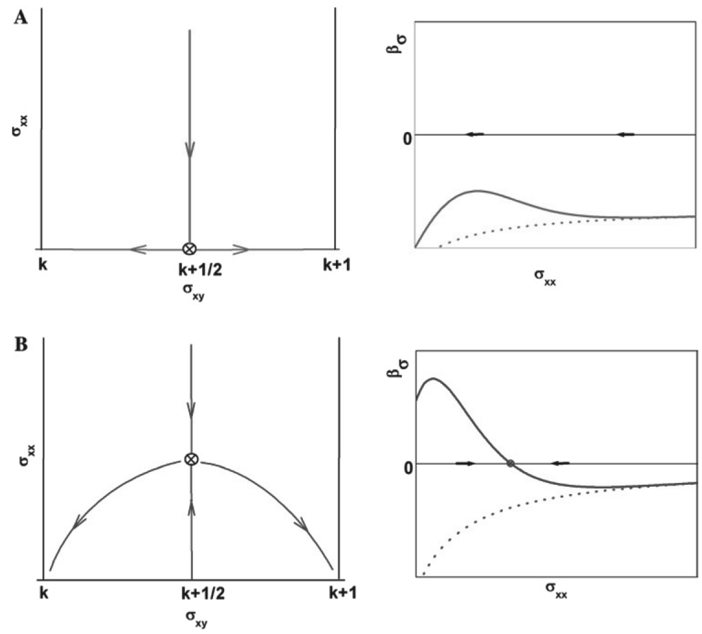


Figure 2.8: The renormalization group flow diagram for different values of field components n and m. (a) The results for large values $n, m \gtrsim 1$. (b) The results for small values $n, m \lesssim 1$.

N expansion can be used as a stage setting for what happens in the strong coupling symmetric phase. The results of Ref. [46] indicate that although the transition at $\theta = \pi$ is a first order one, there is nevertheless a diverging length scale in the problem

$$\xi = \xi_0 |\theta - \pi|^{-1/2}, \tag{2.8.4}$$

where the exponent $1/2$ equals $1/D$ with $D = 2$ the dimension of the system. The following scaling results have been found [46]

$$\sigma_{xx}(X,Y) = \frac{e^{-X}}{e^{-2X} + 1 + 2Ye^{-X}} Y, \tag{2.8.5}$$

$$\sigma_{xy}(X,Y) = k(\nu_f) + \frac{1 + Ye^{-X}}{e^{-2X} + 1 + 2Ye^{-X}}, \tag{2.8.6}$$

with

$$X = \frac{L^2}{\xi_0^2} |\theta(\nu_f) - \pi| = \pm \left(\frac{L}{\xi}\right)^2, \tag{2.8.7}$$

$$Y \propto \frac{\sinh X}{X} \exp\left(-\frac{L}{\xi_M}\right). \tag{2.8.8}$$

These scaling results are in many ways the same as those expected for second order transitions and indicate that the θ vacuum generically displays all the fundamental features of the quantum Hall effect.

Equations (2.8.5)-(2.8.8) are valid in the regime where $L \gg \xi_M$ where ξ_M denotes the correlation length that describes the cross-over between the weak coupling Goldstone singularities at short distances (described by Eqs (2.8.2)) and the quantum Hall singularities (described by Eqs (2.8.5)-(2.8.8)) that generally occur at much larger distances only. An overall sketch of the renormalization in the σ_{xx}, σ_{xy} conductivity plane is given in Fig. 2.8a. We see that the infrared of the system is generally controlled by the *stable* quantum Hall fixed points with $\sigma_{xy} = k(\nu_f)$ and the *unstable* fixed points located at $\sigma_{xy} = k(\nu_f) + 1/2$ that describe the singularities of the *plateau transitions*.

2.8.2 $0 \leqslant m, n \lesssim 1$

Figure 2.8b illustrates how for small values of m and n the instanton contributions to the β functions produce an infrared zero along the line $\theta = \pi$ with $\sigma_{xx} = \mathcal{O}(1)$. This indicates that the transition at $\theta = \pi$ now becomes a true *second order* one. In Fig. 2.9 we have plotted the lines in the m, n plane that separate the regimes of *first order* and *second order* transitions.

To proceed we first address the replica limit $m = n = 0$ where the theory describes the physics of the disordered electron gas. The results for the β and γ functions can

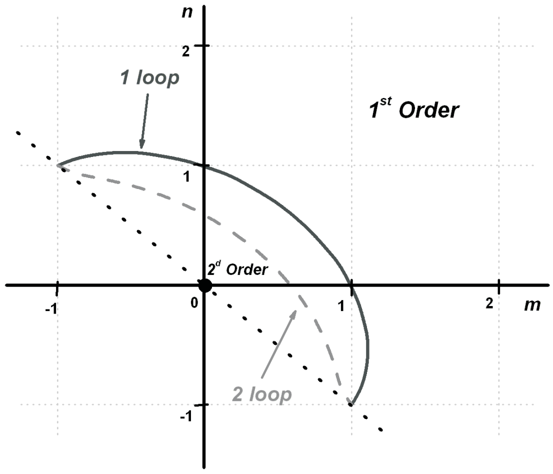


Figure 2.9: Nature of the transition at $\sigma_{xy} = k + 1/2$. The solid red line separates the regions of the first and second order transitions as predicted by β_σ, β_θ, Eqs (2.8.2). The dashed blue line is obtained by extending the perturbative contributions to β_σ to include the two-loop results.

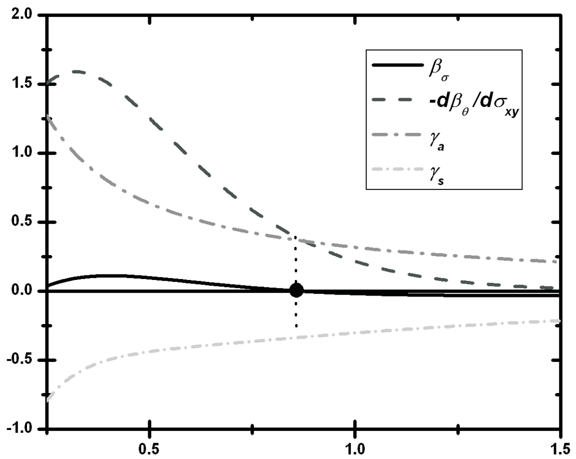


Figure 2.10: The plot of β_σ, $-d\beta_\theta/d\sigma_{xy}$, γ_a and γ_s as functions of σ_{xx} at $\sigma_{xy} = k+1/2$. Blue dot denotes the fixed point.

be summarized as follows

$$\begin{aligned} \beta_\sigma &= \frac{1}{2\pi^2\sigma_{xx}} + \frac{16\pi}{e}\sigma_{xx}^2 e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \\ \beta_\theta &= \frac{16\pi}{e}\sigma_{xx}^2 e^{-2\pi\sigma_{xx}}\sin 2\pi\sigma_{xy}, \\ \gamma_s &= -\frac{1}{\pi\sigma_{xx}} - \frac{8\pi}{e}\sigma_{xx} e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \\ \gamma_a &= \frac{1}{\pi\sigma_{xx}} + \frac{8\pi}{e}\sigma_{xx} e^{-2\pi\sigma_{xx}}\cos 2\pi\sigma_{xy}, \\ \gamma_h &= 0. \end{aligned} \tag{2.8.9}$$

Here we have extended the quantity β_σ to include the perturbative results obtained to two loop order.

We see that along the lines where σ_{xy} is an integer, the instanton contributions generally enhance the tendency of the system toward *Anderson localization*. Like the theory with large values of m and n, this means that the β_σ function renders more *positive* in the presence of instantons. Unlike the previous case, however, the γ_h function is now identically *zero* whereas the γ_a function becomes manifestly *positive*. As already mentioned earlier, these results are extremely important and dictated by the physics of the problem.

When σ_{xy} is close to half-integer values the perturbative and non-perturbative contributions generally carry the opposite sign. The critical infrared fixed point $\beta_\sigma = 0$ at $\sigma^*_{xx} \approx 0.88$ (Fig. 2.10) indicates that the electron gas *de-localizes*. Notice that the situation is in many ways identical to the *mobility edge* problem in $2+\epsilon$ spatial dimensions except for the fact that the transition is now being approached from the *insulating* side only.

The consequences of this fixed point for the physics of the electron gas can be summarized as follows.

1. Localization length. Introducing the quantities

$$\Delta_\theta = \sigma_{xy}^0 - k(\nu_f^*) - \frac{1}{2} = \nu_f - \nu_f^*, \qquad \Delta_\sigma = \sigma_{xx}^0 - \sigma_{xx}^* \tag{2.8.10}$$

which denote the linear environment of the critical fixed point then the divergent localization length ξ of the electron gas can be expressed as

$$\xi = \xi_0 |\Delta_\theta|^{-\nu}. \tag{2.8.11}$$

Here, ξ_0 is an arbitrary length scale determined by the microscopics of the electron gas and the localization length exponent ν is obtained as

$$\nu = -\left[\left(\frac{d\beta_\theta}{d\sigma_{xy}}\right)^*\right]^{-1} \approx 2.8. \tag{2.8.12}$$

2. Transport parameters. Next, the *ensemble averaged* transport parameters of the electron gas can be written as regular functions of two scaling variables X and Y (see Ref. [22])

$$\sigma_{xx} = \sigma_{xx}(X,Y), \qquad \sigma_{xy} = \sigma_{xy}(X,Y), \tag{2.8.13}$$

where

$$X = \left(\frac{L}{\xi_0}\right)^{y_\theta} \Delta_\theta, \qquad Y = \left(\frac{L}{\xi_0}\right)^{y_\sigma} \Delta_\sigma \tag{2.8.14}$$

with

$$y_\theta = \nu^{-1} \approx 0.36, \qquad y_\sigma = -\left(\frac{d\beta_\sigma}{d\sigma_{xx}}\right)^* \approx -0.17. \tag{2.8.15}$$

Besides the aforementioned results obtained from the large N expansions there exists, until to date, no knowledge on the explicit form of the scaling functions. Nevertheless, there indications which tell us that the functions $\sigma_{xx}(X,Y)$ and $\sigma_{xy}(X,Y)$ are, in fact, very similar for all the cases of interest. For example, recent experiments on quantum criticality in the quantum Hall regime have shown that the scaling functions for the true, interacting electron gas are given by [53]

$$\sigma_{xx}(X,Y) = \frac{e^{-X}}{e^{-2X} + 1 + 2Ye^{-X}}, \tag{2.8.16}$$

$$\sigma_{xy}(X,Y) = k(\nu_f) + \frac{1 + Ye^{-X}}{e^{-2X} + 1 + 2Ye^{-X}}. \tag{2.8.17}$$

Notice that these results are strikingly similar to those obtained from the large N expansion, in spite of the fact that the two systems in question are physically totally different. Both systems are realizations of an instanton vacuum, however. The results therefore indicate that the list of *super universal* features of the theory is likely to be extended to include the actual form of the scaling functions $\sigma_{xx}(X,Y)$ and $\sigma_{xy}(X,Y)$.

3. Inverse participation ratio. The electronic wave functions $\Psi_E(\mathbf{r})$ near the Fermi energy E define a quantity

$$P^{(2)} = \left\langle \int d\mathbf{r} |\Psi_E(\mathbf{r})|^4 \right\rangle, \tag{2.8.18}$$

which is called the *inverse participation ratio.* Here, the brackets denote the average with respect to the impurity ensemble. $P^{(2)}$ is a measure for the inverse of the volume that is taken by these electronic levels. It is expected to be zero for *extended* states and finite for *localized* states. Hence, this quantity can be used as an alternative probe for Anderson localization. In the language of the non-linear sigma model we can expressed this quantity in terms of the antisymmetric operator O_a according to

$$P^{(2)} \propto \left\langle z_h O_a(Q) \right\rangle_{z_h \to 0} = \xi^{-D_2} f_2 \left(z_h \xi^d, \Delta_\sigma \xi^{y_\sigma} \right) |_{z_h \to 0}. \tag{2.8.19}$$

Here the expectation is with respect to the non-linear sigma model in the presence of the operator O_h. The exponent D_2 equals

$$D_2 = D - \gamma_a^* \approx 1.67 \tag{2.8.20}$$

with $D = 2$ denoting the dimension of the system. It is interesting to remark that the numerical value of D_2 is largely determined by the perturbative contribution to γ_a, the instanton part merely contributing an amount of roughly three percent.

4. Multifractality. Following Wegner [74] the generalized *inverse participation ratio* is defined as follows

$$P^{(q)} = \left\langle \int d\mathbf{r} |\Psi_E(\mathbf{r})|^{2q} \right\rangle. \tag{2.8.21}$$

The mapping of $P^{(q)}$ onto the non-linear sigma model now involves composite operators with q matrix field variables Q. On the basis of this mapping one expects a scaling behavior of the form

$$P^{(q)} \propto \xi^{-(q-1)D_q} f_q(z_h \xi^d, \Delta_\sigma \xi^{y_\sigma}). \tag{2.8.22}$$

The generalized dimension D_q can be written in terms of the anomalous dimension γ_q of the composite operators according to

$$D_q = D - \frac{\gamma_q^*}{q-1}. \tag{2.8.23}$$

The following perturbative expression is known [96]

$$\gamma_q = \frac{q(q-1)}{2\pi\sigma_{xx}} + O(\sigma_{xx}^{-2}). \tag{2.8.24}$$

Based on our results, Eqs (2.7.62) and (2.7.63), we ready generalize the expression for γ_q to include the effect of instantons. We obtain

$$\gamma_q = \frac{q(q-1)}{2\pi\sigma_{xx}} - \frac{8\pi}{e}\sigma_{xx} B\Big(\pi\sigma_{xx} q(q-1)\Big) e^{-2\pi\sigma_{xx}}, \tag{2.8.25}$$

where function $B(x)$ is obtained from the $\beta_0 \to 0$ limit of the $\tilde{\mathcal{H}}(\sigma_{xx})$ function and is equal to

$$B(x) = 1 - \frac{1 - e^{-x}}{x}. \tag{2.8.26}$$

The validity of Eq. (2.8.25) is likely to be limited to the case of small values of q only, presumably $q \ll 10$. This is so because the higher order terms in the perturbative series for γ_q are rapidly increasing with increasing values of q. [96] The result for D_q becomes

$$D_q = D - \frac{q}{2}\gamma_a^* \left[1 - \frac{\gamma_a^*}{2q(q-1)} B\left(\frac{q(q-1)}{\gamma_a^*}\right)\right],$$

where $\gamma_a^* = 1/\pi\sigma_{xx}^* \approx 0.36$. An important feature of these results is that $\gamma_q^* \to 0$ or $D_q \to D$ as q approaches zero. This property permits one to express the multifractal properties of generalized inverse participation ratios in terms of the $f(\alpha)$ singularity spectrum as follows. Introducing the variable

$$\alpha_q = D_q - (q-1)\frac{dD_q}{dq} \tag{2.8.27}$$

then $f(\alpha_q)$ is given by

$$f(\alpha_q) = q\alpha_q - (q-1)D_q. \tag{2.8.28}$$

To proceed let us first consider the perturbative contributions to the generalized dimension D_q, Eqs (2.8.23) and (2.8.24). The result can be cast in the familiar form

$$f(\alpha) = D - \frac{(\alpha - \alpha_0)^2}{4(\alpha_0 - D)}, \tag{2.8.29}$$

where

$$\alpha_0 = D + \frac{1}{2}\gamma_a^*. \tag{2.8.30}$$

The result of Eq. (2.8.29) generally describes the $f(\alpha)$ singularity spectrum near its maximum value at $\alpha = \alpha_0$ only. Next, by taking into account the effects from instantons (see Eq. (2.8.25)) we find that the expression of Eq. (2.8.30) is only slightly modified in that the quantity α_0 is now given by

$$\alpha_0 = D + \frac{3}{8}\gamma_a^*. \tag{2.8.31}$$

Whereas Eq. (2.8.30) leads to a numerical value $\alpha_0 = 2.18$, Eq. (2.8.31) gives $\alpha_0 = 2.14$. Again we find that the instanton contribution is numerically a small fraction of the final answer. We attribute important significance to these results since they are an integral part of the final conclusion of this chapter which says that the quantum critical aspects of the electron gas are within the range of weak coupling expansion techniques.

As a final remark it should be mentioned that the tails of the $f(\alpha)$ singularity spectrum, which are generally controlled by the D_q with large values of q, is beyond the scope of the present analysis.

Table 2.3: The critical and multifractal exponents for the noninteracting electrons.

Exponent	Instantons	Numerical results
ν	2.8 ± 0.4	2.5 ± 0.5 [76], 2.34 ± 0.04 [77] 2.3 ± 0.08 [78], 2.4 ± 0.1 [79] 2.2 ± 0.1 [80], 2.43 ± 0.18 [81]
y_σ	-0.17 ± 0.02	-0.38 ± 0.04 [88], -0.4 ± 0.1 [89]
D_2	1.67 ± 0.03	1.62 ± 0.04 [82], 1.62 ± 0.02 [83] 1.40 ± 0.02 [85]
α_0	2.14 ± 0.02	2.30 ± 0.07 [86], 2.29 ± 0.02 [87] 2.260 ± 0.003 [89]

2.8.3 Comparison with numerical work

In Table 2.3 we compare our results for the critical indices ν, y_σ, D_2 and α_0 with those extracted from numerical simulations on the electron gas. The agreement is in many respects spectacular, indicating that our instanton analysis captures some of the most essential and detailed physics of the problem.

The study of the multifractal aspects of the problem - as done in the previous Section - already indicates that higher order instanton effects are numerically of minor importance. To generally understand why the effects of multi-instanton configurations do not carry much weight, however, one might argue that the very existence of the critical fixed point at $\sigma_{xx} = \sigma_{xx}^*$ and $\sigma_{xy} = k(\nu_f) + 1/2$ implies that the corrections due to multi-instantons typically involve factors like $e^{-4\pi\sigma_{xx}^*}$ which are negligible as compared to the leading order result $e^{-2\pi\sigma_{xx}^*}$. To specify the thought we introduce the quantity

$$\epsilon = \frac{16\pi}{e}\sigma_{xx}e^{-2\pi\sigma_{xx}}. \tag{2.8.32}$$

In the neighborhood of the lines $\sigma_{xy} = k(\nu_f) + 1/2$ the renormalization group func-

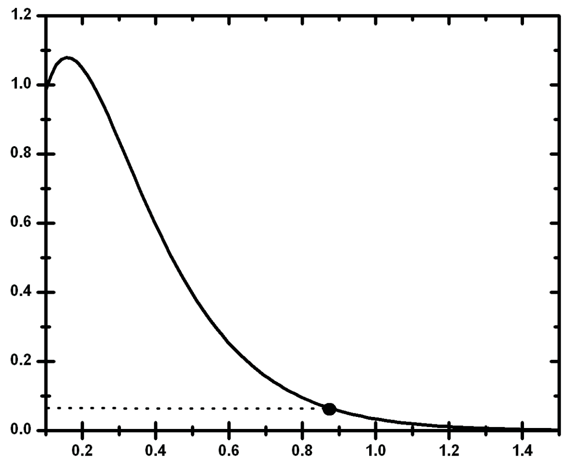


Figure 2.11: The dependence of ϵ on the σ_{xx} for the model of free electrons $m = n = 0$.

tions (2.8.9) can then be written in the form of an ordinary ϵ-expansion

$$\beta_\sigma = -\epsilon\sigma_{xx} + \frac{1}{2\pi^2\sigma_{xx}} + [[\beta_\sigma]]\,, \tag{2.8.33}$$

$$\frac{d\beta_\theta}{d\sigma_{xy}} = -2\pi\epsilon\sigma_{xx} + \left[\left[\frac{d\beta_\theta}{d\sigma_{xy}}\right]\right], \tag{2.8.34}$$

$$\gamma_s = \frac{\epsilon}{2} - \frac{1}{\pi\sigma_{xx}} + [[\gamma_s]]\,, \tag{2.8.35}$$

$$\gamma_a = -\frac{\epsilon}{2} + \frac{1}{\pi\sigma_{xx}} + [[\gamma_a]]\,, \tag{2.8.36}$$

$$\gamma_h = 0. \tag{2.8.37}$$

The brackets $[[\cdots]]$ in the β and γ functions generally stand for the higher order terms that contain the more complex contributions from the multi-instanton configurations. It is clear, however, that in order for the one-instanton approach to be successful the parameter ϵ must be a small quantity. By the same token, the expressions $[[\cdots]]$ should be well approximated by inserting the leading order corrections as obtained from ordinary perturbative expansions [57]

$$[[\beta_\sigma]] = \frac{3}{8\pi^4\sigma_{xx}^3} + \cdots\,, \tag{2.8.38}$$

$$\left[\left[\frac{d\beta_\theta}{d\sigma_{xy}}\right]\right] = 0, \tag{2.8.39}$$

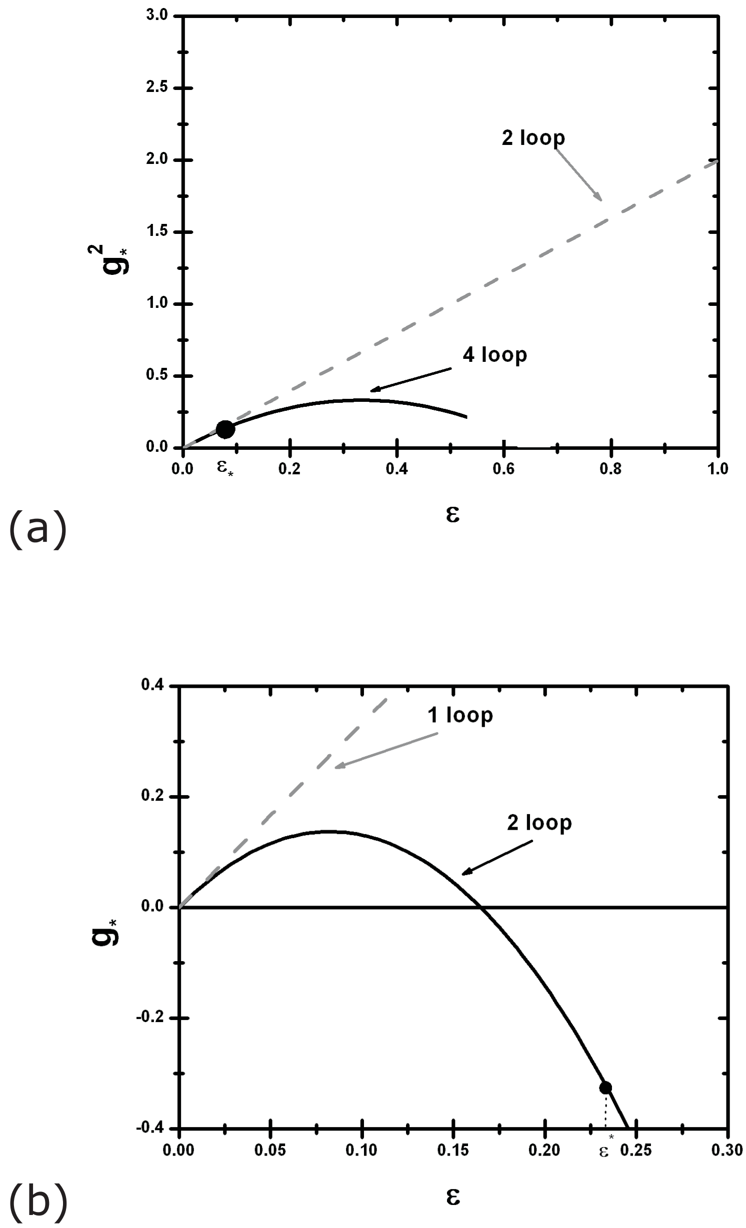


Figure 2.12: The dependence of the critical point $g_* = 1/\pi\sigma_{xx}^*$ from ϵ. (a) The model of free electrons, $m = n = 0$. (b) The model with $m = n = 0.3$.

and [96]

$$[[\gamma_s]] = -\frac{3(1+c)}{8\pi^3\sigma_{xx}^3} + \cdots , \tag{2.8.40}$$

$$[[\gamma_a]] = \frac{3}{8\pi^3\sigma_{xx}^3} + \cdots . \tag{2.8.41}$$

Here, the constant c has not yet been computed explicitly. The critical fixed point can then be obtained formally as an expansion in powers of ϵ,

$$(\pi\sigma_{xx}^*)^{-2} = 2\epsilon\left(1 - \frac{3}{2}\epsilon\right). \tag{2.8.42}$$

Similarly, the critical indices can be obtained as

$$\begin{aligned}
y_\theta &= \sqrt{2\epsilon}\Big(1 + \frac{3}{4}\epsilon\Big), \\
y_\sigma &= -\sqrt{2\epsilon}\Big(1 - \frac{3}{2}\sqrt{2\epsilon} + \epsilon\Big), \\
\gamma_s^* &= -\sqrt{2\epsilon}\Big(1 - \frac{1}{4}\sqrt{2\epsilon} + \frac{3c}{4}\epsilon\Big), \\
\gamma_a^* &= \sqrt{2\epsilon}\Big(1 - \frac{1}{4}\sqrt{2\epsilon}\Big), \\
\gamma_h^* &= 0 .
\end{aligned} \tag{2.8.43}$$

A simple computation next tells us that at the fixed point σ_{xx}^* the parameter ϵ equals $\epsilon \approx 0.07$ which is indeed a "*small*" quantity in every respect. In Fig. 2.11 we have plotted the curve ϵ with varying values of σ_{xx} according to Eq. (2.8.32). We see that the fixed point values σ_{xx}^* and ϵ^* are located well inside the "exponential tail" region where the curve is dominated by the instanton factors $e^{-2\pi\sigma_{xx}}$. In Fig. 2.12(a) we have plotted the critical value $g_*^2 = (\pi\sigma_{xx}^*)^{-2}$ with varying values of ϵ in two-loop approximation where the correction term $[[\beta_\sigma]]$ in β_σ (see Eqs (2.8.33)) has been dropped, as well as in four-loop approximation where this term is retained. The results clearly indicate that the theory with $m = n = 0$ lies well inside the regime of validity of the weak coupling expansions as considered in this chapter.

Finally, we have used the highest order correction terms in the series of Eq. (2.8.43) as an estimate for the uncertainty in the exponent values based on instantons (see Table 2.3).

2.8.4 Continuously varying exponents and a conjecture

The critical behavior at $\theta = \pi$ changes continuously as the value of m and n increases. In Figs 2.13 and 2.14 we plot our instanton results for the exponents ν and y_σ in the interval $0 \leqslant m = n \lesssim 0.3$. In Fig. 2.15 the results are presented for the three different exponents γ_i^* with varying $m = n$. Of interest is the critical end-point $m = n \approx 0.3$ or, more generally, the boarder line in the m, n plane that separates the regimes of *second order* and *first order* phase transitions (Fig. 2.9). From the mechanism by which the critical fixed point in β_σ is generated (see Fig. 2.8) it is clear that the boarding line is defined by the points m, n where the exponent y_σ renders *marginal*.

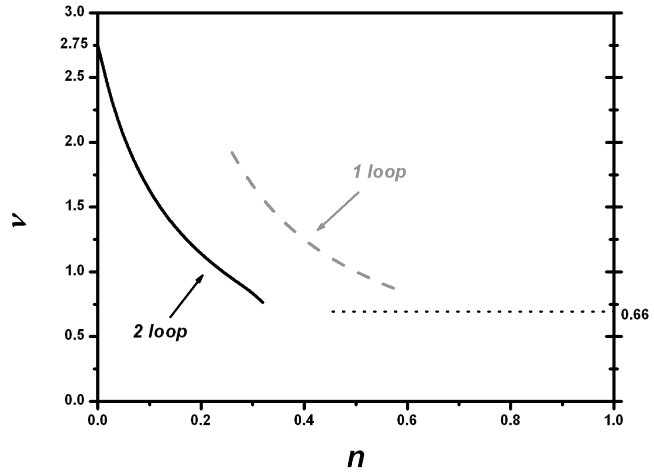


Figure 2.13: The correlation length exponent ν with varying values of $0 \leqslant m = n \lesssim 1$. For comparison we have plotted the value $\nu = 2/3$ which is known to be the exact result for $m = n = 1$, see text.

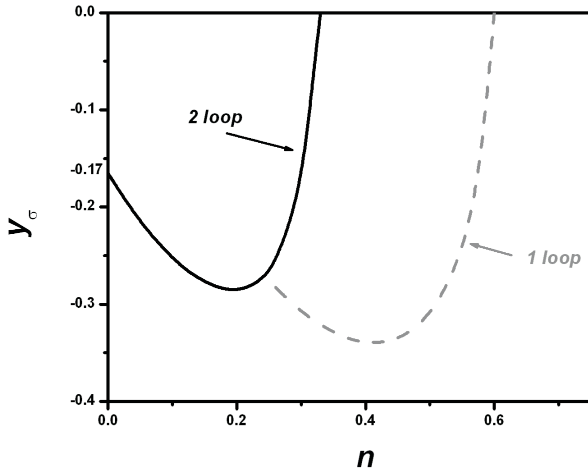


Figure 2.14: The irrelevant exponent y_σ with varying values of $0 \leqslant m = n \lesssim 1$, see text.

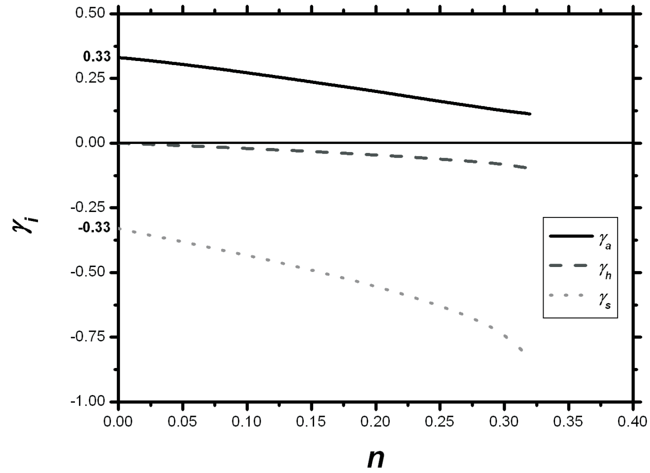


Figure 2.15: The anomalous dimension γ_s^*, γ_a^* and γ_h^* with varying values of $0 \leqslant m = n \lesssim 1$, see text.

Along this line the expansion of the β functions about the fixed point values σ_{xx}^* and $\sigma_{xy}^* = k(\nu_f) + 1/2$ can be written as follows

$$\begin{aligned} \beta_\sigma &= -\alpha_2 \left(\sigma_{xx} - \sigma_{xx}^*\right)^2 \\ \beta_\theta &= -\nu^{-1}\left(\sigma_{xy} - k(\nu_f) - \frac{1}{2}\right). \end{aligned} \tag{2.8.44}$$

Here, ν equals the correlation length exponent and α_2 is a positive constant. Notice that the variable $\sigma_{xx} - \sigma_{xx}^* > 0$ scales to zero in the infrared and, hence, this quantity is *marginally irrelevant*. On the other hand, the variable $\sigma_{xx} - \sigma_{xx}^* < 0$ increases with increasing length scales and this quantity is therefore *marginally relevant*. A characteristic feature of marginally relevant/irrelevant scaling variables is that the critical correlation functions of the system are no longer given by simple power laws but, rather, they acquire logarithmic corrections.

The problem, however, is that the exact location of this line in the m, n plane is beyond the scope of the present analysis. It is easy to see, for example, that the quantity ϵ (Eq. 2.8.32), unlike the theory with $m = n = 0$, cannot be considered as a "small" parameter when the values of m and n increase. For illustration we compare in Fig. 2.12 the critical fixed point $g_* = 1/\pi\sigma_{xx}^*$ with varying values of ϵ for two different values of $m = n$. Whereas the theory with $m = n = 0$ lies well inside the range of validity of the "ϵ expansion", this is no longer the case when $m = n \approx 0.3$. At the same time, the expansion is no longer controlled by the exponential tails $e^{-2\pi\sigma_{xx}}$ when the values of m and n increase. All this indicates that the exact location of the critical boarding line in the m, n plane necessarily involves a detailed knowledge of multi-instanton effects.

In spite of these and other complications, however, there are very good reasons to believe that our instanton results display all the qualitative features of quantum criticality in the theory with small m, n. It is well known, for example, that the $O(3)$ model at $\theta = \pi$ exhibits a second order phase transition. Since the (algebraic) correlation functions turn out to have logarithmic corrections in this case [97] it is natural to associate the point $m = n = 1$ with the aforementioned boarder line in the m, n plane. This means that the regime of second order phase transitions actually spans the interval $0 \leqslant m = n \leqslant 1$ in Figs 2.13 - 2.15, rather than $0 \leqslant m = n \leqslant 0.3$ as predicted by our instanton analysis. At the same time one expects the global phase diagram (Fig. 2.9) to be slightly modified and replaced by Fig. 2.3 in the exact theory.

2.9 Conclusions

The results of this chapter are an integral part of the general statement which says the physics of the quantum Hall effect is a *super universal* strong coupling feature of the topological concept of an instanton vacuum in asymptotically free field theory. We have shown that the instanton gas unequivocally describes the cross-over between the *Goldstone* phase where perturbation theory applies and the completely non-perturbative regime of the *quantum Hall effect* that generally appears in the limit of much larger distances only.

As a major technical advance we have obtained not only the non-perturbative β functions of the theory but also the anomalous dimension or γ function associated with mass terms. Amongst many other things, the results of this chapter lay the foundation for a non-perturbative analysis of the electron gas that includes the effects of electron-electron interactions.

Whereas the theory with finite $m, n > 1$ has been discussed in detail in Ref. [46], on the basis of the large N expansion, the main objective of the present analysis has been the theory with $0 \leqslant m, n \leqslant 1$. In this case one expects a *second order* phase transition at $\theta = \pi$ that is characterized by a finite value of σ_{xx}^* as well as continuously varying exponents. Although much of the global phase structure of the theory was either known or anticipated in previous papers on the subject [20, 21, 46], we shall show that the technical advances made in this chapter are nevertheless extremely important. The main difference with the previous situation is that certain ambiguities in the theory have been removed, notably the *definition* of the β functions. Moreover, by extending the theory to include the renormalization group γ functions we have not only resolved some of the outstanding problems in the instanton methodology, but also facilitated detailed comparisons between the predictions of the theory on the one hand, and the data known from numerical experiments on the free electron gas on the other.

We have seen that the results of the theory with $m = n = 0$ agree remarkably well with the exponent values extracted from numerical simulations (Table 2.3). Notice that this is the first time - ever since the θ parameter was introduced in the theory of the quantum Hall effect by one of us - that accurate estimates for the critical indices of the plateau transitions have been obtained *analytically*. The advances reported in this chapter are obviously important since they teach us something fundamental about the strong coupling problems in QCD where the algebra is the same, but experiments are impossible.

2.A Computation of the matrix elements

For the computations of the quantum fluctuations for the mass terms O_i we have used different matrix elements between the eigenfunctions of the $O^{(a)}$ operators. Let us define matrix elements of a function $f(\eta,\theta)$ as

$$ {}_{(a)}\left\langle J,M|f(\eta,\theta)|M^{'},J^{'}\right\rangle_{(b)} = \int d\eta d\theta\, \Phi^{(a)}_{J,M}(\eta,\theta) f(\eta,\theta)\bar{\Phi}^{(b)}_{J',M'}(\eta,\theta), \qquad (2.A.1) $$

where $a,b = 0,1,2$. By using the following relation for the Jacobi polynomials [98]

$$ (2n+\alpha+\beta)P_n^{(\alpha-1,\beta)}(x) = (n+\alpha+\beta)P_n^{(\alpha,\beta)}(x) - (n+\beta)P_n^{(\alpha,\beta)}(x), \qquad (2.A.2) $$

and the normalization condition [98]

$$ \int_{-1}^{1} dx(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x) $$
$$ = \delta_{n,m}2^{\alpha+\beta+1}\frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)\Gamma(n+1)\Gamma(\alpha+\beta+n+1)} \qquad (2.A.3) $$

we shall find the following results for different matrix elements. For e_0 and e_1 functions we obtain

$$ {}_{(0)}\langle J,M|e_1^*|M-1,J\rangle_{(1)} = \frac{1}{\sqrt{2}}\sqrt{\frac{J+M}{2J+1}}, \qquad (2.A.4) $$

$$ {}_{(0)}\langle J,M|e_1^*|M-1,J+1\rangle_{(1)} = \frac{1}{\sqrt{2}}\sqrt{\frac{J-M+1}{2J+1}}, \qquad (2.A.5) $$

$$ {}_{(0)}\langle J,M|e_0|M,J\rangle_{(1)} = -\frac{1}{\sqrt{2}}\sqrt{\frac{J-M}{2J+1}}, \qquad (2.A.6) $$

$$ {}_{(0)}\langle J,M|e_0|M,J+1\rangle_{(1)} = \frac{1}{\sqrt{2}}\sqrt{\frac{J+M+1}{2J+1}}. \qquad (2.A.7) $$

The matrix elements of e_0^2 are as follows

$$ {}_{(1)}\left\langle J,M|e_0^2|M,J-1\right\rangle_{(1)} = -\frac{\sqrt{(J-M-1)(J+M)}}{2(2J-1)}, \qquad (2.A.8) $$

$$ {}_{(1)}\left\langle J,M|e_0^2|M,J\right\rangle_{(1)} = \frac{1}{2}\left[1+\frac{2M+1}{4J^2-1}\right], \qquad (2.A.9) $$

$$ {}_{(1)}\left\langle J,M|e_0^2|M,J+1\right\rangle_{(1)} = -\frac{\sqrt{(J+M+1)(J-M)}}{2(2J+1)}, \qquad (2.A.10) $$

$$ {}_{(2)}\left\langle J,M|e_0^2|M,J\right\rangle_{(2)} = \frac{1}{2}\frac{M+1}{J(J+1)}. \qquad (2.A.11) $$

The matrix elements of the $e_0 e_1$ function are given as

$$
\begin{aligned}
{}_{(1)}\langle J, M-1|e_0 e_1|M, J-1\rangle_{(1)} &= \frac{\sqrt{(J-M)(J-M-1)}}{2(2J-1)},\\
{}_{(1)}\langle J, M-1|e_0 e_1|M, J\rangle_{(1)} &= \frac{\sqrt{(J-M)(J+M+1)}}{4J^2-1},\\
{}_{(1)}\langle J, M-1|e_0 e_1|M, J+1\rangle_{(1)} &= \frac{\sqrt{(J+M)(J+M+1)}}{2(-2J-1)},\\
{}_{(0)}\langle J, M|e_0 e_1^*|M-1, J-1\rangle_{(0)} &= \sqrt{\frac{(J+M-1)(J+M)}{4(2J-1)(2J+1)}},\\
{}_{(0)}\langle J, M+1|e_0 e_1^*|M, J+1\rangle_{(0)} &= -\sqrt{\frac{(J-M)(J-M+1)}{4(2J+1)(2J+3)}}.
\end{aligned}
\tag{2.A.12}
$$

Finally, we find the matrix elements of the e_0^4 function

$$
{}_{(2)}\langle J, M|e_0^4|M, J\rangle_{(2)} = \frac{(M+1)(3M-J(J+1)(M-3)) - J^2(J+1)^2}{2J(J+1)(2J-1)(2J+3)}. \tag{2.A.13}
$$

We have used the following results for the summation over M

$$
\sum_{M=-J}^{J-1} {}_{(1)}\langle J, M|e_0^2|M, J\rangle_{(1)} = \frac{2J}{2}, \tag{2.A.14}
$$

$$
\sum_{M=-J-1}^{J-1} {}_{(2)}\langle J, M|e_0^2|M, J\rangle_{(2)} = \frac{2J+1}{2}, \tag{2.A.15}
$$

$$
\sum_{M=-J-1}^{J-1} {}_{(2)}\langle J, M|e_0^4|M, J\rangle_{(2)} = \frac{2J+1}{3}, \tag{2.A.16}
$$

and

$$
\sum_{M=-J}^{J-1} {}_{(1)}\langle J, M|e_0^2|e_1|^2|M, J\rangle_{(1)} = \frac{J}{3}, \tag{2.A.17}
$$

$$
\sum_{M=-J-1}^{J-1} {}_{(2)}\langle J, M|e_0^2|e_1|^2|M, J\rangle_{(2)} = \frac{2J+1}{6}. \tag{2.A.18}
$$

2.B Renormalization around the trivial vacuum with the help of Pauli-Villars procedure

The σ_{xx} renormalization

In order to find the renormalization of the σ_{xx} conductivity we should compute the average in Eq. (2.2.41). Using the parameterization $Q = T_0^{-1} q T_0$ with the global unitary matrix $T_0 \in U(m) \times U(n)$ and expanding the q to the second order in w, we obtain

$$\sigma'_{xx} = \sigma_{xx} + \frac{\sigma_{xx}^2}{16mn} \int d\mathbf{r} \nabla^2 \sum_{\alpha,\gamma=1}^{m} \sum_{\beta,\delta=1}^{n} \Big[\langle v^{\alpha\beta}(\mathbf{r}) v^{\dagger\delta\alpha}(\mathbf{r}') \rangle_0 \langle v^{\gamma\delta}(\mathbf{r}') v^{\dagger\beta\gamma}(\mathbf{r}) \rangle_0$$
$$+ \langle v^{\alpha\beta}(\mathbf{r}) v^{\dagger\beta\gamma}(\mathbf{r}') \rangle_0 \langle v^{\gamma\delta}(\mathbf{r}') v^{\dagger\delta\alpha}(\mathbf{r}) \rangle_0 \Big], \tag{2.B.1}$$

where a point $\mathbf{r}'$ can be chosen arbitrary since the averages are depend only on the difference of the coordinates. Now by going from (x, y) to (η, θ) coordinates and performing the averages, we find

$$\sigma'_{xx} = \sigma_{xx} - 2\pi\beta_0 \int\limits_{\eta\theta} O^{(0)} \mathcal{G}_0(\eta\theta; \eta'\theta') \mathcal{G}_0(\eta'\theta'; \eta\theta). \tag{2.B.2}$$

Integrating over η, θ and introducing the Pauli-Villars masses as above, we leads to the following result

$$\sigma'_{xx} = \sigma_{xx} - 2\pi\beta_0 \left[\sum_{J=1} \frac{1}{E_J^{(0)}} + \sum_{f=1}^{K} \hat{e}_f \sum_{J=0} \frac{E_J^{(0)}}{(E_J^{(0)} + \mathcal{M}_f^2)^2} \right]$$
$$\times \sum_{M=-J}^{J} \Phi_{JM}^{(0)}(\eta', \theta') \bar{\Phi}_{JM}^{(0)}(\eta', \theta'). \tag{2.B.3}$$

It is worth mentioning that the Jacobi polynomial $P_{J-M}^{M,M}(\eta)$ is proportional to the Gegenbauer polynomial $C_{J-M}^{M+1/2}(\eta)$. By using the summation theorem [98]

$$C_J^{\lambda}(\cos\phi\cos\phi' + z\sin\phi\sin\phi') = \frac{\Gamma(2\lambda - 1)}{\Gamma^2(\lambda)} \sum_{M=0}^{J} \frac{2^{2M}\Gamma(J - M + 1)}{\Gamma(J + M + 2\lambda)} \Gamma^2(M + \lambda)$$
$$\times (2M + 2\lambda - 1) \sin^M\phi \sin^M\phi' C_{J-M}^{M+\lambda}(\cos\phi) C_{J-M}^{M+\lambda}(\cos\phi') C_M^{\lambda-1/2}(z) \tag{2.B.4}$$

with $z = 1$ and $\lambda = 1/2$, we find that the projection operator

$$\sum_{M=-J}^{J} \Phi_{JM}^{(0)}(\cos\phi, \theta) \bar{\Phi}_{JM}^{(0)}(\cos\phi', \theta) = \frac{2J + 1}{4\pi} C_J^{1/2} \left(\cos(\phi - \phi')\right). \tag{2.B.5}$$

Since $C_J^{1/2}(1) = 1$, we obtain

$$\sigma'_{xx} = \sigma_{xx} - \frac{\beta_0}{2} \lim_{\Lambda\to\infty} \left[\sum_{J=3/2}^{\Lambda} \frac{2J(J^2 - \frac{1}{4})}{(J^2 - \frac{1}{4})^2} + \sum_{f=1}^{K} \hat{e}_f \sum_{J=1/2}^{\Lambda} \frac{2J(J^2 - \frac{1}{4})}{(J^2 - \frac{1}{4} + \mathcal{M}_f^2)^2} \right]. \tag{2.B.6}$$

Finally, evaluation of the sums above yields

$$\sigma'_{xx} = \sigma_{xx} - \frac{\beta_0}{2} \left(Y_{\text{reg}}^{(0)} + 1 \right) = \sigma_{xx} \left(1 - \frac{\beta_0}{\sigma_{xx}} \ln \mathcal{M} e^{\gamma} \right). \tag{2.B.7}$$

The z_i renormalization

The renormalized quantities z_i' with $i = a, s, h$ are defined by Eqs (2.2.47). Using the parameterization $Q = T_0^{-1} q T_0$ with the global unitary matrix $T_0 \in U(m) \times U(n)$ and expanding q to the second order in w, we find

$$z_i' = z_i \left(1 + \frac{\gamma_i^{(0)} \pi}{2mn} \sum_{\alpha=1}^{m} \sum_{\beta=1}^{n} \langle v^{\alpha\beta} v^{\dagger\beta\alpha} \rangle_0 \right) . \tag{2.B.8}$$

Here the average $\langle \cdots \rangle_0$ is defined with the respect to the action $\delta S_\sigma^{(0)}$. The averages yield

$$z_i' = z_i \left(1 + \frac{2\pi\gamma_i^{(0)}}{\sigma_{xx}} \mathcal{G}_0(\eta\theta; \eta\theta) \right) . \tag{2.B.9}$$

By using Eq.(2.B.5), we find

$$z_i' = z_i \left(1 + \frac{\gamma_i^{(0)}}{2\sigma_{xx}} Y_{\text{reg}}^{(0)} \right) . \tag{2.B.10}$$

We finally obtain

$$z_i' = z_i \left(1 + \frac{\gamma_i^{(0)}}{\sigma_{xx}} \ln \mathcal{M} e^{\gamma - 1/2} \right) . \tag{2.B.11}$$

Chapter 3

Non-Fermi liquid theory for disordered metals near two dimensions

3.1 Introduction

The integral quantum Hall regime has traditionally been viewed as a (nearly) free particle localization problem with interactions playing only a minor role. [1] Although it is well known that many features of the experimental data, taken from low mobility heterostructures, [99] can be explained as the behavior of free particles, a much sharper formulation of the problem is obtained by considering the quantum Hall plateau transitions. [22] Following the experimental work by H.P Wei et al., [23] these transitions behave in all respects like a disorder driven metal-insulator transition that is characterized by two independent critical indices, i.e. a *localization length* exponent ν and a *phase breaking length* exponent p. [22] Whereas transport measurements usually provide an experimental value of only the ratio $\kappa = p/2\nu$, it is generally not known how the values of ν and p can be extracted separately.

In spite of the fact that one can not proceed without having a microscopic theory of electron-electron interaction effects, there is nevertheless a strong empirical believe in the literature [100, 101, 102, 103, 104, 105] which says that the zero temperature localization length exponent ν is given precisely by the free electron value $\nu = 2.3$ as obtained from numerical simulations. [77] The experimental situation has not been sufficiently well understood, [53, 63, 64, 65] however, to justify the bold assumption of Fermi liquid behavior. In fact, the progress that has been made over the last few years in the theory of localization and interaction effects clearly indicates that Fermi liquid principles do not exist in general. The Coulomb interaction problem lies in a different universality class of transport phenomena [31] with a previously unrecognized symmetry, called $\mathcal{F}$ invariance. [32, 33, 34, 35] The theory relies in many ways on the approach as initiated by Finkelstein [30] and adapted to the case of the spin polarized or spinless electrons. [31] By reconciling the Finkelstein theory with the topological concept of an instanton vacuum [19] and the Chern Simons statistical gauge fields, [68] the foundations have been laid for a complete renormalization theory that unifies the quantum theory of metals with that of the abelian quantum Hall states. [32, 33, 34, 35]

3.1.1 A historical problem

The unification of the integral and fractional quantum Hall regimes is based on the assumption that Finkelstein approach [30, 31] is renormalizable and generates a strong coupling, *insulating* phase with a massgap. However, the traditional analyses of the Finkelstein theory have actually not provided any guarantee that this is indeed so.

In spite of Finkelstein's pioneering and deep contributions to the field, it is well known that the conventional momentum shell renormalization schemes do not facilitate any computations of the quantum theory beyond one loop order. At the same time, application of the more advanced technique of dimensional regularization has led to conceptual difficulties with such aspects like *dynamical* scaling. [60] One can therefore not rule out the possibility that there are complications, either in the idea of renormalizability, or in other aspects of the theory such as the Matsubara frequency technique.

Nothing much has been clarified, however, by repeating similar kinds of analysis in a different formalism, like the Keldysh technique. [61, 62] What has been lacking all along is the understanding of a fundamental principle that has prevented the Finkelstein approach from becoming a fully fledged field theory for localization and interaction effects.

3.1.2 $\mathcal{F}$ invariance

In Ref. [32] it has been shown that the Finkelstein action has an exact symmetry ($\mathcal{F}$ invariance) that is intimately related to the electrodynamic $U(1)$ gauge invariance of the theory. $\mathcal{F}$ invariance is the basic mechanism that protects the renormalization of the problem with infinitely ranged interaction potentials such as the Coulomb potential. Moreover, it has turned out that the infrared behavior of physical observables can only be extracted from $\mathcal{F}$ invariant quantities and correlations, and these include the linear response to external potentials. Arbitrary renormalization group schemes break the $\mathcal{F}$ invariance of the action and this generally complicates the attempt to obtain the temperature and/or frequency dependence of physical quantities such as the conductivity and specific heat.

Quantum Hall physics is in many ways a unique laboratory for investigating and exploring the various different consequences of $\mathcal{F}$ invariance. For example, one of the longstanding questions in the field is whether and how the theory *dynamically* generates the *exact* quantization of the Hall conductance. As we mention in the previous chapter, the important progress has been made recently by demonstrating that the instanton vacuum, on the strong coupling side of the problem, generally displays massless excitations at the edge of the system. [46] These massless edge excitations are identical to those described by the more familiar theory of chiral edge bosons. [34] The theory of massless edge excitations implies that the concept of $\mathcal{F}$

invariance retains its fundamental significance all the way down to the regime of strong coupling.

3.1.3 Outline of this chapter

In this paper we put the concept of $\mathcal{F}$ invariance at work and evaluate the renormalization behavior of the Finkelstein theory at a two loop level. As shown in Ref. [33], the technique of dimensional regularization is a unique procedure, not only for the computation of critical indices, but also for extracting the dynamical scaling functions. In fact, the metal-insulator transition in $2+\epsilon$ spatial dimensions is the only place in the theory where the temperature and/or frequency dependence of physical observables can be obtained explicitly. This motivates us to further investigate the problem in $2+\epsilon$ dimensions and use it as a stage setting for the much more complex problem of the quantum Hall plateau transitions that we consider in the next chapter.

The final results of this chapter are remarkably similar to those of the more familiar classical Heisenberg ferromagnet. [92] For example, unlike the free electron gas, the Coulomb interaction problem displays a conventional phase transition (metal-insulator transition) in $2+\epsilon$ dimensions with an ordinary order parameter. The theory is therefore quite different from that of free electrons which has a different dimensionality and displays, as is well known, [96] anomalous or multifractal density fluctuations near criticality (see Section 2.8.2).

It is important to bear in mind, however, that the analogy with the Heisenberg model is rather formal and it fails on many other fronts. For example, the classification of critical operators is very different from what we used in the previous chapter for generalized CP^{N-1} models (ordinary sigma models). Moreover, the Feynman diagrams of the Finkelstein theory are more complex, involving internal frequency sums which indicate that the theory effectively exists in $2+1$ space-time dimensions, rather than in two spatial dimensions alone. The complexity of $\mathcal{F}$ invariant systems is furthermore illustrated by the lack of such principles like Griffith analyticity that facilitates a discussion of the symmetric phase in conventional sigma models. [107, 43] The dynamics of the strong coupling insulating phase of the electron gas is distinctly different from that of the Goldstone (metallic) phase and controlled by different operators in the theory. [108]

This chapter is organized as follows. After introducing the formalism (Section 3.2) we embark on the details of the two loop contributions to the conductivity in Section 3.4. We employ an $\mathcal{F}$-invariance-breaking parameter α to regularize the infinite sums over frequency. This methodology actually provides numerous self consistency checks and a major part of the computation consists of finding the ways in which the various singular contributions in α cancel each other. The actual computation of the diagrams is described in the Appendices which contain the list of the momentum and frequency integrals that are used in the main text of this chapter. In tables 3.1 and 3.2 we summarize how the different singular contributions in α cancel each other. Table 3.3 lists the various finite contributions to the pole term in ϵ. The final result for the β function is given by Eqs (3.4.30) - (3.4.34). In Section 3.5 we summarize the consequences for scaling and present the renormalization group flow diagram in $2+\epsilon$ dimensions in Section 3.6. We end this chapter with a conclusion (Section 3.7).

3.2 Effective parameters

3.2.1 The action

The generalized replica non-linear sigma model involves unitary matrix field variables $Q^{\alpha\beta}_{nm}(\mathbf{r})$ that obey the following constraints

$$Q = Q^\dagger, \quad \operatorname{tr} Q = 0, \quad Q^2 = 1. \tag{3.2.1}$$

The superscripts $\alpha, \beta = 1, \cdots, N_r$ represent the *replica* indices and the subscripts n, m are the indices of the *Matsubara* frequencies $\omega_k = \pi T(2k+1)$ with $k = n, m$. A convenient representation in terms of unitary matrices $\mathcal{T}(\mathbf{r})$ is obtained by writing

$$Q(\mathbf{r}) = \mathcal{T}^{-1}(\mathbf{r})\Lambda\mathcal{T}(\mathbf{r}), \quad \Lambda^{\alpha\beta}_{nm} = \operatorname{sign}(\omega_n)\delta^{\alpha\beta}\delta_{nm}. \tag{3.2.2}$$

The effective action for the two-dimensional interacting electron gas in the presence of disorder and an external vector potential $\mathbf{A}$ can be written as follows [32]

$$\mathcal{Z} = \int \mathcal{D}[Q] \exp S, \qquad S = S_\sigma + S_F + S_h. \tag{3.2.3}$$

Here, S_σ is the *free electron* action

$$S_\sigma[Q, \mathbf{A}] = -\frac{\sigma_{xx}}{8} \int d\mathbf{r} \operatorname{tr}[\mathbf{D}, Q]^2. \tag{3.2.4}$$

We remind that the quantity σ_{xx} represents the *meanfield* values for the *dissipative* conductance in units e^2/h respectively. The $\mathbf{D}$ is covariant derivative

$$\mathbf{D} = \nabla - i\widehat{\mathbf{A}}, \tag{3.2.5}$$

with

$$\widehat{\mathbf{A}} = \sum_{\alpha,n} \mathbf{A}^\alpha_n I^\alpha_n, \tag{3.2.6}$$

and $\mathbf{A}^\alpha_n$ is the Fourier transform of the homogeneous external vector potential $\mathbf{A}^\alpha(\tau)$ from imaginary time τ, $\mathbf{A}^\alpha(\tau) = \sum_n \mathbf{A}^\alpha(\nu_n)\exp(-i\nu_n\tau)$. Here $\nu_n = 2\pi T n$ is the Matsubara frequency with T being the temperature. Matrix

$$(I^\alpha_n)^{\beta\gamma}_{km} = \delta^{\alpha\beta}\delta^{\alpha\gamma}\delta_{k,n+m} \tag{3.2.7}$$

is the Matsubara representation of the $U(1)$ generator $\exp(-i\nu_n\tau)$.

Next, S_F contains the *singlet interaction* term [30, 32]

$$S_F[Q] = \pi T z \int d\mathbf{r}\, O_F[Q], \tag{3.2.8}$$

where

$$O_F[Q] = c \sum_{\alpha n} \operatorname{tr} I^\alpha_n Q \operatorname{tr} I^\alpha_{-n} Q + 4 \operatorname{tr} \eta Q - 6 \operatorname{tr} \eta\Lambda. \tag{3.2.9}$$

Here, z is the so-called *singlet interaction amplitude* and c the *crossover* parameter which allows the theory be interpolated between the case of electrons with Coulomb interaction ($c = 1$) and the free electron case ($c = 0$). It is worthwhile to mention that the quantity $\alpha = 1 - c$ breaks the $\mathcal{F}$ invariance of the theory and we shall eventually be interested below in the limit where α goes to zero. The singlet interaction term involves a matrix

$$\eta_{nm}^{\alpha\beta} = n\delta^{\alpha\beta}\delta_{nm} \tag{3.2.10}$$

that is used to represent the set of the Matsubara frequencies ω_n. Finally, the term

$$S_h[Q] = \frac{\sigma_{xx}h_0^2}{4}\int d\mathbf{r}\,\mathrm{tr}\,\Lambda Q. \tag{3.2.11}$$

is not a part of the theory but we shall use it later on as a convenient infrared regulator of the theory.

3.2.2 $\mathcal{F}$ invariance and $\mathcal{F}$ algebra

Unlike the free particle problem ($c = 0$) that we consider in a great details in the previous chapter, the theory with electron-electron interactions ($0 < c \leqslant 1$) is mainly complicated by the fact that the range of Matsubara frequency indices m, n must be taken from $-\infty$ to $+\infty$, along with the replica limit $N_r \to 0$. Under these circumstances one can show that the singlet interaction term fundamentally affects the ultra violet singularity structure of the theory (the renormalization group β and γ functions) which is one of the peculiar features of the theory of electron-electron interactions. [30, 33] Moreover, the problem with *infinite range interactions* ($c = 1$) such as the Coulomb interaction displays an exact global symmetry named $\mathcal{F}$ invariance [32]. This means that S_F is invariant under electrodynamic $U(1)$ gauge transformations which are spanned by the matrices I_n^α. This symmetry is broken by the problem with *finite range interactions* ($0 < c < 1$). In order to retain the $U(1)$ algebra in truncated frequency space with a cut-off N_m a set of algebraic rules has been developed named $\mathcal{F}$ algebra. [32] These rules permit one to proceed in finite frequency space where the index n runs from $-N_m$ to $N_m - 1$, i.e the matrix field variables Q have a finite size

$$Q(\mathbf{r}) = \mathcal{T}^{-1}(\mathbf{r})\Lambda\mathcal{T}(\mathbf{r}), \quad \mathcal{T}(\mathbf{r}) \in U(2N) \tag{3.2.12}$$

where $N = N_r N_m$. The two limits of the theory, $N_r \to 0$ and then $N_m \to \infty$ respectively, are taken at the end of all computations. The main purpose of $\mathcal{F}$ algebra is to ensure that electrodynamic $U(1)$ gauge invariance as well as $\mathcal{F}$ invariance are preserved by the renormalization group, both perturbatively and at a non-perturbative level.

3.2.3 Physical observables

Next, for a detailed understanding of interaction effects it is clearly necessary to develop a quantum theory for the *observable* parameters σ'_{xx}, z' and c'. At the same time it is extremely important to show that the response quantities defined by the background field procedure are precisely the same as those obtained from ordinary linear response theory. This will be done further where we embark on some of the principal results of $\mathcal{F}$ algebra.

In this Section we recollect the $\mathcal{F}$ invariant expressions for the observable parameters that will be used in the remainder of this chapter. As pointed out in Ref. [33] the main advantage of working with $\mathcal{F}$ invariant quantities is that they facilitate renormalization group computations at finite temperatures and frequencies. They are furthermore valid in the entire range $0 \leqslant c \leqslant 1$ and simpler to work with in general. In the second part of this Section we briefly recall the known results of the theory in $2+\epsilon$ spatial dimensions.

Kubo formula

The response quantity σ'_{xx} for arbitrary values of c can be expressed in terms of current-current correlations according to [21, 33]

$$\sigma'(n) = \langle O_1 \rangle + \langle O_2 \rangle, \tag{3.2.13}$$

where

$$O_1 = -\frac{\sigma_{xx}}{4n} \mathrm{tr}[I_n^\alpha, Q(\mathbf{r})][I_{-n}^\alpha, Q(\mathbf{r})] \tag{3.2.14}$$

and

$$O_2 = \frac{\sigma_{xx}^2}{4nD} \int d\mathbf{r} \mathrm{tr} I_n^\alpha Q(\mathbf{r}) \nabla Q(\mathbf{r}) \mathrm{tr} I_{-n}^\alpha Q(\mathbf{r}') \nabla Q(\mathbf{r}'), \tag{3.2.15}$$

with $D = 2+\epsilon$. Here and from now onward the expectations are defined with the respect to the theory of Eqs (3.2.3)- (3.2.9) and we assume the *spherical boundary conditions* (see previous chapter).

Specific heat

A natural definition of the observable quantity z' is obtained through the derivative of the thermodynamic potential with respect to temperature, i.e.

$$\frac{\partial \ln Z}{\partial \ln T} = \pi T z' \int d\mathbf{r}\, O_F[\Lambda] \tag{3.2.16}$$

which is directly related to the specific heat of the electron gas. [33] Equivalently we can write

$$z' = z \frac{\langle O_F[Q] \rangle}{O_F[\Lambda]}. \tag{3.2.17}$$

Finally, the observable quantity c' may be obtained from the general conditions that are imposed on *static response* of the system. These general conditions imply that quantity $z\alpha = z(1-c)$ remains unaffected by the quantum fluctuations. [30, 32, 33] The c' is therefore determined by the relation

$$z'(1-c') = z(1-c) \qquad \text{or} \qquad z'\alpha' = z\alpha. \tag{3.2.18}$$

β and γ functions

Before embarking of the computation of the observable σ'_{xx} in the two-loop approximation we shall first recapitulate the known results of the perturbative renormalization group for disordered spinless metals in $2+\epsilon$ dimensions. [30, 33] Let μ' denote the momentum scale associated with the observable theory then the quantities $\sigma'_{xx} = \sigma_{xx}(\mu')$, $z' = z(\mu')$ and $c' = c(\mu')$ can be expressed in terms of the renormalization group β and γ functions according to (see Section 2.2.4)

$$\sigma'_{xx} = \sigma_{xx} + \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu} \beta_\sigma(\sigma_{xx}, c), \tag{3.2.19}$$

$$z' = z - \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu} \gamma_z(\sigma_{xx}, c) z, \tag{3.2.20}$$

$$z'\alpha' = z\alpha, \tag{3.2.21}$$

where

$$\beta_\sigma(\sigma_{xx}, c) = -\epsilon\sigma_{xx} + \beta_0(c) + \frac{\beta_1(c)}{\sigma_{xx}} + O(\sigma_{xx}^{-2}), \tag{3.2.22}$$

$$\gamma_z(\sigma_{xx}, c) = \frac{c\gamma_0}{\sigma_{xx}} + \frac{c\gamma_1(c)}{\sigma_{xx}^2} + O(\sigma_{xx}^{-3}). \tag{3.2.23}$$

The one-loop results are known for arbitrary value of the crossover parameter c and are given by [30, 33]

$$\beta_0(c) = \frac{2}{\pi}\left(1 + \frac{\alpha}{c}\ln\alpha\right), \qquad \gamma_0 = -\frac{1}{\pi}, \tag{3.2.24}$$

whereas the two-loop results for β_σ were obtained for $c = 0$ only and for γ_z for $c = 0$ and $c = 1$. In the case of electrons with the Coulomb interaction ($c = 1$) the results are as follows [33]

$$\gamma_1(1) = -\frac{3}{\pi^2} - \frac{1}{6} \approx -0.47, \tag{3.2.25}$$

For the case of free electrons ($c = 0$) the two-loop results is known [92]

$$\beta_1(0) = \frac{1}{2\pi^2}, \tag{3.2.26}$$

$$\gamma_1(0) = 0. \tag{3.2.27}$$

3.2.4 The h_0 field

Although we are interested, strictly speaking, in evaluating $\sigma'(n)$ with varying values of external frequencies ν_n and temperature, the computation simplifies dramatically if we put these parameters equal to zero in the end and work with a finite value of the h_0 field instead. This procedure has been analyzed in exhaustive detail in our previous work and, in what follows, we shall greatly benefit from the technical advantages that make the two-loop analysis of the conductivity possible. We shall return to finite frequency and temperature problem in the end of this paper (Section 3.5).

The infrared regularization by the h_0 field relies on the following statement

$$\sigma_{xx} h_0^2 \langle Q(\mathbf{r})\rangle = \sigma'_{xx} h'^2 \Lambda, \tag{3.2.28}$$

which says that there is an effective mass h' in the problem that is being induced by the presence of the h_0 field. It is very well known that, since the quantity $\langle Q(\mathbf{r})\rangle$ is not a gauge invariant object, the definition of the h' field is singular as α goes to zero and the theory is generally not renormalizable. However, the effective parameter σ' is truly defined in terms of the effective mass h' rather than the bare parameter h_0. Hence, all the non-renormalizable singularities are removed from the theory, provided we express σ' in terms of the h' rather than the h_0 (see Section 2.2.4). We shall show that the ultraviolet singularities of the theory can be extracted directly from the final result for $\sigma'(h')$. On the other hand, we can make use of our previous results [33] and express the final answer in terms of frequencies and temperature, rather than the mass h'.

3.3 Linear response versus background field procedure

With the introduction of $\mathcal{F}$ *algebra* it has become possible to show that observable quantities σ'_{xx}, σ'_{xy}, z' and c' which are usually obtained by means of *background field* procedures or *momentum shell* procedures are, in fact, precisely the same as the expressions for the conductances at zero temperature that one derives from ordinary linear response theory in the external vector potential (Eq. (3.2.13)). [32] For the purpose of next chapter we also extend the discussion to the conductivity σ'_{xy}. By the same procedure we prove results (3.2.17) and (3.2.18). In this section we present the arguments for the special case where the infrared of the system is regulated by a finite size L rather than by the infrared regulator h_0^2.

3.3.1 Linear response theory

The response of the system to an external vector potential $\mathbf{A}$ can generally be written in terms of an effective action $S_{\text{eff}}[\mathbf{A}]$ according to

$$\exp S_{\text{eff}}[\mathbf{A}] = \int \mathcal{D}[Q] \exp\Big(S_\sigma[Q,\mathbf{A}] + S_F[Q]\Big). \tag{3.3.1}$$

In the presence of the Hall conductivity σ_{xy} the free electron part (3.2.4) becomes [32]

$$S_\sigma[Q,\mathbf{A}] = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\, \text{tr}[\mathbf{D},Q]^2 + \frac{\sigma_{xy}}{8}\int d\mathbf{r}\, \text{tr}\, \varepsilon_{jk} Q[D_j,Q][D_k,Q]. \tag{3.3.2}$$

Since we are interested in the global response at zero temperature and frequency it suffices to take a spatially independent $\mathbf{A}^\alpha(\nu_n)$ and consider a small range of values $\nu_n = 2\pi T n \approx 0$ only. The response parameters σ'_{xx} and σ'_{xy} are then defined by the following general form of the effective action

$$S_{\text{eff}}[\mathbf{A}] = -\int d\mathbf{r} \sum_{\alpha,n>0} n\Big[\sigma'_{xx}\delta_{jk} + \sigma'_{xy}\varepsilon_{jk}\Big] A_j(\nu_n) A_k(-\nu_n). \tag{3.3.3}$$

By using this expression for the left hand side of Eq. (3.3.1) it is easy to derive result (3.2.13) for σ'_{xx} and the following result for σ'_{xy}

$$\sigma'_{xy} = \sigma_{xy} + \frac{\sigma_{xx}^2}{4nD} \int d\mathbf{r}\, \varepsilon_{ab} \langle \operatorname{tr} I_n^\alpha Q(\mathbf{r}) \nabla_a Q(\mathbf{r}) \operatorname{tr} I_{-n}^\alpha Q(\mathbf{r}') \nabla_b Q(\mathbf{r}') \rangle. \tag{3.3.4}$$

Eqs. (3.2.13) and (3.3.4) are some of the most fundamental quantities of the theory since they can generally be used for studies at finite temperature and frequency rather than finite sample sizes. Moreover, they facilitate an analysis of mesoscopic fluctuations as well as important self-consistency checks in practical computations such as the replica limit $N_r = 0$ and $N_m \to \infty$.

However, the complications primarily arise if one wants to make sure that the Finkelstein formalism preserves the fundamental symmetries of the interacting electron gas, in particular the electrodynamic $U(1)$ gauge invariance as well as $\mathcal{F}$ invariance which are properly defined in infinite Matsubara frequency space only. As we shall see next, these complications automatically arise in the attempt to lay the bridge between linear response theory and the effective action for the edge modes.

3.3.2 $\mathcal{F}$ invariance

To deal with electrodynamic gauge invariance in *finite* frequency space we start out by embedding the matrix variables Q of size $2N_rN_m \times 2N_rN_m$ in a much larger matrix space of size $2N_rN'_m \times 2N_rN'_m$ with $1 \ll N_m \ll N'_m$. All matrix manipulations will be carried out from now onward in the space of *large* matrices whereas the unitary rotations Q effectively retain their size $2N_rN_m \times 2N_rN_m$ which we term *small*.

Let us next introduce the quantity $\varphi_n^\alpha(\mathbf{r}) = \mathbf{A}^\alpha(\nu_n) \cdot \mathbf{r}$. We can then express the vector potential $\hat{\mathbf{A}}$ in terms of the *large* unitary matrix $\hat{\varphi} = \hat{\varphi}(\mathbf{r})$ according to

$$\hat{\mathbf{A}} = \nabla \hat{\varphi} = i W^{-1} \nabla W, \qquad W = \exp(-i\hat{\varphi}). \tag{3.3.5}$$

Following the rules of $\mathcal{F}$ *algebra* [32] the unitary matrix W just stands for an electrodynamic $U(1)$ gauge transformation in Matsubara frequency notation. The free electron part of the action (3.3.2) can be expressed in terms of the W rotation on the matrix field variable Q according to

$$\begin{aligned} S_\sigma[Q, \mathbf{A}] &= S_\sigma[W^{-1}QW] = -\frac{\sigma_{xx}}{8} \int d\mathbf{r} \operatorname{tr}[\nabla(W^{-1}QW)]^2 \\ &+ \frac{\sigma_{xy}}{8} \int d\mathbf{r} \operatorname{tr} \varepsilon_{jk} Q \nabla_j (W^{-1}QW) \nabla_k (W^{-1}QW). \end{aligned} \tag{3.3.6}$$

Next we split the quantity $O_F[Q]$ into an $\mathcal{F}$ invariant part $O_s[Q]$ and a symmetry breaking part

$$O_F[Q] = O_s[Q] + O_\eta[Q], \tag{3.3.7}$$

where [32]

$$O_s[Q] = zc\Big(\sum_{\alpha n} \mathrm{tr}\, I_n^\alpha Q\, \mathrm{tr}\, I_{-n}^\alpha Q + 4\,\mathrm{tr}\,\eta Q - 6\,\mathrm{tr}\,\eta\Lambda\Big)$$
$$= zc{\sum_{\alpha n}}' \mathrm{tr}[I_n^\alpha, Q][I_{-n}^\alpha, Q] \qquad (3.3.8)$$
$$O_\eta[Q] = z\alpha\left\{4\,\mathrm{tr}\,\eta Q - 6\,\mathrm{tr}\,\eta\Lambda\right\}. \qquad (3.3.9)$$

The statement of $\mathcal{F}$ *invariance* now says that $O_s[Q]$ is gauge invariant [32]

$$O_s[Q] = O_s[W^{-1}QW]. \qquad (3.3.10)$$

On the other hand, as long as one evaluates the theory at zero temperature and finite system sizes, the response parameters σ'_{xx} and σ'_{xy} remain unchanged if one inserts the W rotation into the quantity $O_\eta[Q]$, i.e. the replacement

$$O_\eta[Q] \to O_\eta[W^{-1}QW] \qquad (3.3.11)$$

does not affect the statement of Eq. (3.3.3) where the σ'_{xx} and σ'_{xy} depend on the system size L. Linear response theory at zero temperature and finite system sizes is therefore formally the same thing as evaluating the theory in the presence of a gauge field W

$$\exp \tilde{S}_{\mathrm{eff}}[\mathbf{A}] = \int \mathcal{D}[Q] \exp\left(S_\sigma[W^{-1}QW] + S_F[W^{-1}QW]\right). \qquad (3.3.12)$$

The main reason for introducing the two different cut-offs $1 \ll N_m \ll N'_m$ in *finite* Matsubara frequency space is to ensure that Eqs. (3.3.6), (3.3.11) and (3.3.12) display the *exact* same symmetries that are known to exist in the theory where N_m and N'_m are being sent off to infinity.

3.3.3 Background field formalism

It is clear the the statement of Eq. (3.3.12) is non-trivial only due to the fact that that we work at zero temperature and with fixed boundary conditions on the matrix field variable Q. If on the other hand we were to work with *finite* temperatures and *infinite* system sizes L then Eq. (3.3.12) is merely a statement of electrodynamic $U(1)$ gauge invariance which is clearly very different from Eq. (3.3.1).

Notice that Eq. (3.3.12) is not yet quite the same as the *back ground field* methodology that previously has been studied intensively for renormalization group purposes. This is because the quantities Q and $W^{-1}QW$ by construction belong to different manifolds for any finite value of N_m and N'_m. However, in order for the W rotation in Eqs. (3.3.10), (3.3.11) and (3.3.12) to represent an *exact* electrodynamic $U(1)$ gauge transformation it is imperative that the results do not fundamentally depend on the details of how the frequency cut-offs N_m and N'_m go to infinity. Moreover, the statement of Eq. (3.3.12) renders highly non-trivial if one recognizes that the unitary matrix W can in general be written as the product of two distinctly different matrices t and U_0

$$W = \exp(-i\hat{\varphi}) = U_0\, t, \qquad U_0 \in U(N') \times U(N'), \qquad (3.3.13)$$

where $N' = N_r N'_m$. Here, t is a "small" background matrix field in the true sense of the word

$$t = \exp\left(\frac{i}{2}[\hat{\varphi}, \Lambda]\Lambda \ + \ \ldots\right) \tag{3.3.14}$$

whereas the "large" generators of W are all collected together in the $U(N') \times U(N')$ gauge U_0 which can be written as

$$U_0 = \exp\left(\frac{i}{2}\{\hat{\varphi}, \Lambda\}\Lambda\right). \tag{3.3.15}$$

Next we consider the change of variables

$$U_0^{-1} Q U_0 \to Q. \tag{3.3.16}$$

It is clear that this transformation preserves the spherical boundary conditions and leaves the measure of the functional integral unchanged. Equation (3.3.12) can therefore be represented as follows

$$\exp \tilde{S}_{\text{eff}}[\mathbf{A}] = \int \mathcal{D}[Q] \exp\Big(S_\sigma[t^{-1}Qt] + S_F[t^{-1}Qt]\Big) \tag{3.3.17}$$

which precisely corresponds to the background field methodology with the "small" matrix field t given explicitly by Eq. (3.3.14). This, then, leads to the principle result of this Section which says that Eq. (3.3.17) in the limit where $N_m, N'_m \to \infty$ and $T = 0$ is identically the same as linear response theory Eqs. (3.2.13) and (3.3.4).

Eq. (3.3.14) together with Eq. (3.3.17) can be used to derive different or alternative expressions for the quantities σ'_{xx} and σ'_{xy} which are completely equivalent to those given by Eqs. (3.2.13) and (3.3.4). Here we do not list these expressions but instead we simply verify the correctness of the effective action of Eq. (3.3.3). Since Eq. (3.3.17) has the same form as the effective action for the edge modes we can immediately write down the following general result

$$\tilde{S}^0_{\text{eff}}[\mathbf{A}] = -\frac{\sigma'_{xx}}{8} \int d\mathbf{r}\, \text{tr}(\nabla q)^2 + \frac{\sigma'_{xy}}{8} \int d\mathbf{r}\, \text{tr}\, \varepsilon_{jk} q \nabla_j q \nabla_k q, \tag{3.3.18}$$

where the superscript "0" denotes the result at $T = 0$. Eq. (3.3.18) can be obtained, as before, by expanding in the gradients of the slowly varying matrix field $q = t^{-1}\Lambda t$. By inserting expression (3.3.14) for t in Eq. (3.3.18) we obtain

$$\tilde{S}^0_{\text{eff}}[\mathbf{A}] = -\int d\mathbf{r} \sum_{\alpha, n > 0} n\Big[\sigma'_{xx}\delta_{jk} + \sigma'_{xy}\varepsilon_{jk}\Big] \nabla_j \varphi^\alpha_n \nabla_k \varphi^\alpha_{-n}. \tag{3.3.19}$$

The following identities have been used

$$\text{tr}[\hat{I}^\alpha_n, \Lambda][\hat{I}^\alpha_{-n}, \Lambda] \ = \ -4n, \tag{3.3.20}$$

$$\text{tr}\, \Lambda[\hat{I}^\alpha_n, \hat{I}^\alpha_{-n}] \ = \ 2n. \tag{3.3.21}$$

We see that we recover the same results as those in Eq. (3.3.3).

3.3.4 The quantities z' and c'

For completeness we next extend the results of the background field methodology to include the terms obtained by expanding to lowest order in T

$$\begin{aligned}\exp \tilde{S}_{\text{eff}}[\mathbf{A}] &= \exp\left(\tilde{S}^0_{\text{eff}}[\mathbf{A}]\right)\Big(1 + Tz'c' \int d\mathbf{r} \sum_{\alpha n}{}' \operatorname{tr}[I^\alpha_n, q][I^\alpha_{-n}, q] \\ &\qquad + Tz\alpha \int d\mathbf{r}\, (4 \operatorname{tr} \eta q - 6 \operatorname{tr} \eta\Lambda)\Big). \qquad (3.3.22)\end{aligned}$$

These results indicate that the quantity zc is renormalized whereas the statement $z\alpha = z'\alpha'$ is a physical constraint that should in general be imposed upon the theory. Eq. (3.3.22) has been verified in the theory of perturbative expansions. In next chapter we explicitly check the validity of this statement at a non-perturbative level. As a final remark, it should be mentioned that by taking $q = \Lambda$ in Eq. (3.3.22) one immediately obtains the expression for z', Eq. (3.2.17).

3.4 Computation of conductivity in $D = 2 + \epsilon$ dimensions

3.4.1 Introduction

To define a theory for perturbative expansions we use the following parameterization

$$Q = \begin{pmatrix} \sqrt{1 - qq^\dagger} & q^\dagger \\ q & -\sqrt{1 - q^\dagger q} \end{pmatrix}. \qquad (3.4.1)$$

The action can be written as an infinite series in the independent fields $q^{\alpha\beta}_{n_1 n_2}$ and $[q^\dagger]^{\alpha\beta}_{n_4 n_3}$. We use the convention that Matsibara indices with odd subscripts: $n_1, n_3, ...$, run over non-negative integers, whereas those with even subscripts: $n_2, n_4, ...$, run over negative integers. The parameterization (3.4.1) introduces the nontrivial Jacobian for integration over q and $q^\dagger$ fields. Fortunately, it does not contribute in the ϵ-expansion procedure. [42]

The propagators can be written in the form [60, 33]

$$\langle q^{\alpha\beta}_{n_1 n_2}(p)[q^\dagger]^{\gamma\delta}_{n_4 n_3}(-p)\rangle = \frac{4}{\sigma_{xx}} \delta^{\alpha\delta}\delta^{\beta\gamma}\delta_{n_{12}, n_{34}} D_p(n_{12}) \left(\delta_{n_1 n_3} + \delta^{\alpha\beta}\kappa^2 zc D^c_p(n_{12})\right), \qquad (3.4.2)$$

where

$$\begin{aligned}[D_p(n_{12})]^{-1} &= p^2 + h_0^2 + \kappa^2 z n_{12}, \qquad (3.4.3)\\ [D^c_p(n_{12})]^{-1} &= p^2 + h_0^2 + \kappa^2 \alpha z n_{12}. \qquad (3.4.4)\end{aligned}$$

Here we use the notation $n_{12} = n_1 - n_2$ and $\kappa^2 = 8\pi T/\sigma_{xx}$.

The expression for the DC conductivity is known to one loop order [33]

$$\sigma'^{\text{one}}_{xx} = \sigma_{xx} + \frac{8\Omega_D h_0^\epsilon}{\epsilon}, \qquad \Omega_D = \frac{S_D}{2(2\pi)^D}, \qquad (3.4.5)$$

where $S_D = 2\pi^{D/2}/\Gamma(D/2)$ is the surface of a D dimensional unit sphere.

3.4.2 The two-loop theory

To proceed we need the following terms obtained by expanding the action (2.1) in terms of q and $q^\dagger$ fields:

$$S_{\rm int}^{(3)} = -\pi Tzc\int d\mathbf{r}\sum_{\beta,m>0}\left\{{\rm tr}I_m^\beta q^\dagger\,{\rm tr}I_{-m}^\beta[q,q^\dagger]+{\rm tr}I_{-m}^\beta q\,{\rm tr}I_m^\beta[q,q^\dagger]\right\}, \quad (3.4.6)$$

$$S_{\rm int}^{(4)} = \frac{1}{4}\pi Tzc\int d\mathbf{r}\left\{2\sum_{\beta,m>0}{\rm tr}I_{-m}^\beta[q,q^\dagger]\,{\rm tr}I_m^\beta[q,q^\dagger]+\sum_\beta({\rm tr}I_0^\beta[q,q^\dagger])^2\right\}, \quad (3.4.7)$$

$$\begin{aligned} S_0^{(4)} &= \frac{\sigma_{xx}}{32}\int_{p_i}\delta\left(\sum_{i=1}^4\mathbf{p}_i\right)\sum_{n_1n_2n_3n_4}^{\beta\gamma\delta\mu} q_{n_1n_2}^{\beta\gamma}(p_1)(q^\dagger)_{n_2n_3}^{\gamma\delta}(p_2)q_{n_3n_4}^{\delta\mu}(p_3)(q^\dagger)_{n_4n_1}^{\mu\beta}(p_4)\\ &\times\left\{(\mathbf{p}_1+\mathbf{p}_2)\cdot(\mathbf{p}_3+\mathbf{p}_4)+(\mathbf{p}_2+\mathbf{p}_3)\cdot(\mathbf{p}_1+\mathbf{p}_4)-\kappa^2z_0(n_{12}+n_{34})-2h_0^2\right\}, \end{aligned} \quad (3.4.8)$$

where the superscripts is equal to the number of q or $q^\dagger$ fields and

$$\int_p \equiv \int\frac{d^D\mathbf{p}}{(2\pi)^D}. \quad (3.4.9)$$

In addition, we need the following terms obtained by expanding the expression for the conductivity, Eq. (3.2.13),

$$O_1^{(2)} = -\frac{\sigma_{xx}}{2}{\rm tr}\left\{I_n^\alpha q^\dagger I_{-n}^\alpha q+I_{-n}^\alpha q^\dagger I_n^\alpha q-2(I_n^\alpha\Lambda I_{-n}^\alpha+I_{-n}^\alpha\Lambda I_n^\alpha)[q,q^\dagger]\right\}, \quad (3.4.10)$$

$$O_1^{(3)} = \frac{\sigma_{xx}}{4}{\rm tr}\left\{I_n^\alpha(q+q^\dagger)I_{-n}^\alpha qq^\dagger-I_{-n}^\alpha(q+q^\dagger)I_n^\alpha q^\dagger q\right\}, \quad (3.4.11)$$

$$O_1^{(4)} = \frac{\sigma_{xx}}{16}{\rm tr}\left\{(I_n^\alpha\Lambda I_{-n}^\alpha+I_{-n}^\alpha\Lambda I_n^\alpha)[qq^\dagger q,q^\dagger]-2I_n^\alpha[q,q^\dagger]I_{-n}^\alpha[q,q^\dagger]\right\}, \quad (3.4.12)$$

$$O_2^{(4)} = \frac{\sigma_{xx}^2}{4D}\int d\mathbf{r}{\rm tr}I_n^\alpha(q\nabla q^\dagger+q^\dagger\nabla q){\rm tr}I_{-n}^\alpha(q\nabla q^\dagger+q^\dagger\nabla q), \quad (3.4.13)$$

$$\begin{aligned} O_2^{(5)} &= \frac{\sigma_{xx}^2}{8D}\int d\mathbf{r}\Big\{{\rm tr}I_n^\alpha(q\nabla q^\dagger+q^\dagger\nabla q){\rm tr}I_{-n}^\alpha q(\nabla q^\dagger)q\\ &\quad+{\rm tr}I_{-n}^\alpha(q\nabla q^\dagger+q^\dagger\nabla q){\rm tr}I_n^\alpha q^\dagger(\nabla q)q^\dagger\Big\}, \end{aligned} \quad (3.4.14)$$

$$\begin{aligned} O_2^{(6)} &= \frac{\sigma_{xx}^2}{16D}\int d\mathbf{r}\Big\{{\rm tr}I_n^\alpha\Lambda q^\dagger(\nabla q)q^\dagger){\rm tr}I_{-n}^\alpha\Lambda q(\nabla q^\dagger)q)\\ &\quad+{\rm tr}I_n^\alpha(q\nabla q^\dagger+q^\dagger\nabla q){\rm tr}I_{-n}^\alpha(qq^\dagger\nabla(qq^\dagger)+q^\dagger q\nabla(q^\dagger q))\\ &\quad+{\rm tr}I_{-n}^\alpha(q\nabla q^\dagger+q^\dagger\nabla q){\rm tr}I_n^\alpha(qq^\dagger\nabla(qq^\dagger)+q^\dagger q\nabla(q^\dagger q))\Big\}. \end{aligned} \quad (3.4.15)$$

Next we give the complete list of two loop contributions to the conductivity as follows

$$\begin{aligned} \sigma_{xx}'^{\rm two}(n) &= \Big\langle O_1^{(4)}+O_1^{(3)}S_{\rm int}^{(3)}+O_1^{(2)}(S_{\rm int}^{(4)}+S_0^{(4)}+\frac{1}{2}(S_{\rm int}^{(3)})^2)\\ &+\; O_2^{(6)}+O_2^{(5)}S_{\rm int}^{(3)}+O_2^{(4)}(S_{\rm int}^{(4)}+S_0^{(4)}+\frac{1}{2}(S_{\rm int}^{(3)})^2)\Big\rangle. \end{aligned} \quad (3.4.16)$$

The computations of the terms in Eq. (3.4.16) are straightforward but lengthy and tedious. In what follows we present the expressions in terms of the momentum integrals, frequency sums and propagators D_p, D_p^c for each term in Eq. (3.4.16) separately, along with the final answer. In the Appendix 3.A we give the complete list of integrals and symbols that we shall make use of here.

3.4.3 Computation of contractions in Eq. (3.4.16)

The first contraction is given as

$$\langle O_1^{(4)}\rangle = \frac{2}{\sigma_{xx}}\left(\int_p D_p(0)\right)^2 + \frac{2a^2}{\sigma_{xx}}\left(\sum_{m>0}\int_p DD_q^c(m)\right)^2 = 8\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2}(1+\ln^2\alpha) \quad (3.4.17)$$

with $DD_q^c(m) \equiv D_q(m)D_q^c(m)$. Next,

$$\begin{aligned}\langle O_1^{(3)}S_{\text{int}}^{(3)}\rangle \;=\; & -\frac{8}{\sigma_{xx}}\kappa^2 zc\int_{p,q}\sum_{k>0}\Big[D_{p+q}^c(0)D_q(k)D_p(k) + \kappa^2 zc\sum_{m>0}D_p^c(m)DD_q^c(k) \\ & \times D_{p+q}(k+m)\Big] = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon}\left(4S_0 + 4A_{00}^0\right) = 16\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2}\Big[-(1+\ln^2\alpha) \\ & +\epsilon(1+\zeta(3)/2)\Big], \end{aligned} \quad (3.4.18)$$

where $\zeta(x)$ is the Riemann zeta-function. The third contraction yields

$$\begin{aligned}\left\langle O_1^{(2)}(S_{\text{int}}^{(4)} + S_0^{(4)} + \frac{1}{2}(S_{\text{int}}^{(3)})^2)\right\rangle \;=\; & \frac{4}{\sigma_{xx}}\kappa^2 zc\int_{p,q}\sum_{k>0}\Big[D_{p+q}^c(0)D_q(k)D_p(k) \\ & +\kappa^2 zc\sum_{m>0}D_p^c(m)D_q^c(k)D_{p+q}^2(k+m) + \kappa^2 zc\sum_{m>0}(1+\kappa^2 zcmD_p^c(m)) \\ & \times DD_q^c(k)D_{p+q}^2(k+m)\Big] = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon}\left(-2S_0 - 2D_1 - 2T_{01} - 2A_{1,0}^0\right) \\ & = 4\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2}\left[2(1+\ln^2\alpha) - \epsilon\left(2+\zeta(3)+\pi^2/3\right)\right]. \end{aligned} \quad (3.4.19)$$

The next one,

$$
\begin{aligned}
\langle O_2^{(6)}\rangle &= \frac{(-4)}{\sigma_{xx}D}\int_{p,q} p^2\Bigg\{D_p(0)D_q(0)D_{p+q}(0) - 4(\kappa^2 zc)^2\sum_{k,m>0} D^2D_p^c(m)\hat{S}_m DD_q^c(k) \\
&\quad -(\kappa^2 zc)^2\sum_{k,m>0}\Big[D_p(k+m)DD_q^c(m)DD_{p+q}^c(k) + 2DD_p^c(k+m)DD_q^c(m) \\
&\quad \times D_{p+q}(k)\Big]\Bigg\} = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon}\left[S_1 + 4\left(\frac{2\ln\alpha}{\epsilon} + B_1\right) + C_{01} + 2C_{00}\right] \\
&\quad = 4\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2}\Big[16\ln\alpha - 2 + \frac{\epsilon}{2}\Big(-4\ln\alpha - \frac{\pi^2}{3} + \frac{\pi^2}{2}\ln 2 + \frac{\pi^4}{12} + \frac{11\zeta(3)}{2} \\
&\quad + \frac{\pi^2}{3}\ln^2 2 - \frac{1}{3}\ln^4 2 - 7\zeta(3)\ln 2 - 8\,\mathrm{li}_4\Big(\frac{1}{2}\Big)\Big)\Big]. \qquad (3.4.20)
\end{aligned}
$$

Here $D^n D_q^c(m) \equiv D_q^n(m)D_q^c(m)$ and $\mathrm{li}_n(x) = \sum_{k=1}^{\infty} x^k/k^n$ is the polylogarithmic function ($\mathrm{li}_4(1/2) \approx 0.517$), and we have introduced an operator $\hat{S}_m$ which acts only on frequency k according to the following rule $\hat{S}_m f(k) = f(k) + f(k+m)$. The fifth contraction is as follows

$$
\begin{aligned}
\langle O_2^{(5)} S_{\mathrm{int}}^{(3)}\rangle &= \frac{16}{\sigma_{xx}D}\kappa^2 zc\int_{p,q}\Bigg\{\mathbf{p}\cdot(\mathbf{p}-\mathbf{q})\sum_{k>0} D_{p+q}^c(0)D_p^2(k)D_q(k) + \kappa^2 zcp^2 \\
&\quad \times\sum_{k,m>0} D_{p+q}^c(m)\Big[D_p^2(k+m)DD_q^c(k) + D^2D_p^c(k+m) \\
&\quad \times D_q(k)\Big] - \kappa^2 zc\,(\mathbf{p}\cdot\mathbf{q})\sum_{k,m>0}\Big[DD_p^c(m)\hat{T}_m DD_{p+q}^c(k)D_q(k+m) \\
&\quad + D_{p+q}^c(m)D^2D_p^c(k+m)D_q(k+2m)\Big]\Bigg\} = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon}\Big(-4S_{00} - 4A_{01}^1 \\
&\quad -4H_0 - 4C_0 - 4A_0\Big) = 4\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2}\Big[-8\ln\alpha + 4 + \frac{\epsilon}{2}\Big(4\ln^2\alpha + 20\ln\alpha \\
&\quad -12 + 4\zeta(3) + \frac{4\pi^2}{3} - 4A_0 + 4C_0^{'}\Big)\Big], \qquad (3.4.21)
\end{aligned}
$$

where we have introduced yet another operator $\hat{T}_m$ which acts only on frequency k but now according to the rule $\hat{T}_m f(k) = f(k) - f(k+m)$. The result for next contraction can be written as

$$\begin{aligned}
\langle O_2^{(4)} S_0^{(4)} \rangle &= \frac{8\kappa^4 z^2 c^2}{\sigma_{xx} d} \int_{p,q} p^2 \sum_{k,m>0} \Big\{ 3D^3 D_p^c(m) \hat{S}_m D_q^c(k) + 3D^2 D_p^c(m) \hat{S}_m D D_q^c(k) \\
&\qquad + 2\kappa^2 zck D^2 [D_p^c]^2(m) \hat{T}_m \left[D_p(m) D_q^c(k) + D D_q^c(k) \right] \Big\} \\
&= 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon} \Big(-3T_{10}^0 - 3T_{11}^0 - \frac{12\ln\alpha}{\epsilon} - 6B_1 - 2T_{20}^0 + 2T_{21}^0 - 4T_{10}^1 + 4B_2 \Big) \\
&= 4\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2} \Big[4(\ln\alpha - 1)^2 - 2 + \frac{\epsilon}{2}\Big(\frac{2}{\alpha} - 2\ln^2\alpha + \ln\alpha + \frac{44}{3} \Big) \Big]. \qquad (3.4.22)
\end{aligned}$$

Finally,

$$\begin{aligned}
\langle O_2^{(4)} (S_{\text{int}}^{(4)} + \frac{1}{2}(S_{\text{int}}^{(3)})^2) \rangle \;&=\; \frac{16\kappa^4 z^2 c^2}{\sigma_{xx} D} \int_{p,q} (\mathbf{p}\cdot\mathbf{q}) \sum_{k>0} k D_{p+q}(0) D_p^2(k) D_q^2(k) \\
&- \frac{16\kappa^2 zc}{\sigma_{xx} D} \int_{p,q} p^2 \sum_{k>0} \Big[2\kappa^2 zck D_{p+q}^c(0) D_p^3(k) D_q(k) - D_p^3(k) D_q(k) \Big] \\
&+ \frac{16\kappa^4 z^2 c^2}{\sigma_{xx} D} \int_{p,q} (\mathbf{pq}) \sum_{k,m>0} \Big[2(1 + \kappa^2 zcm D_q(k)) D_{p+q}^c(m) D_p^2(k+m) D D_q^c(k) \\
&- D D_q^c(k) D D_p^c(m) D_{p+q}(k+m) \Big] - \frac{16\kappa^4 z^2 c^2}{\sigma_{xx} D} \int_{p,q} p^2 \sum_{k,m>0} \Big\{ (1 + \kappa^2 zcm \\
&\times D_{p+q}^c(m)) \Big[(2 + \hat{T}_m + \kappa^2 zck \hat{T}_m D_p^c(k)) D^3 D_p^c(k) D_q(k+m) + \frac{1}{2} D_q(k) \\
&\times D_p^3(k+m)(3D_q^c(k) + D_p^c(k+m)) \Big] + \frac{3}{2} D_q^c(m) D_{p+q}^c(k) D_p^3(k+m) + (1 \\
&+ \hat{T}_m + 2\kappa^2 zck \hat{T}_m D_p^c(k)) D_{p+q}^c(m) D^2 D_p^c(k) D_q(k+m) + \kappa^2 zck \hat{T}_m D [D_p^c(m)]^2 \\
&\times D D_q^c(k) D_{p+q}(k+m) \Big\} = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon} \Big(4S_{11} + 4S_{01} + \frac{2}{\epsilon} + 8A_{01} + 8A_{11} - 4C_{11} \\
&+ 4T_{02} + 4A_{10} + 2T_{12} + 2A_1 + 3T_{01} + 3A_{11}^1 + \alpha T_{10}^0 - T_{02} + H_1 + 3D_2 + 4C_1 \\
&+ 8A_2 + 8A_{00} - 4A_3 \Big) = 4\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon^2} \Big[-4\ln^2\alpha - 4 + \frac{\epsilon}{2}\Big(-\frac{2}{\alpha} - 2\ln^2\alpha - 25\ln\alpha \\
&+ \frac{55}{2} - 2\zeta(3) - \frac{8}{3}\pi^2 + 12\ln^2 2 - 44\ln 2 - 4C_0^{'} + 4A_0 + 16G - 8\,\mathrm{li}_2\Big(\frac{1}{2}\Big) \Big) \Big], \\
&\qquad (3.4.23)
\end{aligned}$$

where $G \approx 0.916$ denotes the Catalan constant.

3.4.4 Results of the computations

We proceed by presenting the final answer for all the pole terms in ϵ. By putting the external frequency equal to zero and in the limit $\alpha \to 0$ we obtain

$$\sigma'^{\,\mathrm{two}}_{xx} = 2\frac{\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon}\left(A - 8(2+\ln\alpha)\right). \tag{3.4.24}$$

Here, the A stands for all the terms that are finite in α. The complete list is as follows

$$\begin{aligned} A &= 50 + \frac{1}{6} - 3\pi^2 + \frac{19}{2}\zeta(3) + 16\ln^2 2 - 44\ln 2 + \frac{\pi^2}{2}\ln 2 + 16G \\ &+ \frac{\pi^4}{12} + \frac{\pi^2}{3}\ln^2 2 - \frac{1}{3}\ln^4 2 - 7\zeta(3)\ln 2 - 8\,\mathrm{li}_4(1/2) \approx 1.64. \end{aligned} \tag{3.4.25}$$

Before Eq. (3.4.24) is obtained, one has to deal with a host of other contributions that are more singular in α and/or ϵ. These more singular contributions all cancel one another in the end, however. There are in total six different types of contributions that are more singular than the simple pole term $1/\epsilon$. In Tables 3.1 and 3.2 we list these terms, show where they come from and how they sum up to zero. There is one exception, namely the terms proportional to $\ln\alpha/\epsilon$, and their contribution is written in Eq. (3.4.24). However, these terms are absorbed in the definition of an "effective" h' field. More specifically, from the two-loop computation of the singlet amplitude z we know that the effective h' field is given by [33]

$$h_0^2 \to h'^2 = h_0^2\left(1 - (2+\ln\alpha)\frac{h_0^\epsilon g_0}{\epsilon}\right), \tag{3.4.26}$$

where $g_0 = 4\Omega_D/\sigma_{xx}$ (see Section (2.2.4)). Using this result, as well as Eqs. (3.4.5) and (3.4.24), we can write the total answer for the conductivity as follows

$$\sigma'_{xx} = \sigma_{xx}\left(1 + 2\frac{h'^\epsilon g_0}{\epsilon} + 2A\frac{h'^{2\epsilon}g_0^2}{\epsilon}\right). \tag{3.4.27}$$

Eq. (3.4.27) no longer contains α and is therefore the desired result.

3.4.5 β and γ functions

For completeness we list the two-loop result for the singlet interaction amplitude [33]

$$z' = z\left(1 + \frac{h'^\epsilon g_0}{\epsilon} + \frac{h'^{2\epsilon}g_0^2}{2\epsilon^2}\left[1 - \epsilon(4+\pi^2/3)\right]\right). \tag{3.4.28}$$

Following the methodology of Section 2.2.4 we obtain the following relation between observable and bare theory

$$\frac{1}{g_0} = \frac{1}{g'}Z_1^{-1}(g'h'), \qquad z_0 = z'Z_2(g'h'), \tag{3.4.29}$$

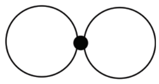
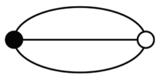
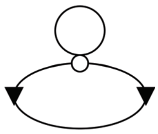
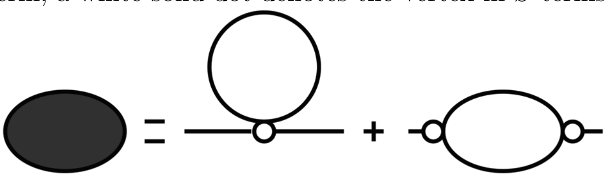
Contractions	Diagrams	$\frac{2}{\epsilon\alpha}$	$\frac{\ln^2\alpha}{(\epsilon/2)^2}$	$\frac{\ln\alpha}{(\epsilon/2)^2}$	$\frac{\ln^2\alpha}{\epsilon/2}$	$\frac{\ln\alpha}{\epsilon/2}$	$\frac{4}{\epsilon^2}$
$\langle O_1^{(4)} \rangle$			2				2
$\langle O_1^{(3)} S_{\text{int}}^{(3)} \rangle$			-4				-4
$\langle O_1^{(2)}(S_{\text{int}}^{(4)} + S_0^{(4)} + \frac{1}{2}(S_{\text{int}}^{(3)})^2) \rangle$			2				2
Total		0	0	0	0	0	0

Table 3.1: The second-loop contributions to the O_1 term in the effective conductivity. The α-dependent and $1/\epsilon^2$ contributions. A black solid dot denotes the vertex in O_1 term, a white solid dot denotes the vertex in S terms, and



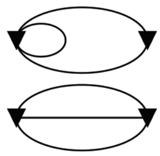
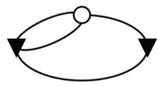
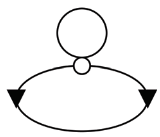
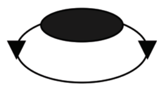
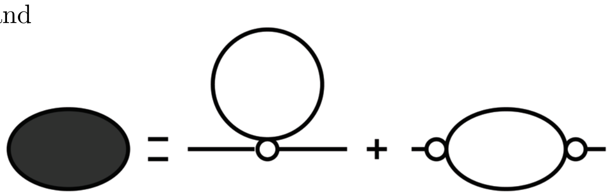
Contractions	Diagrams	$\frac{2}{\epsilon\alpha}$	$\frac{\ln^2\alpha}{(\epsilon/2)^2}$	$\frac{\ln\alpha}{(\epsilon/2)^2}$	$\frac{\ln^2\alpha}{\epsilon/2}$	$\frac{\ln\alpha}{\epsilon/2}$	$\frac{4}{\epsilon^2}$
$\langle O_2^{(6)} \rangle$				16		-4	-2
$\langle O_2^{(5)} S_{int}^{(3)} \rangle^*$				-8	4	20	4
$\langle O_2^{(4)} S_0^{(4)} \rangle$		2	4	-8	-2	1	2
$\langle O_2^{(2)} (S_{\rm int}^{(4)} + \frac{1}{2}(S_{\rm int}^{(3)})^2) \rangle^*$		-2	-4		-2	-25	-4
Total		0	0	0	0	- 8	0

Table 3.2: The second-loop contributions to the O_2 term in the effective conductivity. The α-dependent and $1/\epsilon^2$ contributions. The symbol * denotes that we exclude integrals A_0 and C_0' which cancel in the sum of the two terms. A black solid triangle denotes the current vertex in O_2 term, a white solid dot denotes the vertex in S terms, and



Contractions	$2/\epsilon$
$\langle O_1^{(4)}\rangle$	0
$\langle O_1^{(3)} S_{\rm int}^{(3)}\rangle$	$8+4\zeta(3)$
$\langle O_1^{(2)}(S_{\rm int}^{(4)}+S_0^{(4)}$ $+\frac{1}{2}(S_{\rm int}^{(3)})^2)\rangle$	$-4-2\zeta(3)-2\pi^2/3$
$\langle O_2^{(6)}\rangle$	$-\pi^2/3+\dfrac{\pi^2}{2}\ln 2+\pi^4/12+11\zeta(3)/2+\dfrac{\pi^2}{3}\ln^2 2$ $-\dfrac{1}{3}\ln^4 2-7\zeta(3)\ln 2-8\,\mathrm{li}_4(1/2)$
$\langle O_2^{(5)} S_{\rm int}^{(3)}\rangle^*$	$-12+4\zeta(3)+4\pi^2/3$
$\langle O_2^{(4)} S_0^{(4)}\rangle$	$44/3$
$\langle O_2^{(2)}(S_{\rm int}^{(4)}+\frac{1}{2}(S_{\rm int}^{(3)})^2)\rangle^*$	$55/2-2\zeta(3)-8\pi^2/3+12\ln^2 2$ $-44\ln 2+16G-8\,\mathrm{li}_2(1/2)$
Total	$34+1/6-3\pi^2+19\zeta(3)/2+16\ln^2 2-44\ln 2$ $+\dfrac{\pi^2}{2}\ln 2+16G+\pi^4/12+\dfrac{\pi^2}{3}\ln^2 2$ $-\frac{1}{3}\ln^4 2-7\zeta(3)\ln 2-8\,\mathrm{li}_4(1/2)$

Table 3.3: The two-loop contributions to the O_2 term in the effective conductivity. The $2/\epsilon$ contributions.

where

$$Z_1(\bar{g}) = 1 + 2\frac{\bar{g}}{\epsilon} + 4\frac{\bar{g}^2}{\epsilon^2}(1 + \epsilon A/2) \tag{3.4.30}$$

$$Z_2(\bar{g}) = 1 - \frac{\bar{g}}{\epsilon} - \frac{\bar{g}^2}{2\epsilon^2}\left(1 + \epsilon(\pi^2/6 + 2)\right). \tag{3.4.31}$$

Hence for the β and γ functions which are defined as (see section 2.2.4)

$$\beta(g) = \frac{dg}{d\ln\mu} = \frac{\epsilon g}{1 + g\dfrac{d\ln Z_1(g)}{dg}}, \tag{3.4.32}$$

$$\gamma_z(g) = -\frac{d\ln z}{d\ln\mu} = \beta(g)\frac{d\ln Z_2(g)}{dg}, \tag{3.4.33}$$

the final answer can be written as

$$\beta(g) = \epsilon g - 2g^2 - 4Ag^3, \qquad \gamma_z(g) = -g - (3 + \pi^2/6)g^2. \tag{3.4.34}$$

By using that

$$\beta_\sigma(\sigma_{xx}, 1) = -\pi\sigma_{xx}^2\beta(1/\pi\sigma_{xx}), \tag{3.4.35}$$

we find finally, that (see Eq. (3.2.22))

$$\begin{aligned}\beta_1(1) &= \frac{4}{\pi^2}A = \frac{4}{\pi^2}\Bigg[50 + \frac{1}{6} - 3\pi^2 + \frac{19}{2}\zeta(3) + 16\ln^2 2 - 44\ln 2 + \frac{\pi^2}{2}\ln 2 + 16G \\ &+ \frac{\pi^4}{12} + \frac{\pi^2}{3}\ln^2 2 - \frac{1}{3}\ln^4 2 - 7\zeta(3)\ln 2 - 8\,\mathrm{li}_4(1/2)\Bigg] \approx 0.66. \end{aligned}\tag{3.4.36}$$

3.5 Dynamical scaling

3.5.1 Relation between h' and ω_s

In this Section we combine the two loop computations of this chapter with those of the amplitude z' presented in Ref. [33] and establish the connection between the effective mass h' and the frequency ω_s. For this purpose, recall that the renormalization of the z field was obtained from the derivative of the grand potential Ω with respect to $\ln T$. [33] The result of the computation has been as follows

$$\frac{d\Omega}{d\ln T} = 2\sum_{s>0}\omega_s z M_b(g_0, h_s^2), \tag{3.5.1}$$

where

$$M_b(g_0, h_s^2) = 1 + \frac{h_s^\epsilon g_0}{\epsilon} + \frac{h_s^{2\epsilon} g_0^2}{2\epsilon^2}\left[-1 + \epsilon\left(2 + \pi^2/3\right)\right]. \tag{3.5.2}$$

Here, the frequency enters through the quantity $h_s^2 = \kappa^2 zs = 2\pi\omega_s z g_0/\Omega_d$ which has the dimension of the mass squared. The frequency dependence in $\sigma'_{xx}(s)$ is restored by writing

$$\sigma'_{xx}(s) = \frac{4\Omega_d}{g_0} R_b(g_0, h_s^2) \tag{3.5.3}$$

with

$$R_b(g_0, h_s^2) = 1 + 2\frac{h_s^\epsilon g_0}{\epsilon} + (2A-1)\frac{h_s^{2\epsilon} g_0^2}{\epsilon} \tag{3.5.4}$$

One can easily verify that Eqs (3.5.1)-(3.5.3) lead to the same expressions for Z_1 and Z_2 and, hence, the same β and γ functions as those of the previous Section. Eq. (3.5.3) is therefore the correct result. The relation between h' and ω_s can now be made more explicit by writing

$$h'^2 = h_s^2 M_b(g_0, h_s^2)/R_b(g_0, h_s^2). \tag{3.5.5}$$

Here, h' is the effective mass that is induced by the frequency ω_s and the result is consistent with all previous statements and explicit computations. [33]

3.5.2 The Goldstone phase

Specific heat and AC conductivity

The zero of the function $\beta(g)$, Eq. (3.4.34), determines a critical point $g_* = O(\epsilon)$ that separates the Goldstone or metallic phase ($g < g_*$) from an insulating phase ($g > g_*$). To second order in ϵ we have

$$g_* = \epsilon/2 - A\epsilon^2/2 \approx 0.5\epsilon - 0.82\epsilon^2 \tag{3.5.6}$$

We see that the ϵ^2 contribution is rather large and the expansion can clearly not be trusted for $\epsilon = 1$ or three spatial dimensions. This is a well-known drawback of asymptotic expansions and the two-loop theory is otherwise necessary to completely establish the scaling behavior of the electron gas in $2+\epsilon$ spatial dimensions. To discuss this scaling behavior, we proceed and express Eqs (3.5.1) and (3.5.3) in terms of the renormalized parameters g and z. The results can be written in the following general form

$$\frac{d\Omega}{d\ln T} = 2\sum_{s>0} \mu^{2\epsilon}\omega_s z M(g, \omega_s z), \tag{3.5.7}$$

$$\sigma'_{xx}(s) = \mu^\epsilon \frac{4\Omega_d}{g} R(g, \omega_s z). \tag{3.5.8}$$

The expressions are now finite in ϵ. The AC conductivity is obtained from $\sigma'_{xx}(s)$ by analytic continuation from imaginary frequencies $i\omega_s$ in the upper half-plane to the real frequencies ω. On the other hand, the specific heat of the electron gas can be expressed as [33]

$$c_v = \int_0^\infty d\omega \frac{\partial f_{BE}(\omega)}{\partial T} \omega \rho_{qp}(\omega), \tag{3.5.9}$$

where

$$f_{BE}(\omega) = \frac{1}{e^{\omega/T} - 1} \tag{3.5.10}$$

and

$$\rho_{qp}(\omega) = \frac{z}{\pi}(M(g, i\omega z) + M(g, -i\omega z)) \tag{3.5.11}$$

is the density of states of bosonic quasiparticles indicating that the Coulomb system is unstable with respect to the formation of particle-hole bound states. [110]

Scaling results

Next, from the method of characteristics we can obtain the general scaling behavior of the quantities M and R as usual:

$$M(g, \omega_s z) = M_0(g)G[\omega_s z \xi^d M_0(g)], \tag{3.5.12}$$

$$R(g, \omega_s z) = R_0(g)H[\omega_s z \xi^d R_0(g)]. \tag{3.5.13}$$

Here G and H are unspecified functions, whereas ξ, R_0 and M_0 each have a clear physical significance and are identified as the correlation length, the DC conductivity and $\rho_{qp}(0)$ respectively. They obey the following equations

$$\begin{aligned} (\mu\partial_\mu + \beta(g)\partial_g)\xi(g) &= 0, \\ (\beta(g)\partial_g - \epsilon - \beta(g)/g)R_0(g) &= 0, \\ (\beta(g)\partial_g + \gamma(g))M_0(g) &= 0. \end{aligned} \tag{3.5.14}$$

In the metallic phase ($g < g_*$) the solutions can be written as follows

$$R_0(g) = (1 - g/g_*)^{\epsilon\nu_\epsilon}, \qquad M_0(g) = (1 - g/g_*)^{\beta_\epsilon}, \tag{3.5.15}$$

$$\xi = \mu^{-1}\left(\frac{g}{4\Omega_d}\right)^{1/\epsilon}(1 - g/g_*)^{-\nu_\epsilon}, \tag{3.5.16}$$

where the critical exponents ν_ϵ and β_ϵ are obtained as

$$\nu_\epsilon^{-1} = -\partial_g\beta(g)\Big|_{g=g_*}, \qquad \beta_\epsilon = -\nu_\epsilon\gamma(g_*). \tag{3.5.17}$$

To second order in ϵ the results are

$$\nu_\epsilon^{-1} = \epsilon(1 + A\epsilon) \approx \epsilon + 1.64\epsilon^2 \tag{3.5.18}$$

$$\beta_\epsilon = 1/2 + (\pi^2/24 + 3/4 - A)\epsilon \approx 0.50 - 0.48\epsilon. \tag{3.5.19}$$

Notice that R_0 is not an independent quantity and can be expressed in terms of the correlation length ξ. For example, Eq. (3.5.8) can be written as

$$\sigma'_{xx}(s) = \xi^{-\epsilon}H[\omega_s z \xi^d M_0(g)], \tag{3.5.20}$$

where the function H is the same as in Eq. (3.5.13). Similarly one can write the quantity $h_s'^2$ (Eq. (3.5.5)) as follows:

$$h_s'^2 = \xi^{-2}K[\omega_s z \xi^d M_0(g)], \tag{3.5.21}$$

where the function $K(Y) = 4Y[G(Y)/H(Y)]$. Both the DC conductivity $R_0 \sim \xi^{-\epsilon}$ and the quantity M_0 vanish as one approaches the metal-insulator transition at $g = g_*$. The results are quite familiar from the Heisenberg ferromagnet where M_0 stands for the spontaneous magnetization. Unlike the free electron gas, [19] however, the interacting system with Coulomb interactions has a true order parameter, M_0, which is associated with a non-Fermi liquid behavior of the specific heat.

Equations of state

The explicit results of Section 3.5.1 can be used to completely determine the quantities M and R in the Goldstone phase. They take the form of an "equation of state" [93]

$$\frac{\omega_s z g}{M^{\delta_\epsilon}} = \left(\frac{g_*}{g}\right)^{1/\epsilon} \left(1 + (2\epsilon\nu_\epsilon - 1)(1 - g/g_*) - 2\epsilon\nu_\epsilon \frac{1 - g/g_*}{M^{1/\beta_\epsilon}}\right)^{1/\epsilon},$$

$$\frac{\omega_s z g}{R^{\kappa_\epsilon}} = \left(\frac{g_*}{g}\right)^{1/\epsilon} \left(1 - \frac{1 - g/g_*}{R^{1/2\epsilon\nu_\epsilon}}\right)^{1/\epsilon}. \tag{3.5.22}$$

Here, the exponents δ_ϵ and κ_ϵ can be obtained from the values of ν_ϵ and β_ϵ following the relations

$$d\nu_\epsilon = \beta_\epsilon(\delta_\epsilon + 1), \qquad \epsilon\nu_\epsilon\kappa_\epsilon = \beta_\epsilon\delta_\epsilon. \tag{3.5.23}$$

The universal features of the "equations of state" are the Goldstone singularities at $g = 0$ and the critical singularities near $g = g_*$.

Equations (3.5.22) provide the complete solution of the physical observables, Eqs (3.5.8) and (3.5.9), of the theory in $D = 2 + \epsilon$ dimension. As for the specific heat, we find the usual behavior $c_v = \gamma_0 T$ at $g = 0$. At criticality the following algebraic behavior is found

$$c_v = \gamma_1 T^{1+1/\delta_\epsilon}, \tag{3.5.24}$$

where

$$\delta_\epsilon^{-1} = \frac{\epsilon}{4}\left[1 - \left(A - \frac{\pi^2}{12} - \frac{5}{4}\right)\epsilon\right] \approx 0.25\epsilon + 0.11\epsilon^2. \tag{3.5.25}$$

It is worthwhile to mention that the one-loop result for δ_ϵ^{-1} (contribution of the order of ϵ) has been obtain in Ref. [109].

It is important to remark that the result for the conductivity $\sigma'_{xx}(s)$, Eq. (3.5.8), can also be used to obtain the scaling behavior at finite temperatures. For example, we may, on simple dimensional grounds, substitute T for ω_s in expressions for $h_s'^2, \sigma'_{xx}(s)$ and R. The results, however, strictly describe the Goldstone and critical phases only. The "equations of state" cannot be analytically continued and used to obtain information on the insulating phase. As we already mentioned in the Introduction to this Chapter, the strong coupling phase is controlled by different operators in the theory and has a distinctly different frequency and temperature dependence.

3.6 Renormalization group flows in $D = 2+\epsilon$ dimensions

It is helpful to discuss the structure of the theory in $2+\epsilon$ dimensions for general crossover parameter $0 \leqslant c \leqslant 1$ (see Eqs (3.2.19) and (3.2.21)),

$$\frac{d\sigma_{xx}}{d\ln\mu} = \beta_\sigma(\sigma_{xx}, c). \tag{3.6.1}$$

$$\frac{dc}{d\ln\mu} = \beta_c(\sigma_{xx}, c) = (1-c)\gamma_z(\sigma_{xx}, c). \tag{3.6.2}$$

The renormalization group flow lines in the (σ_{xx}, c) plane are sketched in Fig. 3.1. We see that there are two critical fixed points describing a quantum phase transition between a metal and an insulator. As it was shown in Section 3.5 along the Coulomb line ($c = 1$) the fixed point value is $\sigma^*_{xx} = O(\epsilon^{-1})$. However, along the Fermi liquid line ($c = 0$), the renormalization group function become

$$\beta_\sigma(\sigma_{xx}, 0) = -\epsilon\sigma_{xx} + \frac{1}{2\pi^2\sigma_{xx}} + \frac{3}{8\pi^4\sigma^3_{xx}}, \qquad \gamma_z(\sigma_{xx}, 0) = 0, \tag{3.6.3}$$

where we add the four-loop contribution to the $\beta_\sigma(\sigma_{xx}, 0)$. [57] There is the critical fixed point $\sigma^*_{xx} = (1/\sqrt{2\pi^2\epsilon})(1 + 3\epsilon/4)$ with the critical exponent $\nu_\epsilon^{-1} = 2\epsilon(1 + 3\epsilon/2)$. In general the renormalization is determined, to a major extend, by the global symmetries of the problem. [42] In particular, since $\mathcal{F}$ invariance is retained along the Coulomb line $c = 1$ only and broken otherwise one generally expects, that the problem with finite range interactions $0 < c < 1$ lies in the domain of attraction of the Fermi liquid line $c = 0$ whereas the Coulomb interaction problem $c = 1$ describes a distinctly different, non-Fermi liquid universality class. The results presented in Section (3.5) and renormalization group flow shown in Fig. 3.1 completely prove this scenario.

In two spatial dimensions the metallic phases ($\sigma_{xx} > \sigma^*_{xx}$) disappear altogether indicating that all the states of the (spin polarized or spinless) electron gas are now Anderson localized, independent of the presence of electron-electron interactions. We consider the Coulomb interaction problem ($c = 1$) in two dimensions ($\epsilon = 0$). Since the response parameter σ'_{xx} is independent of the arbitrary momentum scale μ_0 that defines the "renormalized" theory $\sigma_{xx}(\mu_0)$ we immediately obtain from Eq. (3.2.19) the general scaling result

$$\sigma'_{xx} = \sigma_{xx}(\mu') = f_\sigma(\mu'\xi) \tag{3.6.4}$$

where μ' is related to the linear dimension L of the system according to $\mu' = L^{-1}$. The ξ obeys the differential equation

$$(\mu_0\partial_{\mu_0} + \beta_\sigma\partial_{\sigma_{xx}})\,\xi = 0 \tag{3.6.5}$$

and can be identified with a dynamically generated correlation length (localization length) of the system

$$\xi = \mu_0^{-1}\sigma_{xx}^{-\beta_1(1)/\beta_0^2(1)} e^{\sigma_{xx}/\beta_0(1)}. \tag{3.6.6}$$

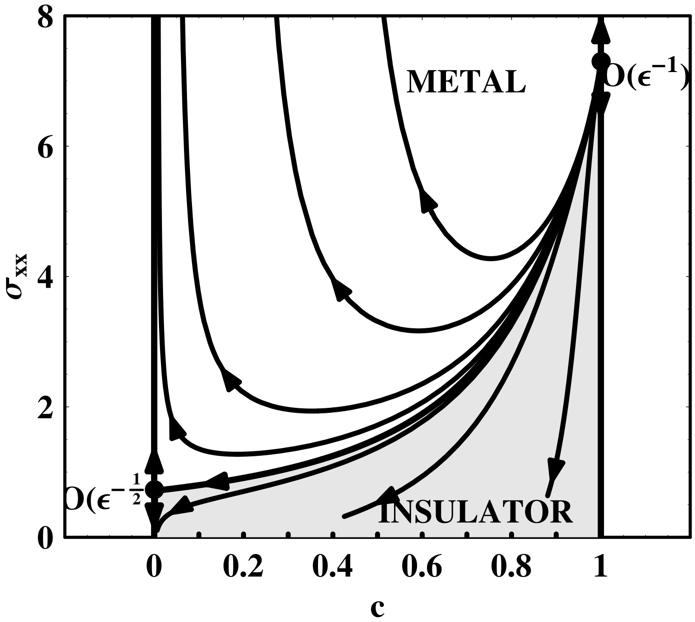


Figure 3.1: Renormalization group flow in terms of σ_{xx} and c. Here $\epsilon = 0.1$.

Next, comparison of Eqs. (3.6.4) and (3.6.6) with the expression of Eq. (3.2.19) leads to the following explicit (weak coupling) result for the scaling function $f_\sigma(X)$ with $X = (\mu'\xi)^{\beta_0(1)}$

$$f_\sigma(X) \approx \ln X + \frac{\beta_1(1)}{\beta_0(1)} \ln\ln X + \frac{\beta_1^2(1)}{\beta_0^2(1)} \frac{\ln\ln X}{\ln X}, \quad X \gg 1. \tag{3.6.7}$$

The statement of exponential localization can now be formulated by saying that in the regime of strong coupling the scaling function $f(X)$ vanishes according to

$$f_\sigma(X) \approx \exp\left(-X^{-1/\beta_0(1)}\right) = \exp\left(-1/(\mu'\xi)\right), \quad X \ll 1. \tag{3.6.8}$$

On the other hand, the quantity z' is quite analogous to the *spontaneous magnetization* (M_0) in the classical Heisenberg ferromagnet. Write

$$z' = zM_0(X) \tag{3.6.9}$$

then the following explicit form of $M_0(X)$ can be extracted

$$M_0(X) = m_0 f_z(X) \tag{3.6.10}$$

where

$$m_0 = \sigma_{xx}^{\gamma_0/\beta_0(1)} \left[1 + \frac{\gamma_0\beta_1(1) - \gamma_1(1)\beta_0(1)}{\beta_0^2(1)\sigma_{xx}}\right] \quad (3.6.11)$$

$$f_z(X) \approx (\ln X)^{-\gamma_0/\beta_0(1)} \Big(1 - \frac{\beta_1(1)\gamma_0}{\beta_0^2(1)} \ln\ln X - \frac{2\beta_1^2(1)\gamma_0}{\beta_0^3(1)} \frac{\ln\ln X}{\ln X}\Big), \qquad X \gg 1. \quad (3.6.12)$$

As we shall demonstrate in the next chapter these naive expectations are fundamentally modified by the $\theta(\nu_f)$ dependence of the theory which is invisible in perturbative expansions of the renormalization group β and γ functions in powers of $1/\sigma_{xx}$. Notice that on the basis of the Mirmin-Wagner-Coleman theorem one would expect that the quantity $f_z(X)$, like $f_\sigma(X)$, vanishes in the regime of strong localization $X \ll 1$.

3.7 Conclusions

In this chapter we have completed the two-loop analysis of the Finkelstein theory with the singlet interaction term. We have reported the detailed computations of the conductivity which is technically the most difficult part of the analysis. We have benefitted from the regularization procedure involving the h_0 field, which has substantially simplified the two-loop computations. Moreover, we have obtained a general relation between the effective masses that are being induced by the h_0 field on the one hand, and the frequency ω_n on the other. This enables one to re-express the final answer in terms of finite frequencies and/or temperature, simply by a substitution of the h_0 regulating field.

By combining the concept of $\mathcal{F}$ invariance with technique of dimensional regularization, we have extracted new physical information on the disordered electron gas with Coulomb interactions in low dimensions. In particular, we now have a non-Fermi liquid theory for the specific heat and dynamical scaling.

As it will be clear from the next chapter, the metal-insulator transition in $2+\epsilon$ dimensions sets the stage for the plateau transitions in the quantum Hall regime.

3.A Computation of integrals

In this Appendix we present the final results for the various integrals listed in Eqs (3.4.17)-(3.4.23). We shall follow the same methodology as used in the two-loop computation of Ref. [33] and employ the standard representation for the momentum and frequency integrals in terms of the Feynman variables x_1, x_2 and x_3. [42] We classify the different contributions in Eqs (3.4.17)-(3.4.23) in different categories, labelled A-integrals, B-integrals etc. In total we have seven different categories, i.e. A, B, C, D, H, S and T respectively, which are discussed separately in Sections 3.A.1-3.A.5 of this Appendix. The last Section, 3.A.8, contains a list of abbreviations and a list of symbols for those integrals that need not be computed explicitly because their various contributions sum up to zero in the final answer.

In Appendix 3.B we present the main computational steps for a specific example, the so-called A_{10}-integral. We show how the integral representation of hypergeometric functions can be used to define both the ϵ expansion and the limit where $\alpha \to 0$.

3.A.1 The A - integrals

Definition

To set the notation, we consider the integral

$$X^{\nu}_{\nu,\eta} = -\frac{2^{1+\nu}(\kappa^2 zc)^{2+\mu}}{\sigma_{xx} D^{\nu}} \int\limits_{pq} p^{2\nu} \sum_{k,m>0} m^{\mu} D^{c}_{p+q}(m) D D^{c}_{p}(k) D^{1+\mu+\eta}_{q}(k+m). \quad (3.A.1)$$

Here, the three indices μ, ν and η generally take on the values $0, 1$. We shall only need those quantities $X^{\nu}_{\nu,\eta}$ which have $\eta = \nu$, however. Using the Feynman trick, one can write (for the notation, see Section 3.A.5)

$$\begin{aligned} X^{\nu}_{\nu,\eta} \quad &= \quad -\frac{2^{1+\nu}(\kappa^2 zc)^{2+\mu}}{\sigma_{xx} D^{\nu}} \int\limits_{pq} p^2 \int\limits_0^{\infty} dm\, m^{\mu} \int\limits_0^{\infty} dk \, \frac{\Gamma(\mu+\eta+4)}{\Gamma(\mu+\eta+1)} \int\limits_{\alpha}^{1} dz \int [] \, x_2 x_3^{\mu+\eta} \\ &\times \left[h_0^2 + q^2 x_{12} + p^2 x_{13} + 2\mathbf{p}\cdot\mathbf{q} x_1 + \kappa^2 zm(\alpha x_1 + x_3) + \kappa^2 zk(z x_2 + x_3)\right]^{-\mu-\eta-4} \end{aligned} \quad (3.A.2)$$

Next, by shifting $\mathbf{q} \to \mathbf{q} - \mathbf{p} x_1/x_{12}$, we can decouple the vector variables $\mathbf{p}$ and $\mathbf{q}$ in the denominator. The integration over k, m, p and q then leads to an expression that only involves the integral over z and the Feynman variables x_1, x_2 and x_3. Write

$$X^{\nu}_{\nu,\eta} = \frac{2\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon} A^{\nu}_{\mu,\eta} \quad (3.A.3)$$

then

$$A^{\nu}_{\mu\eta} = \int\limits_{\alpha}^{1} dz \int [] \frac{x_2 x_3^{1+\mu+\eta}(x_1+x_2)^{\nu}(x_i x_j)^{-1-\nu-\epsilon/2}}{(z x_2 + x_3)(\alpha x_1 + x_3)^{1+\mu}}. \quad (3.A.4)$$

To complete the list of A-integrals, we next define quantities that carry either two indices μ, ν or only a single index μ. Like $A^{\nu}_{\mu\eta}$, they all describe contractions that contain both momentum and frequency integrals. The results are all expressed in terms of integrals over z, x_1, x_2 and x_3.

$$A_{\nu\mu} = \int_\alpha^1 dz(z-\alpha)^{1+\nu-\mu} \int [] \frac{x_1^\mu x_2^{2+\nu-\mu} x_3^\mu (x_1+x_3)^{1-\mu}(x_i x_j)^{-2-\epsilon/2}}{(\alpha x_1+x_3)^{1+\nu}(z x_2+x_3)}, \quad (3.A.5)$$

$$A_0 = \int_\alpha^1 dz(z-\alpha) \int [] \frac{x_2^2 x_1 (x_i x_j)^{-2-\epsilon/2}}{(x_3+z x_2)(z x_2+\alpha x_1+2x_3)}, \quad (3.A.6)$$

$$A_1 = \int_\alpha^1 dz(z-\alpha)^2 \int [] \frac{x_2^3(x_1+x_3)(x_2+x_3)}{(z x_2+x_3)^2}(x_i x_j)^{-2-\epsilon/2} \times \Big[\frac{1}{(\alpha x_1+x_3)^2} - \frac{1}{(z x_2+\alpha x_1+2x_3)^2}\Big], \quad (3.A.7)$$

$$A_2 = \int_\alpha^1 dz(z-\alpha)(1-z) \int [] \frac{x_2^3(x_1+x_3)(x_i x_j)^{-2-\epsilon/2}}{(z x_2+x_3)(\alpha x_1+x_3)(z x_2+\alpha x_1+2x_3)}, \quad (3.A.8)$$

$$A_3 = \int_\alpha^1 dz(z-\alpha) \int [] \frac{x_2^2(x_1+x_3)(x_i x_j)^{-2-\epsilon/2}}{(\alpha x_1+z x_2+2x_3)(z x_2+x_3)}. \quad (3.A.9)$$

ϵ **expansion**

The calculation of integrals is straightforward but tedious and lengthy. Here we only present the final results of those quantities that are needed. The list does not contain the final answer for the A_0-integral because the various contributions to A_0 sum up to zero in the final answer. The same holds for some other integrals that are defined in Section 3.A.5 and that we do not specify any further.

$$A_{00}^{0} = -\frac{\ln^2\alpha}{\epsilon/2} + \zeta(3), \tag{3.A.10}$$

$$A_{10}^{0} = -\frac{\ln^2\alpha + \ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2} + \frac{\pi^2}{6} + \zeta(3), \tag{3.A.11}$$

$$A_{01}^{1} = \frac{\ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2} - 2\ln\alpha - \frac{\pi^2}{3} + 1, \tag{3.A.12}$$

$$A_{11}^{1} = \frac{\ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2} - 2\ln\alpha - \frac{\pi^2}{3}, \tag{3.A.13}$$

$$A_{00} = \frac{\ln\alpha}{\epsilon/2} + \frac{\ln^2\alpha}{2} + 2\ln\alpha + \frac{\pi^2}{3} - 1, \tag{3.A.14}$$

$$A_{10} = -\frac{1}{\alpha} - \frac{2\ln\alpha + 3}{\epsilon/2} - \ln^2\alpha - 5\ln\alpha - \frac{2\pi^2}{3} + 3, \tag{3.A.15}$$

$$A_{01} = -\ln\alpha - \frac{\pi^2}{6} + 1, \tag{3.A.16}$$

$$A_{11} = \frac{\ln\alpha + 2}{\epsilon/2} + \frac{\ln^2\alpha}{2} + 3\ln\alpha + \frac{\pi^2}{2}, \tag{3.A.17}$$

$$\begin{aligned} A_1 &= -\frac{2}{\alpha} + \frac{2\ln^2\alpha + 4\ln\alpha}{\epsilon/2} - 3\ln^2\alpha + 8\ln 2\ln\alpha - \frac{17}{2}\ln\alpha \\ &\quad + 4K_1(\alpha) + 8J_3^{'}(\alpha) - \pi^2 - 2\zeta(3) - 6\ln^2 2 + 10\ln 2 - \frac{1}{2}, \end{aligned} \tag{3.A.18}$$

$$\begin{aligned} A_2 &= -\frac{\ln^2\alpha + 2\ln\alpha}{\epsilon/2} - 2\ln\alpha - 3\ln 2\ln\alpha - J_1(\alpha) - K_1(\alpha) - 2J_3^{'}(\alpha) \\ &\quad + A_0 - \frac{\pi^2}{6} + 1 + \zeta(3) + 3\ln^2 2 - 3\ln 2 - 3\,\mathrm{li}_2(1/2), \end{aligned} \tag{3.A.19}$$

$$A_3 = A_0 - 2\,\mathrm{li}_2\,(1/2) + \frac{\pi^2}{6}. \tag{3.A.20}$$

3.A.2 The B - integrals

Definition

The B-integrals are similarly defined in terms of the variables z, x_1, x_2 and x_3. However, they describe only those contractions that contain frequency sums and no momentum integrals.

$$B_\mu = \int_\alpha^1 \frac{dz}{z^\mu} \int [] \frac{x_1^{\mu-1} x_2 x_3^{-\mu-\epsilon/2} (x_1 + x_2)^{-\mu-\epsilon/2}}{(\alpha x_2 + z x_3 + x_1)}, \tag{3.A.21}$$

ϵ expansion

$$B_1 = \frac{\ln\alpha}{\epsilon/2} + \frac{\ln^2\alpha}{2} + \ln\alpha, \tag{3.A.22}$$

$$B_2 = -\frac{1}{\alpha} + \frac{\ln^2\alpha}{\epsilon/2} + \frac{2\ln\alpha}{\epsilon/2} - 2\ln\alpha - 2. \tag{3.A.23}$$

3.A.3 The C - integrals

Definition

The C-integrals contain one additional integration over y, besides the ones over z and the Feynman variables x_1, x_2 and x_3. They originate from expressions involving integrations over both frequencies and momenta. We distinguish between quantities with two indices μ and ν

$$C_{\mu\nu} = \int_\alpha^1 dzdy \int [] \frac{x_1^\mu x_2 x_3 (x_2+x_3)^{1-\mu}(x_i x_j)^{-2-\epsilon/2}}{(zx_3+x_1)(yx_2+x_1)^\nu (zx_3+yx_2)^{1-\nu}} \tag{3.A.24}$$

and those that carry only a single index ν

$$C_\nu = \int_\alpha^1 dz(1-z)^\nu \int_\alpha^1 dy \int [] \frac{x_2^{2-\nu} x_3^{1+\nu}(x_1+x_2)^\nu (x_i x_j)^{-2-\epsilon/2}}{(zx_3+x_1)(yx_2+x_1)^\nu (zx_3+yx_2+2x_1)}. \tag{3.A.25}$$

ϵ expansion

$$\begin{aligned} C_{00} &= \frac{\ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2} - 2\ln\alpha + \frac{\pi^2}{4}\ln 2 - \frac{\pi^2}{6} + \frac{15}{4}\zeta(3) \\ &\quad -\frac{\pi^4}{24} - \frac{\pi^2}{6}\ln^2 2 + \frac{1}{6}\ln^4 2 + \frac{7}{2}\zeta(3)\ln 2 + 4\,\mathrm{li}_4(\frac{1}{2}), \end{aligned} \tag{3.A.26}$$

$$C_{01} = \frac{2\ln\alpha}{\epsilon/2} - \ln^2\alpha - 4\ln\alpha - 2 - \zeta(3), \tag{3.A.27}$$

$$C_{11} = \zeta(3), \tag{3.A.28}$$

$$C_0 = \frac{\ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2} - 2\ln\alpha - 1 - \zeta(3) - C_0', \tag{3.A.29}$$

$$C_1 = 4\ln 2\ln\alpha + 2J_1(\alpha) - C_0' - 2 - \frac{\zeta(3)}{2} - 4\ln 2 - \frac{\pi^2}{6} + 4G, \tag{3.A.30}$$

where the Catalan constant $G = 0.916\ldots$ appears as the integral

$$G = -\int_0^1 du \frac{\ln u}{1+u^2}. \tag{3.A.31}$$

3.A.4 The D-integrals

Definition

These are integrals over the Feynman variables only. They originate from the contractions which contain sums over both momenta and frequencies.

$$D_\nu = \int [] \frac{x_3^\nu (x_1+x_2)^{\nu-1}(x_i x_j)^{-\nu-\epsilon/2}}{(\alpha x_1 + x_3)(\alpha x_2 + x_3)}. \tag{3.A.32}$$

ϵ expansion

$$D_1 = -\ln^2\alpha - \frac{\pi^2}{6}, \tag{3.A.33}$$

$$D_2 = -2\ln\alpha. \tag{3.A.34}$$

3.A.5 The H - integrals

Definition

The H-integrals involve the variable z and the Feynman variables. All of them originate from contractions with sums over both momenta and frequencies.

$$H_\nu = \int_\alpha^1 dz (z-\alpha)^{2\nu} \int [] \frac{x_2^{2+\nu}(x_1+x_3)(x_i x_j)^{-2-\epsilon/2}}{(\alpha x_1 + z x_2)(z x_2 + x_3)}. \tag{3.A.35}$$

ϵ expansion

$$H_0 = -\ln\alpha + 1, \tag{3.A.36}$$

$$H_1 = -\ln\alpha. \tag{3.A.37}$$

3.A.6 The S - integrals

Definition

These are integrals over the Feynman variables only and they do not not contain the parameter α. All of them originate from the expressions with sums over both momenta and frequencies.

$$S_{\mu\nu} = \int [] \frac{x_1^\mu x_2^{1+\nu-\mu}((2-\nu-\mu)x_1 + x_3)(x_i x_j)^{-2-\epsilon/2}}{(x_2+x_3)^{1+\nu}}, \tag{3.A.38}$$

$$S_\nu = \int [](x_1+x_2)^{-1+2\nu}(x_i x_j)^{-1-\nu-\epsilon/2}. \tag{3.A.39}$$

ϵ expansion

$$S_{00} = -\frac{2}{\epsilon} + 2, \tag{3.A.40}$$

$$S_{01} = -\frac{2}{3\epsilon} + \frac{8}{9}, \tag{3.A.41}$$

$$S_{11} = -\frac{1}{3\epsilon} + \frac{1}{9}, \tag{3.A.42}$$

$$S_{0} = -\frac{2}{\epsilon} + 2, \tag{3.A.43}$$

$$S_{1} = -\frac{4}{\epsilon} + 2. \tag{3.A.44}$$

3.A.7 The T-integrals

Definition

The integrals are over the Feynman variables only. They come from the expressions which only contain sums over frequency.

$$T^{\eta}_{\mu\nu} = \frac{(1-\alpha)^{\eta}}{\alpha^{\mu}} \int [] \frac{x_1^{2-\eta} x_2^{\mu+\eta-1} x_3^{-1-\mu-\epsilon/2} (x_1+x_2)^{-2-\epsilon/2}}{(\alpha x_2 + \nu\alpha x_3 + x_1)}, \tag{3.A.45}$$

$$T_{\mu\nu} = \int [] \frac{x_1^{2\nu-2} (x_1+x_2)^{-\nu-\epsilon} (x_1 + x_3 + (\alpha+\mu) x_2)}{x_3^{\nu+\epsilon/2} (\alpha x_2 + (1+\mu) x_3 + x_1)(x_1 + x_3 + \alpha x_2)}. \tag{3.A.46}$$

ϵ expansion

$$T^{0}_{10} = -\frac{1}{\alpha} + 1, \tag{3.A.47}$$

$$T^{0}_{11} = -\frac{1}{\alpha} + \frac{2}{\epsilon} + \ln\alpha + 1, \tag{3.A.48}$$

$$T^{0}_{20} = \frac{1}{6\alpha^2} - \frac{1}{3\alpha} - \ln\alpha - \frac{11}{12}, \tag{3.A.49}$$

$$T^{0}_{21} = \frac{1}{6\alpha^2} + \frac{2}{3\alpha} + \frac{\ln\alpha + 5/2}{\epsilon} + \frac{\ln^2\alpha}{2} + 4\ln\alpha + \frac{17}{12}, \tag{3.A.50}$$

$$T^{1}_{10} = -\frac{1}{\alpha} - 2\ln\alpha - 2, \tag{3.A.51}$$

$$T_{01} = \frac{\ln\alpha}{\epsilon/2} - \frac{\ln^2\alpha}{2}, \tag{3.A.52}$$

$$T_{02} = \frac{2}{\epsilon}, \tag{3.A.53}$$

$$
\begin{aligned}
T_{12} &= -\frac{3\ln\alpha + 11/2}{\epsilon/2} + \frac{3\ln^2\alpha}{2} + \frac{9\ln\alpha}{2} \\
&\quad -4\ln 2\ln\alpha + \frac{\pi^2}{6} - 4Li_2(\frac{1}{2}) - 12\ln 2 + \frac{27}{4}. \qquad (3.A.54)
\end{aligned}
$$

3.A.8 List of symbols and abbreviations

$$
\begin{aligned}
\int [] &= \int_0^1 dx_1 \int_0^1 dx_2 \int_0^1 dx_3\, \delta(x_1 + x_2 + x_3 - 1), \qquad (3.A.55) \\
x_{ij} &= x_i + x_j, \qquad (3.A.56) \\
x_i x_j &= x_1 x_2 + x_2 x_3 + x_3 x_1. \qquad (3.A.57)
\end{aligned}
$$

$$
\begin{aligned}
K_1(\alpha) &= \int_\alpha^1 dz \int [] \frac{x_2(x_1(x_2+x_3) + x_3^2)(x_i x_j)^{-2-\epsilon/2}}{(zx_2 + x_3)(\alpha x_1 + zx_2 + 2x_3)}, \qquad (3.A.58) \\
J_3'(\alpha) &= \alpha \int_\alpha^1 \frac{dz}{z} \int [] \frac{x_2(x_1(x_2+x_3) + x_3^2)(x_i x_j)^{-2-\epsilon/2}}{(\alpha x_1 + zx_2 + 2x_3)^2}, \qquad (3.A.59) \\
J_1(\alpha) &= \int_\alpha^1 dz \int [] \frac{x_1(x_1+x_3)(x_2+x_3)(x_i x_j)^{-2-\epsilon/2}}{(zx_1 + x_3)(zx_1 + \alpha x_2 + 2x_3)}, \qquad (3.A.60) \\
C_0' &= \int_\alpha^1 dz dy \int [] \frac{x_1 x_2^2 (x_i x_j)^{-2-\epsilon/2}}{(x_3 + yx_2)(zx_1 + yx_2 + 2x_3)}. \qquad (3.A.61)
\end{aligned}
$$

3.B Example of calculation for a typical integral

In this appendix we present the calculation of the integral A_{10} as a typical example. We start with the integral

$$
X_{10} = -\frac{32(\kappa^2 zc)^3}{\sigma_{xx} D} \int_{pq} p^2 \sum_{k,m>0} m D^c_{p+q}(m) D^3 D^c_p(k) D_q(k+m). \qquad (3.B.1)
$$

Using the Feynman trick, one can write

$$
\begin{aligned}
X_{10} &= -\frac{16(\kappa^2 zc)^3}{\sigma_{xx} D} \int_{pq} p^2 \int_0^\infty dm\, m \int_0^\infty dk\, \Gamma(6) \int_\alpha^1 dz (z-\alpha)^2 \int [] \Big[h_0^2 + q^2 x_{13} \\
&\quad + p^2 x_{12} + 2\mathbf{p}\cdot\mathbf{q} x_1 + \kappa^2 zm(\alpha x_1 + x_3) + \kappa^2 zk(zx_2 + x_3) \Big]^{-6}. \qquad (3.B.2)
\end{aligned}
$$

Shifting $\mathbf{q} \to \mathbf{q} - \mathbf{p}x_1/x_{13}$, we can decouple $\mathbf{p}$ and $\mathbf{q}$ in the denominator. We are then able to perform the integration over k, m, p and q, resulting in

$$X_{10} = \frac{8\Omega_D^2 h_0^{2\epsilon}}{\sigma_{xx}\epsilon} A_{10}, \tag{3.B.3}$$

where

$$A_{10} = \int_\alpha^1 dz (z-\alpha)^2 \int [] \frac{x_2^3(x_1+x_3)(x_i x_j)^{-2-\epsilon/2}}{(zx_2+x_3)(\alpha x_1 + x_3)^2}. \tag{3.B.4}$$

Next we write the integral as a sum of four terms

$$\begin{aligned} A_{10} &= \int_\alpha^1 \frac{dz(z-\alpha)^2}{z} \int [] \frac{x_2(x_1+x_3)(x_i x_j)^{-1-\epsilon}}{(\alpha x_1 + x_3)^2} \Bigg\{ 1 - x_1 x_3 (x_i x_j)^{-1} \\ &\quad - \frac{x_3(x_1+x_3)(x_i x_j)^{-1}}{z} + \frac{x_3^2(x_1+x_3)(x_i x_j)^{-1}}{z(zx_2+x_3)} \Bigg\} \\ &= I_0 - I_1 - I_2 + I_3. \end{aligned} \tag{3.B.5}$$

In what follows we retain the full ϵ dependence in the I_0, I_1 and I_2 and it suffices to put $\epsilon = 0$ in the fourth piece I_3. Introducing a change of variables

$$x_1 = \frac{u}{s+1}, \qquad x_2 = \frac{s}{s+1}, \qquad x_3 = \frac{1-u}{s+1}, \tag{3.B.6}$$

where $0 < s < \infty$ and $0 < u < 1$, then the four different pieces can be written as follows

$$I_0 = \left(\frac{1}{2} - 2\alpha\right) \int_0^1 \frac{du}{(\alpha u + 1 - u)^2} \int_0^\infty ds \frac{s(s+1)^\epsilon}{(s + u(1-u))^{1+\epsilon/2}}, \tag{3.B.7}$$

$$I_1 = \frac{1}{2} \int_0^1 du \frac{u(1-u)}{(\alpha u + 1 - u)^2} \int_0^\infty ds \frac{s(s+1)^\epsilon}{(s + u(1-u))^{2+\epsilon/2}}, \tag{3.B.8}$$

$$I_2 = \int_0^1 du \frac{(1-u)}{(\alpha u + 1 - u)^2} \int_0^\infty ds \frac{s(s+1)^\epsilon}{(s + u(1-u))^{2+\epsilon/2}}, \tag{3.B.9}$$

$$\begin{aligned} I_3 &= \int_\alpha^1 dz \left(\frac{z-\alpha}{z}\right)^2 \int_0^1 du \frac{u(1-u)^2}{(\alpha u + 1 - u)^2} \\ &\qquad \times \int_0^\infty ds \frac{(s+1-u)}{(s+u(1-u))^2(\alpha s + 1 - u)^2}. \end{aligned} \tag{3.B.10}$$

The integrals over s in Eqs (3.B.7)-(3.B.10) can now be recognized as integral representations of the hypergeometric function ${}_2F_1$. Write

$$I_0 = \frac{1-4\alpha}{2}\int_0^1 du \frac{[u(1-u)]^{1-\epsilon/2}}{(\alpha u+1-u)^2}\left[-\frac{G_0(u(1-u))}{1+\epsilon/2}+\frac{G_1(u(1-u))}{\epsilon/2}\right], \quad (3.B.11)$$

$$I_1 = -\frac{1}{2}\int_0^1 du \frac{[u(1-u)]^{1-\epsilon/2}}{(\alpha u+1-u)^2}\left[\frac{G_1(u(1-u))}{\epsilon/2}-\frac{G_2(u(1-u))}{1-\epsilon/2}\right], \quad (3.B.12)$$

$$I_2 = -\int_0^1 du \frac{u^{-\epsilon/2}(1-u)^{1-\epsilon/2}}{(\alpha u+1-u)^2}\left[\frac{G_1(u(1-u))}{\epsilon/2}-\frac{G_2(u(1-u))}{1-\epsilon/2}\right], \quad (3.B.13)$$

$$I_3 = \frac{1}{4}\int_\alpha^1 dz \left(\frac{z-\alpha}{z}\right)^2 \int_0^1 \frac{du\;[2uH_3(1-\alpha u)+(1-u)H_4(1-\alpha u)]}{zu+1-u}, \quad (3.B.14)$$

then, in the limit where $\epsilon \to 0$, we can identify the functions G_i and H_i as follows

$$G_0(1-z) = {}_2F_1(1,-\epsilon,-\epsilon/2;z) \to \frac{1+z}{1-z}, \quad (3.B.15)$$

$$G_1(1-z) = {}_2F_1(1,-\epsilon,1-\epsilon/2;z) \to 1+\epsilon\ln(1-z), \quad (3.B.16)$$

$$G_2(1-z) = {}_2F_1(1,-\epsilon,2-\epsilon/2;z) \to 1, \quad (3.B.17)$$

and

$$H_3(z) = {}_2F_1(1,2,3;z) = -\frac{2}{z^2}\left(\ln(1-z)+z\right), \quad (3.B.18)$$

$$H_4(z) = {}_2F_1(1,2,4;z) = \frac{6}{z^3}\left((1-z)\ln(1-z)+z-z^2/2\right). \quad (3.B.19)$$

Using these results we obtain

$$I_0 = -\frac{1}{\alpha}-\frac{\ln\alpha+2}{\epsilon/2}-\frac{\ln^2\alpha}{2}-2\ln\alpha-\frac{\pi^2}{3}, \quad (3.B.20)$$

$$I_1 = \frac{\ln\alpha+2}{\epsilon/2}+\frac{\ln^2\alpha}{2}+2\ln\alpha+\frac{\pi^2}{3}, \quad (3.B.21)$$

$$I_2 = \frac{\ln\alpha+1}{\epsilon/2}+\frac{\ln^2\alpha}{2}+2\ln\alpha+\frac{\pi^2}{3}+1, \quad (3.B.22)$$

$$I_3 = -\ln\alpha. \quad (3.B.23)$$

The final answer is therefore

$$A_{10} = -\frac{1}{\alpha}-\frac{2\ln\alpha+3}{\epsilon/2}-\ln^2\alpha-5\ln\alpha-\frac{2\pi^2}{3}+3 \quad (3.B.24)$$

Chapter 4

θ renormalization and electron-electron interactions

4.1 Introduction

One of the long standing mysteries in the theory of the plateau transitions in the quantum Hall regime is the apparently insignificant or subdominant role that is played by the long range Coulomb interaction between the electrons. The pioneering experiments on quantum criticality in the quantum Hall regimes by H. P. Wei *et al.*, [23] for example, are in many ways a carbon copy of the scaling predictions based on the field theory of Anderson localization in strong magnetic fields. [22] The initial success of the free electron theory has primarily led to a widely spread believe in Fermi liquid type of ideas [100, 101, 102, 103, 104, 105] as well as an extended literature on scaling and critical exponent phenomenology. [76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89]

Except for experimental considerations, however, there exists absolutely no valid (microscopic) argument that would even remotely justify any of the different kinds of free (or *nearly* free) electron scenarios that have frequently been proposed over the years. In fact, Fermi liquid principles are fundamentally in conflict with the novel insights that have more recently emerged from the development of a microscopic theory on interaction effects. [32, 33, 35] These developments are naturally based on the topological concept of an *instanton vacuum* [18] which is very well known to be the fundamental mechanism by which the free electron gas *de-localizes* in two spatial dimensions and in strong magnetic fields. [22] The outstanding and difficult problem that one is faced with is whether or not the topological concepts in quantum field theory retain their significance also when the electron-electron interactions are taken into account. [31]

For a variety of reasons, however, it has taken a very long time before the subject matter gained the physical clarity that it now has. [66] Perhaps the most awkward obstacles were provided by the historical controversies [67] in QCD where the idea of an instanton parameter θ arose first but its meaning remained rather obscure. [111] These controversies have mainly set the stage for the wrong physical ideas and the wrong mathematical objectives. For example, in sharp contrast to the general expectations in the field [112, 113, 114, 39, 40, 41] the fundamental problems do not reside in the conventional aspects of disordered systems such as the replica method or "exact" critical exponent values. A more fundamental issue has emerged, the *massless chiral edge excitations*, [34] that dramatically change the way in which the θ parameter is generally being perceived. [66] A detailed understanding of the physics of the edge has resolved, amongst many other things, the long standing controversies that historically have spanned the subject such as the *quantization of topological charge*, [37] the meaning of *instantons* and *instanton gases* [37, 38] etc. As a result of all this we can now state that the instanton angle θ *generically* displays all the basic features of the quantum Hall effect, independent of the details such as the replica limit. This includes not only the appearance of *gapless* excitations at $\theta = \pi$ but also the most fundamental and much sought after aspect of the theory, the existence of *robust topological quantum numbers* that explain the precision and observability of the quantum Hall effect. [66]

A second major complication in dealing with interaction effects is the notorious complexity of the underlying theory. [30] Although Finkelstein's original ideas in the field have been very illuminating, it has nevertheless taken herculean efforts to understand how the generalized non-linear σ model approach can be studied as a field theory. This includes not only the theory of perturbative expansions [33] but also such basic aspects like the global symmetries of the problem ($\mathcal{F}$ invariance), electrodynamic $U(1)$ gauge invariance as well as the physical observables of the theory. [32] These advances are absolutely necessary if one wants to extend the perturbative theory of localization and interaction effects to include the highly non-trivial consequences of the θ vacuum.

It obviously makes an enormously big difference to know that the instanton vacuum theory of the quantum Hall effect is NOT merely an isolated critical exponent problem that exists in replica field theory or "super symmetric" extensions of free electron approximations alone. Contrary to this widely spread misconception in the literature the fundamental features of the quantum Hall effect actually reveal themselves as a *super universal* consequence of topological principles in quantum field theory that until to date have not been well understood. The concept of *super universality* makes it easier and more natural to comprehend why the basic phenomena of scaling are retained by the electron gas also when the Coulomb interaction between the electrons is taken into account. Moreover, it facilitates the development of a unifying theory that includes completely different phenomena like the fractional quantum Hall regime. Unlike Fermi liquid ideas, however, *super universality* does not necessarily imply that the quantum critical details at $\theta = \pi$ remain the same. The various different applications of the θ vacuum concept do in general have different exponent values at $\theta = \pi$ and, hence, they belong to different *universality classes.*

In this Chapter we revisit the problem of topological excitations (instantons) and θ renormalization [20, 21] in the theory of the interacting electron gas. The results of an early analysis of instanton effects have been reported in a short paper by Pruisken and Baranov. [31] However, much of the conceptual structure of the theory was not known at that time, in particular the principle of $\mathcal{F}$ *invariance* and the appearance of the *massless edge excitations* that together elucidate the fundamental aspects of the

θ vacuum on the strong coupling side. [66] These novel insights unequivocally define the physical observables (i.e. the conductance parameters σ_{xx} and σ_{xy}) that control the dynamics of the θ vacuum at low energies. These physical observables should therefore quite generally be regarded as some of the most fundamental quantities of the theory.

A detailed knowledge of instanton effects on the physical observables of the theory has fundamental significance since it bridges the gap that exists between the *weak coupling* Goldstone singularities at short distances, and the *super universal* features of the quantum Hall effect that generally appear at much larger distances only. The theory of observable parameters, as it now stands, provides the general answer to the "arena of bloody controversies" that historically arose because of a complete lack of any physical objectives of the theory. A prominent and exactly solvable example of these statements is given by the large N expansion of the CP^{N-1} model that, unlike the previous expectations, sets the stage for all the non-perturbative features of the θ parameter that one is interested in Ref. [66].

The main objective of the present chapter is to review the instanton methodology, provide the technical details of the computation and extend the analysis in several ways. Our study of the interacting electron gas primarily relies on the procedure of *spatially varying masses* that has been applied in the context of the ordinary $U(m+n)/U(m) \times U(n)$ non-linear σ model in Chapter 2. The important advantage of this procedure is that it facilitates non-perturbative computations of the renormalization group β and γ functions of the theory. These computations, together with the new insights into the strong coupling features and symmetries of the problem, lay out the complete phase and singularity structure of the disordered electron gas. The results of this paper, which include the non-Fermi liquid behavior of the Coulomb interaction problem, obviously cannot be obtained in any different manner.

This chapter is organized as follows. We start out in Section 2.2 with a brief introduction to the formalism and recall the effective action procedure for massless chiral edge excitations. In Section 4.2.2 we briefly elaborate on the general topological principles that explain the *robust* quantization of the Hall conductance. The general argument is deeply rooted in the methods of quantum field theory and relies on the relation that exists between the conductances on the one hand, and the sensitivity of the interacting electron gas to infinitesimal changes in the boundary conditions on the other. The argument is furthermore based on the relation between Kubo formalism, the background field methodology and the effective action for chiral edge excitations which was described in Section 3.3.

In Section 3.2.3 we give the complete list of physical observables which then serves as the basic starting point for the remainder of this paper. We show the general relationship between the physical observables and the renormalization group β and γ functions.

In Section 4.3 we recall the various different aspects associated with instanton matrix field configurations and embark on the problem of quantum fluctuations. We use the method of spatially varying masses described in Chapter 2 and end the Section with the complete action for the quantum fluctuations in Tables 4.3 and 4.4.

In Section 4.4 together with Appendix 4.A we present the results of detailed computations that deal, amongst many other things, with the technical difficulties associated with the theory in Pauli-Villars regularization, the *replica method* as well as the *infinite sums* over Matsubara frequency indices that are inherent to the problem of electron-electron interactions.

In Section 4.5 we address the various different aspects associated with the integration over zero modes and embark on the general problem of transforming the Pauli-Villars masses in curved space back into flat space following the methodology introduced by 't Hooft. [44] This finally leads to the most important advances of this paper, the renormalization-group β and γ functions which are evaluated at a non-perturbative level. These final results provide a unified theory of the disordered electron gas that includes the effects of both *finite* range electron-electron interactions and *infinite* range interactions such as the Coulomb potential. We end this Chapter with a discussion in Section 4.8.

4.2 Formalism

4.2.1 The action

For convenience of a reader we remind that the generalized replica non-linear sigma model involves unitary matrix field variables $Q^{\alpha\beta}_{nm}(\mathbf{r})$ that obey the following constraints

$$Q = Q^\dagger, \quad \operatorname{tr} Q = 0, \quad Q^2 = 1. \tag{4.2.1}$$

The superscripts $\alpha, \beta = 1, \ldots, N_r$ represent the *replica* indices and the subscripts n, m are the indices of the *Matsubara* frequencies $\omega_k = \pi T(2k+1)$ with $k = n, m$. A convenient representation in terms of unitary matrices $\mathcal{T}(\mathbf{r})$ is obtained by writing

$$Q(\mathbf{r}) = \mathcal{T}^{-1}(\mathbf{r})\Lambda\mathcal{T}(\mathbf{r}), \quad \Lambda^{\alpha\beta}_{nm} = \operatorname{sign}(\omega_n)\delta^{\alpha\beta}\delta_{nm}. \tag{4.2.2}$$

The effective action for the two-dimensional interacting electron gas in the presence of disorder and a perpendicular magnetic field can be written as follows [32]

$$Z = \int \mathcal{D}[Q] \exp S, \qquad S = S_\sigma + S_F. \tag{4.2.3}$$

Here, S_σ is the *free electron* action [19]

$$S_\sigma = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\operatorname{tr}(\nabla Q)^2 + \frac{\sigma_{xy}}{8}\int d\mathbf{r}\,\operatorname{tr}\varepsilon_{ab}Q\nabla_a Q\nabla_b Q. \tag{4.2.4}$$

The quantities σ_{xx} and σ_{xy} represent the *meanfield* values for the *longitudinal* and *Hall* conductances in units e^2/h respectively. The symbol $\varepsilon_{ab} = -\varepsilon_{ba}$ stands for the antisymmetric tensor. Next, S_F contains the *singlet interaction* term [30, 32]

$$S_F = \pi T z \int d\mathbf{r}\, O_F[Q], \tag{4.2.5}$$

where

$$O_F[Q] = c\sum_{\alpha n} \operatorname{tr} I^\alpha_n Q \operatorname{tr} I^\alpha_{-n} Q + 4\operatorname{tr}\eta Q - 6\operatorname{tr}\eta\Lambda. \tag{4.2.6}$$

Here, z is the so-called *singlet interaction amplitude*, T the temperature and c the *crossover* parameter which allows the theory be interpolated between the case of electrons with Coulomb interaction ($c = 1$) and the free electron case ($c = 0$). The singlet interaction term involves a matrix

$$(I_n^\alpha)_{km}^{\beta\gamma} = \delta^{\alpha\beta}\delta^{\alpha\gamma}\delta_{k,n+m} \tag{4.2.7}$$

which is the Matsubara representation of the $U(1)$ generator $\exp(-i\omega_n\tau)$ with τ being imaginary time. Matrix

$$\eta_{nm}^{\alpha\beta} = n\delta^{\alpha\beta}\delta_{nm} \tag{4.2.8}$$

is used to represent the set of the Matsubara frequencies ω_n. In what follows we shall regulate of the infrared of the system by the finite size L rather than the infrared regulator h_0^2 that we have used in the previous Chapter.

4.2.2 Quantization of the Hall conductance

As it was shown in Chapter 2 the robust quantization of the Hall conductance can be demonstrated on the basis of very general principles such as mass generation and the fact that the conductances can be expressed in terms of the response of the system to changes in the boundary conditions. Below we briefly repeat the arguments presented in Chapter 2 for noninteracting electrons and extend it to the case of the Coulomb interactions. The subtleties of the argument involve a novel and previously unexpected ingredient of the instanton vacuum concept, however, which has been recognized very recently only. The main problem resides in the σ_{xy} term in Eq. (4.2.4) which is formally identified as the *topological charge* $\mathcal{C}[Q]$ associated with the matrix field configuration Q. Assuming for simplicity the geometry of a square of size $L \times L$ then we can express the topological charge in terms of both a *bulk* integral and an *edge* integral as follows

$$\mathcal{C}[Q] = \frac{1}{16\pi i}\int d\mathbf{r}\,\mathrm{tr}\,\varepsilon_{ab}Q\nabla_a Q\nabla_b Q = \frac{1}{4\pi i}\oint dx\,\mathrm{tr}\,\mathcal{T}\nabla_x\mathcal{T}^{-1}\Lambda. \tag{4.2.9}$$

As we discussed in great details in Chapter 2 the remarkable thing that is usually overlooked is that the matrix field Q generally splits up into distinctly different components, each with a distinctly different topological significance and very different physical properties. For this purpose we introduce a change of variables

$$Q = t^{-1}Q_0 t. \tag{4.2.10}$$

Here, the Q_0 is an arbitrary matrix field with boundary conditions $Q_0 = \Lambda$ at the edge (or, equivalently, $\mathcal{T}_0$ equals an arbitrary $U(N) \times U(N)$ gauge at the edge). The fixed unitary matrix field t generally represents the fluctuations about the special boundary conditions. This change of variables is just a formal way of splitting the topological charge $\mathcal{C}[Q]$ of an arbitrary matrix field configuration Q into an *integral* piece $\mathcal{C}[Q_0]$ and a *fractional* piece $\mathcal{C}[q]$,

$$\mathcal{C}[Q] = \mathcal{C}[Q_0] + \mathcal{C}[q], \qquad q = t^{-1}\Lambda t. \tag{4.2.11}$$

Without a loss in generality we can write

$$\mathcal{C}[Q_0] \in \mathbb{Z}, \qquad -\frac{1}{2} < \mathcal{C}[q] \leqslant \frac{1}{2}. \tag{4.2.12}$$

The main new idea is that the matrix field t or q corresponding to all possible boundary conditions for Q field should be taken as a dynamical variable in the problem, rather than being a fixed quantity that one can choose freely. The reason is that one can generally associate *massless chiral edge excitations* with the fluctuating matrix fields q. These so-called *edge modes* q are distinctly different from the *bulk modes* Q_0 which usually (i.e. for arbitrary values of σ_{xy}) generate dynamically a *mass gap* in the bulk of the system. These various statements immediately suggest that the low energy dynamics of the strong coupling phase is described by an effective action of the matrix field variable q obtained by formally eliminating the *bulk modes* Q_0. This effective action procedure is furthermore based on the fact that the mean field quantity σ_{xy} (which is equal to the filling fraction ν_f of the Landau levels) can in general be split into an *integral edge* part $k(\nu_f)$ and a *fractional bulk* piece $\theta(\nu_f)$ as follows

$$\sigma_{xy} = \nu_f = k(\nu_f) + \frac{\theta(\nu_f)}{2\pi}, \tag{4.2.13}$$

where

$$k(\nu_f) \in \mathbb{Z}, \qquad -\pi < \theta(\nu_f) \leqslant \pi. \tag{4.2.14}$$

In what follows we shall separately consider the theory with $c = 0$ (free particles) and $c = 1$ (Coulomb interactions) both of which are invariant under the action of renormalization group. [33]

Free particles ($c = 0$)

In the absence of external frequencies and at $T = 0$ we can write the action for the free electron gas as follows

$$S = S_\sigma^{\text{edge}}[q] + S_\sigma^{\text{bulk}}[Q], \tag{4.2.15}$$

where

$$S_\sigma^{\text{edge}}[q] = 2\pi i k(\nu_f)\mathcal{C}[q], \tag{4.2.16}$$

$$S_\sigma^{\text{bulk}}[Q] = -\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 + i\theta(\nu_f)\mathcal{C}[Q]. \tag{4.2.17}$$

Provided the matrix field variable t satisfies the classical equations of motion we can obtain an effective action for q by eliminating the bulk matrix field Q_0

$$S_{\text{eff}}[q] = S_\sigma^{\text{edge}}[q] + S_{\text{eff}}^{\text{bulk}}[q], \tag{4.2.18}$$

where

$$\exp S_{\text{eff}}^{\text{bulk}}[q] = \int_{\partial V} \mathcal{D}[Q_0] \exp S_\sigma^{\text{bulk}}[t^{-1}Q_0 t]. \tag{4.2.19}$$

Here the subscript ∂V indicates that the functional integral has to be performed with $Q_0 = \Lambda$ at the edge. The effective action for the bulk can be written as (see Section 2.2.3)

$$S_{\text{eff}}^{\text{bulk}}[q] = -\frac{\sigma'_{xx}}{8}\int d\mathbf{r}\,\text{tr}(\nabla q)^2 + i\theta'\mathcal{C}[q]. \tag{4.2.20}$$

Here, $\sigma'_{xx} = \sigma_{xx}(L)$ and $\theta' = \theta(L)$ play the role of response parameters that measure the sensitivity of the system to an infinitesimal change in the boundary conditions. For exponentially localized states these parameters are expected to vanish for large enough L and the effective action is now given by the edge piece (4.2.16) alone. This one dimensional action is known to describe *massless chiral edge excitations.* [34] To obtain a suitably regulated action for the edge we may proceed by stacking many blocks of size $L \times L$ on top of one another to form an infinite strip (see Fig. 4.1). The action for the quantum Hall state is then defined along infinite edges and can be written as [32]

$$S_{\text{eff}}[q] = \frac{k(\nu_f)}{2}\oint dx\,\text{tr}\,t\nabla_x t^{-1}\Lambda + \pi T\rho_{\text{edge}}\oint dx\,\text{tr}\,\eta q. \tag{4.2.21}$$

Here we have introduced a frequency term to regulate the infrared and $\sigma'_{xy} = k(\nu_f)$ indicates that the Hall conductance is robustly quantized.

At this stage several remarks are in order. First of all, from an explicit (non-perturbative) computation of the response parameters σ'_{xx} and θ' we know that the argument generally fails for $\theta(\nu_f) = \theta' = \pi$ where the mass gap vanishes and the system is quantum critical. This happens at the center of the Landau bands where a transition takes place between adjacent quantum Hall plateaus.

Secondly, it is important to keep in mind that the aforementioned argument for an exact quantization of the Hall conductance is entirely based on the fact that the edge modes q are *massless*. The beauty of the effective action procedure is that it unequivocally demonstrates that the so-called *spherical boundary conditions* (i.e. $Q_0 = \Lambda$ at the edge) are dynamically generated by the system itself, independent of any weak coupling arguments such as finite action requirements and independent of N_r and N_m. The bulk components Q_0 have mathematically very interesting properties in that they are a realization of the formal homotopy theory result

$$\pi_2(SU(2N)/S(U(N)\times U(N))) = \pi_1(S(U(N)\times U(N))) = \mathbb{Z}. \tag{4.2.22}$$

The integer $\mathbb{Z}$ is equal to the topological charge $\mathcal{C}[Q_0]$ which is identified as the jacobian for the mapping of the manifold $U(2N)/U(N)\times U(N)$ onto the two-dimensional plane. Physically the quantization of $\mathcal{C}[Q_0]$ represents the *quantization of flux* and the integer $k(\nu_f)\mathcal{C}[Q_0]$ can be interpreted in terms of a *discrete number* of electrons that have crossed the Fermi energy at the edge of the system.

Notice that except for the massless chiral edge modes there exists no compelling reason to believe why the topological charge $\mathcal{C}[Q_0]$ and, hence, the Hall conductance is *robustly quantized*. In fact, the quantization of topological charge has been one of the longstanding and controversial issues in quantum field theory [37] that have fundamentally complicated the development of a microscopic theory of the quantum Hall effect.

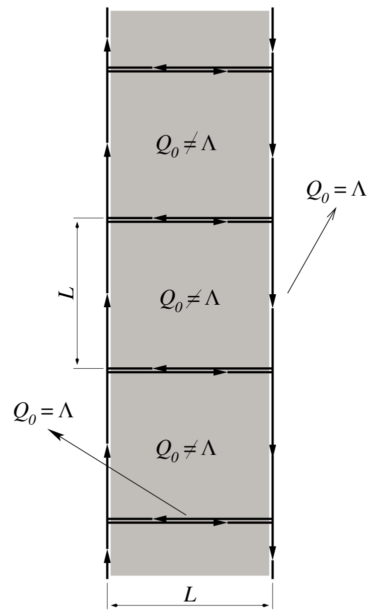


Figure 4.1: Geometry of an infinite strip

Coulomb interaction ($c = 1$)

An extension of the effective action procedure to the problem with the long range Coulomb interaction is by no means obvious. The argument relies, to a major extend, on the detailed knowledge obtained from an explicit analysis of the Finkelstein approach which shows that the theory undergoes structural changes in the limit where $N_r \to 0$ and $N_m \to \infty$. The action is more complicated and now given by

$$S = S_\sigma^{\text{edge}}[q] + S_\sigma^{\text{bulk}}[Q] + S_F[Q], \tag{4.2.23}$$

where $c = 1$ is inserted in the expression for $S_F[Q]$. Elimination of the matrix field variable Q_0 leads to the definition of the effective action

$$\exp S_{\text{eff}}^{\text{bulk}}[q] = \int_{\partial V} \mathcal{D}[Q_0] \exp\left(S_\sigma^{\text{bulk}}[t^{-1}Q_0t] + S_F[t^{-1}Q_0t]\right). \tag{4.2.24}$$

On the basis of symmetries one can write down the following explicit result (see Section 2.2.3)

$$S_{\text{eff}}^{\text{bulk}}[q] = -\frac{\sigma'_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla q)^2 + i\theta'\mathcal{C}[q] + \mathcal{O}(T). \tag{4.2.25}$$

Here, the response parameters $\sigma'_{xx} = \sigma_{xx}(L)$ and $\theta' = \theta(L)$ are evaluated in the limit where $T \to 0$. It is important to emphasize that S_F cannot be omitted from Eqs. (4.2.23)-(4.2.25). The reason is, as we already mentioned before, that this term fundamentally affects the ultra violet singularity structure of the theory. [30, 33]

The remaining part of the argument proceeds along similar lines as before. Provided the system with Coulomb interactions generates a mass gap, both parameters σ'_{xx} and θ' should vanish for L large enough. The complete action for the quantum Hall state has been obtained previously but only for the much simpler case where the Fermi energy is located in a Landau gap. The parameters σ_{xx}, θ as well as the singlet interaction amplitude z are identically to zero from the start in this case and the result is [34]

$$\begin{aligned} S_{\text{eff}}[q] &= \frac{k(\nu_f)}{2} \oint dx \operatorname{tr} t \nabla_x t^{-1} \Lambda + \frac{\pi^2}{2} T \rho_{\text{edge}} \oint dx O_F[q] \\ &- \frac{\pi}{4} T k(\nu_f) \oint dx \oint dy \operatorname{tr} I^\alpha_{-n} q(x) v_{\text{eff}}(x-y) \operatorname{tr} I^\alpha_n q(y). \end{aligned} \quad (4.2.26)$$

As before we have $\sigma'_{xy} = k(\nu_f)$. Here, the quantity $v_{\text{eff}}(x-y)$ contains the Coulomb interaction $U_0(x-y) = 1/|x-y|$. The Fourier transform is given by

$$v^{-1}_{\text{eff}}(p_x) = \frac{k(\nu_f)}{2\pi\rho_{\text{edge}}} \left[1 + \rho_{\text{edge}} \int dp_y \, U_0 \left(\sqrt{p_x^2 + p_y^2} \right) \right]. \quad (4.2.27)$$

It can been shown that Eq. (4.2.26) is equivalent to the action for $k(\nu_f)$ *chiral bosons* in $1+1$ space time dimension [34]

$$\begin{aligned} S_{\text{chiral}} &= -\frac{i}{4\pi} \sum_{i=1}^{k(\nu_f)} \int_0^{1/T} d\tau \oint dx \, \partial_x \phi_i \left(\partial_\tau - i \frac{k(\nu_f)}{2\pi\rho_{\text{edge}}} \partial_x \right) \phi_i \\ &- \frac{1}{8\pi^2} \sum_{i,j=1}^{k(\nu_f)} \int_0^{1/T} d\tau \oint dx \oint dy \, \partial_x \phi_i(x) U_0(x-y) \partial_y \phi_j(x'). \end{aligned} \quad (4.2.28)$$

This result has been used as the starting point for a first principle derivation of the complete Luttinger liquid theory of the edge for the abelian quantum Hall states. [35]

4.2.3 Physical observables

General remarks

The fundamental problem that one is faced with is how the interacting electron gas manages to go from an ordinary metallic state at short distances to the quantum Hall state described by Eqs (4.2.26) - (4.2.28) that one generally expects at much larger distances only. Pursuing a satisfactory understanding of this phenomenon it is obviously the same thing as developing a microscopic theory for the *observable* parameters σ'_{xx}, σ'_{xy} or θ'. We have seen that these transport parameters automatically appear as an integral aspect of the effective action procedure of massless chiral edge excitations. As we have shown in the previous Chapter that σ'_{xx} and σ'_{xy} can quite generally be identified with the Kubo formulae for the conductances. One of the main difficulties, however, is that a detailed knowledge of the transport parameters generally involves a detailed knowledge the interaction terms that are linear in T. These terms, for all practical purposes act, like unusual mass terms that affect the renormalization behavior of the electron gas at $T = 0$. This means, amongst many other things, that the *dynamic generation* of the quantum Hall state is not merely a statement that is made on the transport quantities σ'_{xx} and σ'_{xy} alone. In particular, the arguments of the previous Section must be extended to also include the interactions terms indicated by $\mathcal{O}(T)$ in Eq. (4.2.25). Before one could even think of any attempt in this direction it is absolutely necessary to first recapitulate the perturbative renormalization group results obtained from the *background field* methodology which is a slight modification of the effective action procedure. For computational purposes it is convenient to add a $U(N) \times U(N)$ invariant mass term to Eq. (4.2.24) such that the effective action for arbitrary values of c now becomes

$$e^{S_{\text{eff}}^{\text{bulk}}[q]} = \int \mathcal{D}[Q] \exp\Big(S_\sigma^{\text{bulk}}[t^{-1}Qt] + S_F[t^{-1}Qt] + h^2 \int d\mathbf{r}\, \text{tr}\, \Lambda Q \Big). \quad (4.2.29)$$

Provided the length scale induced by the h^2 field is much smaller than the actual sample size one can evaluate $S_{\text{eff}}^{\text{bulk}}[q]$ in a systematic fashion by using the standard methods for perturbative expansions in $2+\epsilon$ dimensions. Assuming that the "background field" t obeys the classical equations of motions of S_σ then an explicit computation of $S_{\text{eff}}^{\text{bulk}}[q]$ to lowest orders in an expansion in powers of both the gradients and T leads to the following expression

$$\begin{aligned} S_{\text{eff}}^{\text{bulk}}[q] &= -\frac{\sigma'_{xx}}{8} \int d\mathbf{r}\, \text{tr}(\nabla q)^2 + i\theta \mathcal{C}[q] \\ &+ \pi T z' \int d\mathbf{r} \left(c' \sum_{\alpha n} \text{tr}\, I_n^\alpha q\, \text{tr}\, I_{-n}^\alpha q + 4\, \text{tr}\, \eta q - 6\, \text{tr}\, \eta \Lambda \right). \end{aligned} \quad (4.2.30)$$

Eq.(4.2.30) is of the same form as the original action except that the bare parameters are now replaced by the "observable" ones that contain radiative corrections. Notice that $\theta' = \theta$ which is a direct consequence of the fact that the θ term is invisible in perturbative expansions. The results nevertheless indicate that besides the quantities σ'_{xx} and θ' one should generally extend the list of observable parameters to include z' and c' as well. Eq. (4.2.30) is consistent with the results originally obtained by Finkelstein in a different physical context. [30] Important for our purposes is the fact the quantities σ'_{xx}, z' and c' depend on the h field only through an "induce" momentum scale $\mu' = \mu'(h) \ll L^{-1}$. On the other hand, in the limit $h \to 0$, the infrared of Eq. (4.2.29) is regulated by the sample size L rather than h. Under these circumstances one expects that the results of Eq. (4.2.30) are unchanged except that the momentum scale μ' is now fixed by L rather than h.

Kubo formula

The response quantities σ'_{xx} and θ' for arbitrary values of c can be expressed in terms of current-current correlations according to [21, 33] (see Eqs. (3.2.13) and (3.3.4))

$$\begin{aligned} \sigma'_{xx} &= -\frac{\sigma_{xx}}{4n} \left\langle \text{tr}[I_n^\alpha, Q(\mathbf{r})][I_{-n}^\alpha, Q(\mathbf{r})] \right\rangle \\ &+ \frac{\sigma_{xx}^2}{8n} \int d\mathbf{r} \left\langle \text{tr}\, I_n^\alpha Q(\mathbf{r}) \nabla Q(\mathbf{r})\, \text{tr}\, I_{-n}^\alpha Q(\mathbf{r}') \nabla Q(\mathbf{r}') \right\rangle, \end{aligned} \quad (4.2.31)$$

$$\sigma'_{xy} = \sigma_{xy} + \frac{\sigma_{xx}^2}{8n} \int d\mathbf{r} \varepsilon_{ab} \langle \text{tr}\, I_n^\alpha Q(\mathbf{r}) \nabla_a Q(\mathbf{r})\, \text{tr}\, I_{-n}^\alpha Q(\mathbf{r}') \nabla_b Q(\mathbf{r}') \rangle. \quad (4.2.32)$$

Here and from now onward the expectations are defined with the respect to the theory of Eq. (4.2.3)-(4.2.6) and we assume spherical boundary conditions.

Specific heat

A natural definition of the observable quantity z' is obtained through the derivative of the thermodynamic potential with respect to temperature which is directly related to the specific heat of the electron gas. [33] Write

$$\begin{aligned}\frac{\partial \ln \Omega}{\partial \ln T} &= \pi T z \int d\mathbf{r} \langle O_F[Q] \rangle \\ &= \pi T z' \int d\mathbf{r}\, O_F[\Lambda] \end{aligned} \tag{4.2.33}$$

then the expression for z' becomes (see Section 3.3)

$$z' = z \frac{\langle O_F[Q] \rangle}{O_F[\Lambda]}. \tag{4.2.34}$$

The expression for remaining observable c' is determined by the general condition imposed on the *static response* of the system which says that the quantity $z\alpha = z(1-c)$ remains unaffected by the quantum fluctuations. [30, 32, 33] The second equation therefore reads as follows

$$z'(1-c') = z(1-c) \qquad \text{or} \qquad z'\alpha' = z\alpha. \tag{4.2.35}$$

In the previous Chapter we have worked out the perturbative expressions for the observable parameters. In what follows we proceed and employ Eqs. (4.2.34) and (4.2.35) for non-perturbative computational purposes as well. A justification of this procedure is given in Section 4.7 where we embark on the various different subtleties associated with instanton calculus.

β and γ functions

The expressions of the previous Sections facilitate renormalization group studies that include not only ordinary perturbative expansions but also the non-perturbative effects of instantons. Let μ' denote the momentum scale associated with the observable theory then the quantities $\sigma'_{xx} = \sigma_{xx}(\mu')$, $z' = z(\mu')$ and $c' = c(\mu')$ can be expressed in terms of the renormalization group β and γ functions according to (see chapter 3)

$$\sigma'_{xx} = \sigma_{xx} + \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu} \beta_\sigma(\sigma_{xx}, c), \tag{4.2.36}$$

$$\theta' = \theta + \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu} \beta_\theta(\sigma_{xx}, c), \tag{4.2.37}$$

$$z' = z - \int_{\mu_0}^{\mu'} \frac{d\mu}{\mu} \gamma_z(\sigma_{xx}, c) z, \tag{4.2.38}$$

$$z'\alpha' = z\alpha. \tag{4.2.39}$$

Here β_θ and γ_z functions are given by Eqs. (3.2.22)-(3.2.23) and $\beta_\theta = 0$.

4.2.4 Sensitivity to boundary conditions

Let us next come back to the background field procedure of Eqs (4.2.29) and (4.2.30) and see how one in general can learn something about the problem on the strong coupling side. We are, in particular, interested to see how the perturbative results of the renormalization group can be reconciled with Thouless' idea of exponential localization. This means that upon entering the strong coupling phase the bulk of the system renders insensitive to changes in the boundary conditions. We specialize from now onward to the theory in two spatial dimensions and discard, for the time being, θ term. We shall start the discussion with the Coulomb interaction problem since it displays many of the features that are familiar from the ordinary non-linear σ model. We then proceed with the less familiar aspects that are very specific to the physics of the electron gas.

Weak versus strong coupling

Putting $c = c' = 1$ then the effective action for the edge field variable q, Eqs (4.2.24) and (4.2.25), can be written as follows

$$
\begin{aligned}
S_{\text{eff}}^{\text{bulk}}[q] = & - \frac{\sigma'_{xx}}{8} \int d\mathbf{r} \, \text{tr}(\nabla q)^2 \\
& + \frac{1}{2} \pi T z' \int d\mathbf{r} {\sum_{\alpha n}}' \text{tr}[I_n^\alpha, q][I_{-n}^\alpha, q].
\end{aligned}
\tag{4.2.40}
$$

This result displays $\mathcal{F}$ invariance as it should. In terms of the effective action procedure for edge excitations one can say that upon approaching the strong coupling insulating phase the bulk parts of the action generally vanish. More precisely, $S_{\text{eff}}^{\text{bulk}}[q]$ in Eq. (4.2.40) becomes invariant under the replacement of q by the following expression

$$
q = t^{-1} \Lambda t \rightarrow t^{-1} \mathcal{T}_0^{-1} \Lambda \mathcal{T}_0 t. \tag{4.2.41}
$$

Here, $\mathcal{T}_0$ stands for an arbitrary unitary matrix field that reduces to a $U(N) \times U(N)$ gauge at the edge of the system.

Coulomb interaction

If it were true that the electron gas behaves in all respects like an ordinary σ model then the effective action procedure for edge excitations would precisely follow the general expectations in quantum field theory and the discussion would be closed. There are, however, important and well known differences that generally complicate the matter. Technically speaking these differences arise as a peculiarity of the theory in the *replica limit* where the Mirmin-Wagner-Coleman theorem no longer applies. From a physical point of view these differences are a direct consequence of certain general constraints that are deeply rooted in the theory of quantum transport.

First, in as far as the Coulomb interaction problem is concerned, it is important to emphasize that there are higher dimensional terms in the action that one in general cannot ignore. Besides the singlet interaction term S_F one should include also the

Coulomb term S_U which can be written as [32]

$$S_U[Q] = -\pi T \int \sum_{\alpha n} \operatorname{tr} I_n^\alpha Q(\mathbf{r}) U(\mathbf{r}-\mathbf{r}') \operatorname{tr} I_{-n}^\alpha Q(\mathbf{r}'). \tag{4.2.42}$$

The potential U in momentum space is given by

$$U(p) = \rho/(1 + U_0(p)\rho) \tag{4.2.43}$$

where $\rho = \partial n/\partial \mu$ denotes the thermodynamic density of states and $U_0(p) \propto |p|^{-1}$ is the Coulomb potential in two dimensions. Although "irrelevant" in the context of weak coupling expansions, the Coulomb term is nevertheless important if one wishes to describe the full dynamical response of the electron gas as well as the electrodynamic $U(1)$ gauge invariance of the theory.

The most remarkable feature of the Coulomb term, however, is that it generally does not "renormalize". This means that S_U does not acquire any quantum or "radiative" corrections. This important aspect of the problem can be demonstrated, just like the statement made earlier on $z\alpha$, on the basis of general principles of transport theory. To see how these general statements complicate the idea of having "edge excitations" in the strong coupling regime we next consider an extended background field procedure that includes the effects of the Coulomb term S_U. The results, when fully written out, now become

$$\begin{aligned} S_{\text{eff}}^{\text{bulk}}[q] = & - \frac{\sigma'_{xx}}{8} \int \operatorname{tr}(\nabla q)^2 \\ & + \frac{1}{2}\pi T z' \int {\sum_{\alpha n}}' \operatorname{tr}[I_n^\alpha, q][I_{-n}^\alpha, q] \\ & - \pi T \int \sum_{\alpha n} \operatorname{tr} I_n^\alpha q(\mathbf{r}) U(\mathbf{r}-\mathbf{r}') \operatorname{tr} I_{-n}^\alpha q(\mathbf{r}'). \end{aligned} \tag{4.2.44}$$

Since the last term in Eq. (4.2.44) is length scale independent the bulk part of the action can no longer vanish as one increases the linear dimension of the system. In different words, scaling results alone no longer indicate how the interacting electron gas manages to render insensitive to boundary conditions as one enters into the strong coupling phase.

Free particles

This conflict between the scaling results of the theory and the invariance statement made on the strong coupling phase, Eq. (4.2.41), becomes only more pronounced in the problem with finite range interactions $0 < c, c' < 1$. This problem lies in the domain of attraction of the Fermi liquid line $c = c' = 0$ in which case the effective action $S_{\text{eff}}^{\text{bulk}}[q]$ can be written as

$$S_{\text{eff}}^{\text{bulk}}[q] = -\frac{\sigma'_{xx}}{8} \int d\mathbf{r} \operatorname{tr}(\nabla q)^2 + \pi T z \int d\mathbf{r} \left(4 \operatorname{tr} \eta q - 6 \operatorname{tr} \eta \Lambda\right). \tag{4.2.45}$$

The scaling behavior of the quantity σ'_{xx} in Eq. (4.2.45) is very similar to the results obtained for the Coulomb interaction problem, Eqs. (3.6.4) - (3.6.12). On the other hand, the quantity $z' = z$ now is length scale independent. This well known result for free particles is intimately related to the existence of anomalously broad density distributions (*multi-fractality*), a phenomenon that is absent in the problem with infinitely ranged interactions such as the Coulomb potential. [118]

Effective action for the edge.

It is in general not difficult to see how the result of Eq. (4.2.45) can in general be consistent with the existence of a strong coupling phase and, in particular, the invariance statement made in Eq. (4.2.41). For this purpose we consider the classical equations of motion

$$\sigma'_{xx} \nabla_j (q \nabla_j q) + 2\pi T z [\eta, q] = 0 \tag{4.2.46}$$

which must be satisfied everywhere except precisely at the edge of the system. As long as we work in the regime where σ'_{xx} is large we can neglect the second term in this equation and proceed by treating all the terms in $S_{\text{eff}}^{\text{bulk}}[q]$ that are linear and higher order in T as a perturbation. The solutions to Eq. (4.2.46) are then quite generally determined by the matrix field variable q at the edge of the system which therefore can be taken as independent degrees of freedom. As one approaches the strong coupling phase, however, the value of σ'_{xx} decreases and eventually renders exponentially small for large enough system sizes L. Under these circumstances one can no longer discard the frequency term in Eq. (4.2.46) since it now becomes the dominant one. Therefore, upon increasing the linear dimension of the system the saddle point structure of the theory changes such that eventually, as one enters into the strong coupling regime, the frequency term in Eq. (4.2.45) gets replaced by

$$\begin{aligned} \pi T z \int d\mathbf{r} \, (4 \, \text{tr} \, \eta q - 6 \, \text{tr} \, \eta \Lambda) \rightarrow & \\ & \pi T z \left(-2 \int d\mathbf{r} \, \text{tr} \, \eta \Lambda + \Delta \oint dx \, \text{tr} \, \eta q \right) \end{aligned} \tag{4.2.47}$$

where Δ is a phenomenological constant that otherwise depends on the microscopic details of the edge.

To summarize the results of the previous Sections one can say that there are generally distinctly different *mechanisms* at work by which the effective action procedure for disordered metals as defined in Eqs (4.2.19) and (4.2.24) eventually turns into a proper one dimensional action for edge excitations. It is next of interest to complete the argument and digress on the type of edge excitations that one would normally expect on the basis of the perturbative theory of localization and interaction effects alone. For example, the most general action for the edge for free electrons that is consistent with the symmetries of the problem as well as the invariance statement of Eq. (4.2.41) reads as follows (discarding constants)

$$S_{\text{eff}}[q] \quad = \quad g \oint dx \, \text{tr} (\nabla_x q)^2 + \pi T z \Delta \oint dx \, \text{tr} \, \eta q \tag{4.2.48}$$

where g, like Δ, is a phenomenological quantity. This action defines an Anderson localization problem in one dimension and indicates that the mass gap of "edge excitations" is even larger than that of "bulk excitations".

It is clear that these naive expectations are fundamentally modified by the presence of the $\theta(\nu)$ term in the action which is invisible in perturbative expansions.

Armed with the insights obtained from the perturbative renormalization group we next embark - for the remainder of this paper - on the problem of instantons.

4.3 Instantons

In this Section we recapitulate the instanton analysis for the Grassmannian non-linear σ model (Sections 4.3.1 and 2.4). Following the methodology of Chapter 2 we use the spatially varying masses which essentially adapts the interaction part of the action S_F to the metric of a sphere (Section 4.3.3). In Section 4.3.4 we derive the complete action for the small oscillator problem that will be used as a starting point for the remainder of this paper.

4.3.1 Introduction

The action S_σ

On the basis of the Polyakov-Schwartz inequality [20]

$$\frac{1}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla Q)^2 \geqslant 2\pi|\mathcal{C}[Q]| \tag{4.3.1}$$

one can construct stable matrix field configurations (instantons) for each of the discrete topological sectors labelled by the integer $\mathcal{C}[Q]$. The classical action S_σ is finite

$$S_\sigma^{\mathrm{inst}} = -2\pi\sigma_{xx}|\mathcal{C}[Q]| + i\theta\mathcal{C}[Q]. \tag{4.3.2}$$

The single instanton configuration with the topological charge $\mathcal{C}[Q] = \pm 1$ which is of interest to us can be represented as follows [20, 31]

$$Q_{\mathrm{inst}}(\mathbf{r}) = \mathcal{T}_0^{-1}\Lambda_{\mathrm{inst}}(\mathbf{r})\mathcal{T}_0, \qquad \Lambda_{\mathrm{inst}}(\mathbf{r}) = \Lambda + \rho(\mathbf{r}). \tag{4.3.3}$$

Here, the matrix $\rho_{nm}^{\alpha\beta}(\mathbf{r})$ has four non-zero matrix elements only

$$\begin{aligned} \rho_{00}^{11} &= -\rho_{-1-1}^{11} = -\frac{2\lambda^2}{|z-z_0|^2+\lambda^2}, \\ \rho_{0-1}^{11} &= \bar{\rho}_{-10}^{11} = \frac{2\lambda(z-z_0)}{|z-z_0|^2+\lambda^2}, \end{aligned} \tag{4.3.4}$$

with $z = x + iy$. The manifold of instanton parameters consists the quantity z_0 denoting the *position* of the instanton, the parameter λ which equals the *scale size* as well as the global unitary rotation $\mathcal{T}_0$ which describes the *orientation* in the coset space $U(2N)/U(N)\times U(N)$. These parameters do not change the value of the classical action S_σ^{inst}. The *anti-instanton* with $\mathcal{C}[Q] = -1$ is simply obtained by complex conjugation.

The action S_F

In the presence of mass terms like the *singlet interaction* term S_F the idea of stable topologically non-trivial field configurations becomes generally more complicated. The minimum action requirement, for example, immediately tells us that the global matrix $\mathcal{T}_0$ is now restricted to run over the subgroup $U(N) \times U(N)$ only. Instead of Eq. (4.3.3) we therefore write

$$Q_{\text{inst}}(\mathbf{r}) = U^{-1}\Lambda_{\text{inst}}(\mathbf{r})U = \Lambda + U^{-1}\rho(\mathbf{r})U \tag{4.3.5}$$

with $U \in U(N) \times U(N)$. Next, by substituting Eq. (4.3.5) into Eq.(4.2.5) one can split S_F into a topologically trivial piece and an instanton peace as follows

$$S_F[Q_{\text{inst}}] = S_F[\Lambda] + S_F^{\text{inst}}, \tag{4.3.6}$$

where

$$S_F[\Lambda] = -2\pi Tz \int d\mathbf{r} \operatorname{tr} \eta\Lambda, \tag{4.3.7}$$

and

$$S_F^{\text{inst}}[U] = \pi Tz \int d\mathbf{r} \Big[c \sum_{\alpha n} \operatorname{tr} I_n^\alpha U^{-1}\rho U \operatorname{tr} I_{-n}^\alpha U^{-1}\rho U + 4 \operatorname{tr} \eta U^{-1}\rho U \Big]. \tag{4.3.8}$$

Similarly we can write the classical contribution to the thermodynamic potential as the sum of two pieces

$$\Omega^{\text{class}} = \Omega_0^{\text{class}} + \Omega_{\text{inst}}^{\text{class}}, \tag{4.3.9}$$

where Ω_0^{class} is the contribution of the trivial vacuum

$$\Omega_0^{\text{class}} = S_F[\Lambda] = -2\pi Tz \int d\mathbf{r} \operatorname{tr} \eta\Lambda \tag{4.3.10}$$

and Ω_{inst} is the instanton piece

$$\Omega_{\text{inst}}^{\text{class}} = \int_{\text{inst}} \exp\left(-2\pi\sigma_{xx} \pm i\theta + S_F^{\text{inst}}[U]\right). \tag{4.3.11}$$

The subscript "inst" indicates that the integral is over the manifold of instanton parameters z_0, λ and U.

One of the main complications next is that the action $S_F^{\text{inst}}[U]$ is not finite but, rather, it diverges logarithmically in the size of the system. Although these and other complications associated with mass terms are quite well known, the resolution that has been proposed is formal at best and useless for practical purposes. [45] There is in this respect a true advantage to be gained if one follows up on the idea of *spatially varying masses* which has been introduced and analyzed in great detail in Chapter 2. This methodology not only extends the formalism developed for the massless theory in a natural fashion, but also lends itself to a non-perturbative analysis of the renormalization group β and γ functions of the theory. Before embarking on the specific problem of the interacting electron gas it is necessary to first recapitulate some of the main results obtained in Chapter 2 for the ordinary Grassmannian manifold. This will be done in the Sections below where we generalize the harmonic oscillator problem to include an arbitrary range of Matsubara frequencies. The most important results are written in Tables 4.3 and 4.4 which contains the complete action of quantum fluctuations about the single instanton.

4.3.2 Quantum fluctuations

Preliminaries

To obtain the most general matrix field variable Q with topological charge equal to unity we first rewrite the instanton solution Λ_{inst} in Eqs. (4.3.4) and (4.3.5) as a unitary rotation R about the trivial vacuum Λ

$$\Lambda_{\text{inst}} = R^{-1}\Lambda R. \tag{4.3.12}$$

From now onward we use the following notation for an arbitrary matrix A

$$A^{\alpha\beta}_{mn} = \begin{pmatrix} A^{\alpha\beta}_{n_1 n_3} & A^{\alpha\beta}_{n_1 n_2} \\ & \\ A^{\alpha\beta}_{n_2 n_1} & A^{\alpha\beta}_{n_2 n_4} \end{pmatrix}. \tag{4.3.13}$$

Here, the n_i with *odd* subscripts i denote the indices for *positive* Matsubara frequencies. Similarly, the *even* subscripts i refer to the *negative* Matsubara frequencies. Hence, the indices n_1 and n_3 run over the set of *non-negative* integers $\{0, 1, 2, \dots\}$. The indices n_2 and n_4 run over the set of *negative* integers $\{-1, -2, -3, \dots\}$. Fully written out the different frequency blocks of the unitary matrix $R^{\alpha\beta}_{mn}$ now become

$$R^{\alpha\beta}_{n_1 n_3} = \delta^{\alpha\beta}\delta_{n_1 n_3}\left[1 + (\bar{e}_1 - 1)\delta^{\alpha 1}\delta_{n_1,0}\right] \tag{4.3.14}$$

$$R^{\alpha\beta}_{n_2 n_4} = \delta^{\alpha\beta}\delta_{n_2 n_4}\left[1 + (e_1 - 1)\delta^{\alpha 1}\delta_{n_2,-1}\right] \tag{4.3.15}$$

$$R^{\alpha\beta}_{n_1 n_2} = \delta^{\alpha\beta}\delta^{\alpha 1}\delta_{n_1,0}\delta_{n_2,-1}\left[e_0\right] \tag{4.3.16}$$

$$R^{\alpha\beta}_{n_2 n_1} = \delta^{\alpha\beta}\delta^{\alpha 1}\delta_{n_1,0}\delta_{n_2,-1}\left[-e_0\right] = -R^{\alpha\beta}_{n_1 n_2} \tag{4.3.17}$$

where the quantities e_0 and e_1 are defined by

$$e_0 = \frac{\lambda}{\sqrt{|z - z_0|^2 + \lambda^2}} \tag{4.3.18}$$

$$e_1 = \frac{z - z_0}{\sqrt{|z - z_0|^2 + \lambda^2}}. \tag{4.3.19}$$

The structure of the matrix $R^{\alpha\beta}_{mn}$ is illustrated in Fig. 4.2. It is a simple matter next to generalize Eq. (4.3.12) and the result is

$$Q = \mathcal{T}_0^{-1} R^{-1} \mathcal{V}\, R\, \mathcal{T}_0. \tag{4.3.20}$$

Here, $\mathcal{T}_0$ denotes a global $U(2N)$ rotation and the matrix $\mathcal{V}$ with $\mathcal{V}^2 = \mathbf{1}$ represents the small fluctuations about the one instanton. Write

$$\mathcal{V} = w + \Lambda\sqrt{1 - w^2} \tag{4.3.21}$$

with

$$w = \begin{pmatrix} 0 & v \\ v^\dagger & 0 \end{pmatrix} \tag{4.3.22}$$

then the matrix $\mathcal{V}$ can formally be written as a series expansion in powers of the $N \times N$ complex matrices v, $v^\dagger$ which are taken as the independent field variables in the problem.

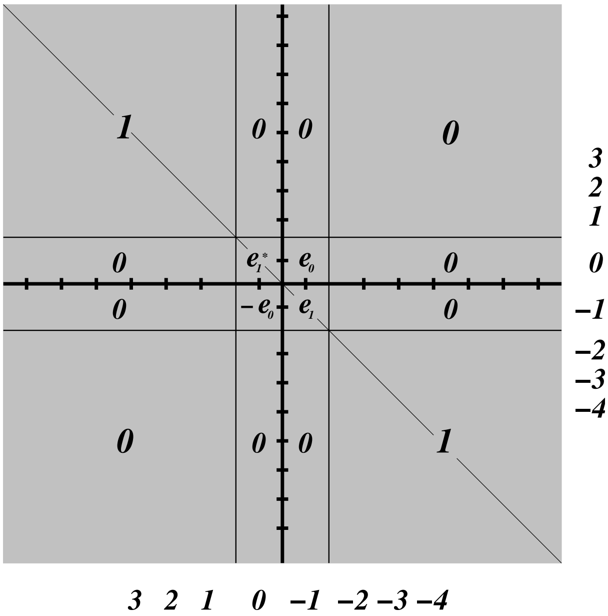


Figure 4.2: The instanton matrix R.

Stereographic projection

Eq. (4.3.20) lends itself to an exact analysis of the small oscillator problem. First we recall the results obtained for the free electron theory (Section 2.4)

$$\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}(\nabla_j Q)^2 = \frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}[\nabla_j + A_j, \mathcal{V}]^2, \tag{4.3.23}$$

where the matrix A_j contains the instanton degrees of freedom

$$A_j = R\mathcal{T}_0 \nabla_j \mathcal{T}_0^{-1} R^{-1} = R\nabla_j R^{-1}. \tag{4.3.24}$$

By expanding the $\mathcal{V}$ in Eq. (4.3.23) to quadratic order in the quantum fluctuations $v, v^\dagger$ we obtain the following results

$$\frac{\sigma_{xx}}{8}\int d\mathbf{r}\,\mathrm{tr}[\nabla_j + A_j, \mathcal{V}]^2 = \frac{\sigma_{xx}}{4}\int d\mathbf{r}\mu^2(\mathbf{r})\Bigg[\sum_{\alpha=2}^{N_r}\sum_{\beta=2}^{N_r}\sum_{n_1 n_2} v^{\alpha\beta}_{n_1 n_2} O^{(0)} v^{\dagger\beta\alpha}_{n_2 n_1}$$
$$+\sum_{\alpha=2}^{N_r}\Bigg(\sum_{n_1 n_2}{}'' v^{1\alpha}_{n_1 n_2} O^{(0)} v^{\dagger\alpha 1}_{n_2 n_1} + \sum_{n_1 n_2}{}'' v^{\alpha 1}_{n_1 n_2} O^{(0)} v^{\dagger 1\alpha}_{n_2 n_1} + \sum_{n_1}{}' v^{1\alpha}_{n_1,-1} O^{(0)} v^{\dagger\alpha 1}_{-1,n_1}$$
$$+\sum_{n_2}{}' v^{\alpha 1}_{0,n_2} O^{(0)} v^{\dagger 1\alpha}_{n_2,0} + \sum_{n_1} v^{\alpha 1}_{n_1,-1} O^{(1)} v^{\dagger 1\alpha}_{-1,n_1} + \sum_{n_2} v^{1\alpha}_{0,n_2} O^{(1)} v^{\dagger\alpha 1}_{n_2,0}\Bigg)$$
$$+\sum_{n_1 n_2}{}'' v^{11}_{n_1 n_2} O^{(0)} v^{\dagger 11}_{n_2 n_1} + \sum_{n_1}{}' v^{11}_{n_1,-1} O^{(1)} v^{\dagger 11}_{-1,n_1} + \sum_{n_2}{}' v^{11}_{0,n_2} O^{(1)} v^{\dagger 11}_{n_2,0}$$
$$+ v^{11}_{0,-1} O^{(2)} v^{\dagger 11}_{-1,0}\Bigg] \tag{4.3.25}$$

The "prime" on the summation signs are defined as follows

$$\sum_{n_1}{}' = \sum_{n_1=1}^{N_m-1}, \quad \sum_{n_2}{}' = \sum_{n_2=-2}^{-N_m}. \tag{4.3.26}$$

The three different operators $O^{(a)}$ with $a = 0, 1, 2$ are given as

$$O^{(a)} = \frac{(r^2+\lambda^2)^2}{4\lambda^2}\left[\nabla_j + \frac{ia}{r^2+\lambda^2}\varepsilon_{jk} r_k\right]^2 + \frac{a}{2}. \tag{4.3.27}$$

The introduction of a measure $\mu^2(\mathbf{r})$ for the spatial integration in Eq. (4.3.25),

$$\mu(r) = \frac{2\lambda}{r^2+\lambda^2}, \tag{4.3.28}$$

indicates that the quantum fluctuation problem is naturally defined on a sphere with radius λ. It is convenient to employ the stereographic projection

$$\eta = \frac{r^2-\lambda^2}{r^2+\lambda^2}, \qquad -1 < \eta < 1 \tag{4.3.29}$$
$$\theta = \tan^{-1}\frac{y}{x}, \qquad 0 \leqslant \theta < 2\pi. \tag{4.3.30}$$

In terms of η, θ the integration can be written as

$$\int d\mathbf{r}\mu^2(r) = \int d\eta d\theta. \tag{4.3.31}$$

Moreover,

$$e_0 = \sqrt{\frac{1-\eta}{2}}, \qquad e_1 = \sqrt{\frac{1+\eta}{2}}e^{i\theta}, \tag{4.3.32}$$

Table 4.1: Counting the total number of zero modes

Operator	Number of fields $v^{\alpha\beta}_{n_1 n_2}$	Degeneracy
$O^{(0)}$	$(N-1)^2$	1
$O^{(1)}$	$2(N-1)$	2
$O^{(2)}$	1	3

and the operators become

$$O^{(a)} = \frac{\partial}{\partial\eta}\left[(1-\eta^2)\frac{\partial}{\partial\eta}\right] + \frac{1}{1-\eta^2}\frac{\partial^2}{\partial^2\theta} - \frac{ia}{1-\eta}\frac{\partial}{\partial\theta} - \frac{a^2}{4}\frac{1+\eta}{1-\eta} + \frac{a}{2}, \tag{4.3.33}$$

with $a = 0, 1, 2$. Finally, using Eq.(4.3.25) we can count the total number of fields $v^{\alpha\beta}$ on which each of the operators $O^{(a)}$ act. The results are listed in Table 4.1.

Energy spectrum

We are interested in the eigenvalue problem

$$O^{(a)}\Phi^{(a)}(\eta,\theta) = E^{(a)}\Phi^{(a)}(\eta,\theta), \tag{4.3.34}$$

where the set of eigenfunctions $\Phi^{(a)}$ are taken to be orthonormal with respect to the scalar product

$$(\bar{\Phi}_1^{(a)}, \Phi_2^{(a)}) = \int d\eta d\theta\, \bar{\Phi}_1^{(a)}(\eta,\theta)\Phi_2^{(a)}(\eta,\theta). \tag{4.3.35}$$

The Hilbert space of square integrable eigenfunctions is given in terms of Jacobi polynomials,

$$P_n^{\alpha,\beta}(\eta) = \frac{(-1)^n}{2^n n!}(1-\eta)^{-\alpha}(1+\eta)^{-\beta}\frac{d^n}{d\eta^n}(1-\eta)^{n+\alpha}(1+\eta)^{n+\beta}. \tag{4.3.36}$$

Introducing the quantum number J to denote the discrete energy levels

$$\begin{aligned} E_J^{(0)} &= J(J+1), & J &= 0,1,\dots \\ E_J^{(1)} &= (J-1)(J+1), & J &= 1,2,\dots \\ E_J^{(2)} &= (J-1)(J+2), & J &= 1,2,\dots \end{aligned} \tag{4.3.37}$$

then the eigenfunctions are labelled by (J, M) and can be written as follows

$$\begin{aligned} \Phi_{J,M}^{(0)} &= C_{J,M}^{(0)} e^{iM\theta}\sqrt{(1-\eta^2)^M} P_{J-M}^{M,M}(\eta), & M &= -J,\cdots,J \\ \Phi_{J,M}^{(1)} &= C_{J,M}^{(1)} e^{iM\theta}\sqrt{(1-\eta^2)^M}\sqrt{1-\eta} P_{J-M-1}^{M+1,M}(\eta), & M &= -J,\cdots,J-1 \\ \Phi_{J,M}^{(2)} &= C_{J,M}^{(2)} e^{iM\theta}\sqrt{(1-\eta^2)^M}(1-\eta) P_{J-M-1}^{M+2,M}(\eta), & M &= -J-1,\cdots,J-1 \end{aligned} \tag{4.3.38}$$

where the normalization constants equal

$$
\begin{aligned}
C^{(0)}_{J,M} &= \frac{\sqrt{\Gamma(J-M+1)\Gamma(J+M+1)(2J+1)}}{2^{M+1}\sqrt{\pi}\Gamma(J+1)}, \\
C^{(1)}_{J,M} &= \frac{\sqrt{\Gamma(J-M)\Gamma(J+M+1)}}{2^{M+1}\sqrt{\pi}\Gamma(J)}, \\
C^{(2)}_{J,M} &= \frac{\sqrt{\Gamma(J-M)\Gamma(J+M+2)(2J+1)}}{2^{M+2}\sqrt{\pi}\Gamma(J)\sqrt{J(J+1)}}.
\end{aligned}
\tag{4.3.39}
$$

Zero modes

From Eq. (4.3.37) we see that the operators $O^{(0)}$ has a zero frequency mode $E^{(0)}_J = 0$ for $J = 0$. Similarly, we have $E^{(1)}_J = E^{(2)}_J = 0$ for $J = 1$. The corresponding eigenfunctions can be written as follows

$$
\begin{aligned}
O^{(0)} &\Longrightarrow \Phi^{(0)}_{0,0} = 1, \\
O^{(1)} &\Longrightarrow \Phi^{(1)}_{1,-1} = \frac{1}{\sqrt{2\pi}}\bar{e}_1, \qquad \Phi^{(1)}_{1,0} = \frac{1}{\sqrt{2\pi}}e_0, \\
O^{(2)} &\Longrightarrow \Phi^{(2)}_{1,-2} = \sqrt{\frac{3}{4\pi}}\bar{e}_1^2, \qquad \Phi^{(2)}_{1,-1} = \sqrt{\frac{3}{2\pi}}e_0\bar{e}_1, \\
&\qquad\;\;\; \Phi^{(2)}_{1,0} = \sqrt{\frac{3}{4\pi}}e_0^2.
\end{aligned}
\tag{4.3.40}
$$

Here, the quantities e_0 and e_1 are defined in Eqs. (4.3.18) and (4.3.19) (see also Eq. (4.3.32)). The number of the zero modes of each $O^{(a)}$ is listed in Table 4.1. The total we find $2(N^2 + 2N)$ zero modes in the problem.

Next, it is important to show that these zero modes precisely correspond to all the instanton degrees of freedom contained in the matrices R and $\mathcal{T}_0$ of Eq. (4.3.20). For this purpose we write the instanton solution as follows

$$
Q_{\text{inst}}(\xi_i) = U^{-1}_{\text{inst}}(\xi_i)\Lambda U_{\text{inst}}(\xi_i). \tag{4.3.41}
$$

Here, $U_{\text{inst}} = R\mathcal{T}_0$ and the ξ_i stand for the parameters z_0, λ and the generators of $\mathcal{T}_0$. An infinitesimal change in the instanton parameters $\xi_i \to \xi_i + \varepsilon_i$ can be written in the form of Eq.(4.3.20) as follows

$$
Q_{\text{inst}}(\xi_i + \varepsilon_i) = U^{-1}_{\text{inst}}(\xi_i)\mathcal{V}_\varepsilon U_{\text{inst}}(\xi_i), \tag{4.3.42}
$$

where to linear order in ε_i we can write

$$
\begin{aligned}
\mathcal{V}_\varepsilon &= \Lambda - \varepsilon_i \left[U_{\text{inst}}\partial_i U^{-1}_{\text{inst}}, \Lambda\right] \\
&= \begin{pmatrix} \mathbf{1} & 2\varepsilon_i \left[U_{\text{inst}}\partial_i U^{-1}_{\text{inst}}\right]^{\alpha\beta}_{n_1 n_2} \\ -2\varepsilon_i \left[U_{\text{inst}}\partial_i U^{-1}_{\text{inst}}\right]^{\alpha\beta}_{n_2 n_1} & -\mathbf{1} \end{pmatrix}.
\end{aligned}
\tag{4.3.43}
$$

Table 4.2: Zero energy modes

$\alpha\ \beta$	$n_1\ n_2$	$O^{(0)}$	$O^{(1)}$	$O^{(2)}$
$\alpha > 1,\ \beta > 1$	$n_1 \geqslant 0,\ n_2 \leqslant -1$	$2it^{\alpha\beta}_{n_1 n_2}\Phi^{(0)}_{0,0}$		
$\alpha > 1,\ \beta = 1$	$n_1 \geqslant 0,\ n_2 = -1$		$2i\sqrt{2\pi}t^{\alpha 1}_{n_1,-1}\Phi^{(1)}_{1,-1}$ $-2i\sqrt{2\pi}t^{\alpha 1}_{n_1,0}\Phi^{(1)}_{1,0}$	
	$n_1 \geqslant 0,\ n_2 < -1$	$2it^{\alpha 1}_{n_1 n_2}\Phi^{(0)}_{0,0}$		
$\alpha = 1,\ \beta > 1$	$n_1 > 0,\ n_2 \leqslant -1$	$2it^{1\beta}_{n_1 n_2}\Phi^{(0)}_{0,0}$		
	$n_1 = 0,\ n_2 \leqslant -1$		$2i\sqrt{2\pi}t^{1\beta}_{0,n_2}\Phi^{(1)}_{1,-1}$ $+2i\sqrt{2\pi}t^{1\beta}_{-1,n_2}\Phi^{(1)}_{1,0}$	
$\alpha = 1,\ \beta = 1$	$n_1 > 0,\ n_2 < -1$	$2it^{\alpha\beta}_{n_1 n_2}\Phi^{(0)}_{0,0}$		
	$n_1 = 0,\ n_2 < -1$		$2i\sqrt{2\pi}t^{11}_{0,n_2}\Phi^{(1)}_{1,-1}$ $+2i\sqrt{2\pi}t^{11}_{-1,n_2}\Phi^{(1)}_{1,0}$	
	$n_1 > 0,\ n_2 = -1$		$2i\sqrt{2\pi}t^{11}_{n_1,-1}\Phi^{(1)}_{1,-1}$ $-2i\sqrt{2\pi}t^{11}_{n_1,0}\Phi^{(1)}_{1,0}$	
	$n_1 = 0,\ n_2 = -1$			$4i\frac{\sqrt{\pi}}{\sqrt{3}}t^{11}_{-1,0}\Phi^{(2)}_{1,-2}$ $+4i\frac{\sqrt{\pi}}{\sqrt{6}}t^{11}_{-1,-1}\Phi^{(2)}_{1,-1}$ $-4i\frac{\sqrt{\pi}}{\sqrt{6}}t^{11}_{0,0}\Phi^{(2)}_{1,-1}$ $-4\frac{\sqrt{\pi}}{\sqrt{6}}\frac{\delta\lambda}{\lambda}\Phi^{(2)}_{1,-1}$ $-4i\frac{\sqrt{\pi}}{\sqrt{3}}t^{11}_{0,-1}\Phi^{(2)}_{1,0}$ $+4i\frac{\sqrt{\pi}}{\sqrt{3}}\frac{\delta\bar{z}_0}{\lambda}\Phi^{(2)}_{1,0}$

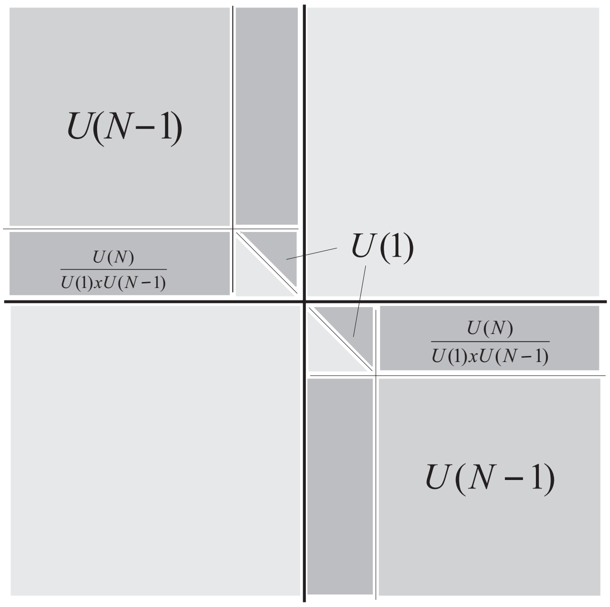


Figure 4.3: The hierarchy of symmetry breaking by the instanton solution

We have written $\partial_i = \partial/\partial\xi_i$. By comparing this expression with Eq.(4.3.20) we see that the fluctuations tangential to the instanton manifold can be expressed in terms of the matrix field variables $v, v^\dagger$ according to

$$v^{\alpha\beta}_{n_1 n_2} = 2\varepsilon_i \left[U_{\rm inst}\partial_i U^{-1}_{\rm inst}\right]^{\alpha\beta}_{n_1 n_2}, \quad (4.3.44)$$

$$\left[v^\dagger\right]^{\alpha\beta}_{n_2 n_1} = -2\varepsilon_i \left[U_{\rm inst}\partial_i U^{-1}_{\rm inst}\right]^{\alpha\beta}_{n_2 n_1}. \quad (4.3.45)$$

To obtain explicit expressions it suffices to expand $\mathcal{T}_0$ about unity

$$\mathcal{T}_0 = 1 + i\,t \quad (4.3.46)$$

and write

$$R(\lambda + \delta\lambda, z_0 + \delta z_0) = R(\lambda, z_0) + \delta\lambda\ \partial_\lambda R + \delta z_0\ \partial_{z_0} R. \quad (4.3.47)$$

The expression for v now becomes

$$v^{\alpha\beta}_{n_1 n_2} = 2i\left[RtR^{-1}\right]^{\alpha\beta}_{n_1 n_2} + 2\delta\lambda\left[R\partial_\lambda R^{-1}\right]^{\alpha\beta}_{n_1 n_2} + 2\delta z_0\left[R\partial_{z_0}R^{-1}\right]^{\alpha\beta}_{n_1 n_2}. \quad (4.3.48)$$

Notice that $v^\dagger$ is just the hermitian conjugate of v as it should be. In Table 4.2 we present the complete list of zero energy modes $v^{\alpha\beta}_{n_1 n_2}$ written in terms of $t^{\alpha\beta}_{mn}$, $\delta\lambda$ and δz_0 as well as the eigenfunctions $\Phi^{(a)}_{JM}$.

In these expressions $t^{\alpha\beta}_{n_1 n_2}$ and $t^{\alpha\beta}_{n_2 n_1}$ denote the generators of $U(2N)/U(N)\times U(N)$. The $t^{\alpha 1}_{n_1,0}$ and $t^{1\alpha}_{0,n_1}$ with $n_1 \neq 0$ and $\alpha = 1$ are the generators of a $U(N)/U(N-1)\times$

$U(1)$ rotation. The same holds for $t^{\alpha 1}_{n_2,-1}$ and $t^{1\alpha}_{-1,n_2}$ with $n_2 \neq -1$ and $\alpha = 1$. Finally, $t^{11}_{0,0} - t^{11}_{-1,-1}$ denotes the $U(1)$ generator corresponding to rotations of the $O(3)$ instanton in the xy plane. The number of instanton degrees of freedom adds up to $2(N^2 + 2N)$ which is that same as the number of zero modes in the problem. The various different generators t of the instanton manifold is illustrated in Fig. 4.3.

4.3.3 Spatially varying masses

In the previous Section we have seen that the instanton problem naturally acquires the geometry of a *sphere*. This clearly complicates the problem of mass terms in the theory which are usually written in *flat* space. To deal with this problem we shall follow the root established in Chapter 2 and shall modify the definition of the singlet interaction term and introduce a spatially varying momentum scale $\mu(\mathbf{r})$ as follows

$$z \to z\mu^2(\mathbf{r}), \qquad zc \to zc\mu^2(\mathbf{r}), \tag{4.3.49}$$

such that the action S_F is now *finite* and can be written as

$$S_F[Q] \to \pi T z \int d\mathbf{r} \mu^2(\mathbf{r}) \Big(c \sum_{\alpha n} \operatorname{tr} I^\alpha_n Q \operatorname{tr} I^\alpha_{-n} Q + 4 \operatorname{tr} \eta Q - 6 \operatorname{tr} \eta \Lambda \Big). \tag{4.3.50}$$

As we shall show below, in Sections 4.4 and 4.5.1, the introduction of a spatially varying momentum scale $\mu(\mathbf{r})$ permits the development of a complete quantum theory of the interacting electron gas that is defined on a sphere. Although the philosophy sofar proceeds along similar lines as those employed in the ordinary Grassmannian model in Chapter 2, it is important to keep in mind that the presence of S_F is itself affecting the ultraviolet singularity structure of the theory. This means that both the physics and the conceptual structure of the problem with interactions are fundamentally different from what one is used to. Moreover, in view of the mathematical peculiarities of the theory, in particular those associated with the limits $N_r \to 0$ and $N_m \to \infty$, it must be shown explicitly that instantons are well defined at a quantum level and that the aforementioned ultraviolet behavior of the interacting electron gas does not depend on the specific geometry that one chooses, i.e. the introduction of $\mu(\mathbf{r})$ in Eq. (4.3.50). In this respect, we shall in what follows greatly benefit from our theory of *observable parameters* since it provides the appropriate framework for a general understanding of the theory at short distances. To study the *ultraviolet* we first address the problem of quantum fluctuations for the special case where unitary matrix $\mathcal{T}_0$ in Eq. (4.3.20) is equal to unity. We will come back to the general case not until Section 4.5 where embark on the *infrared* of the system, notably the various different steps that are needed in order to change the geometry of the system from *curved* space to *flat* space.

4.3.4 Action for the quantum fluctuations

Keeping the remarks of the previous Section in mind we obtain the complete action as the sum of a classical part S^{inst} and a quantum part δS as follows

$$S = S_F[\Lambda] + S^{\text{inst}} + \delta S, \tag{4.3.51}$$

where

$$S^{\text{inst}} = -2\pi\sigma_{xx} + i\theta + S_F^{\text{inst}}. \tag{4.3.52}$$

and

$$\delta S = \delta S^{(0)} + \delta S^{(1)} + \delta S^{(2)} + \delta S^{(2)}_{\text{linear}}. \tag{4.3.53}$$

Here S_F^{inst} stands for the classical action of the modified *singlet interaction* term, Eq. (4.3.8), with $W = 1$ and is given by

$$S_F^{\text{inst}} = \pi T z \int d\mathbf{r}\mu^2(\mathbf{r}) \left(c \sum_{\alpha n} \operatorname{tr} I_n^\alpha \rho \operatorname{tr} I_{-n}^\alpha \rho + 4 \operatorname{tr} \eta\rho \right) = 16\pi^2 T z \Big(\frac{c}{3} - 1 \Big). \tag{4.3.54}$$

Next, the results for δS in Eq. (4.3.53) are classified in four different parts. The complete list of contributions is presented in Tables 4.3 and 4.4. As before we shall use the following notations $n_{12} = n_1 - n_2$ and $\kappa^2 = 8\pi T/\sigma_{xx}$ from now onward. We shall first briefly comment on the different parts of δS.

$\delta S^{(0)}$

This term contains all the fluctuations $v_{mn}^{\alpha\beta}$ with replica indices $\alpha, \beta > 1$ that do not couple to the instanton. $\delta S^{(0)}$ has therefore the same form as the fluctuations about the trivial vacuum.

$\delta S^{(1)}$, $\delta S^{(2)}$

The terms $\delta S^{(1)}$ and $\delta S^{(2)}$ contain all the fluctuations $v_{mn}^{\alpha\beta}$ with either $\alpha = 1$ or $\beta = 1$. $\delta S^{(2)}$ only contains the fluctuations in the first replica channel v_{mn}^{11} and the remaining contributions are collected in $\delta S^{(1)}$. In both $\delta S^{(1)}$ and $\delta S^{(2)}$ we distinguish between the "diagonal" contributions that mainly originate from S_σ (first four lines in Tables 4.3 and 4.4) and the "off-diagonal" ones originating from S_F (fifth and subsequent lines).

$\delta S^{(2)}_{\text{linear}}$

The contributions linear in v and $v^\dagger$ originate from the singlet interaction term S_F and are written in the bottom line of Table 4.3. They can be written in terms of the eigenfunctions $\Phi_{JM}^{(a)}$ as follows

$$\int d\eta d\theta \left(e_0^2 \bar{e}_1 v_{0,-2}^{11} + e_0^2 e_1 v_{-2,0}^{\dagger 11} \right) \;\propto\; \int d\eta d\theta \Big(\bar{\Phi}_{2,1}^{(1)} v_{0,-2}^{11} + \Phi_{2,1}^{(1)} v_{-2,0}^{\dagger 11} \Big), \tag{4.3.55}$$

$$\int d\eta d\theta \left(e_0^2 \bar{e}_1 v_{1,-1}^{11} + e_0^2 e_1 v_{-1,1}^{11} \right) \;\propto\; \int d\eta d\theta \Big(\bar{\Phi}_{2,1}^{(1)} v_{1,-1}^{11} + \Phi_{2,1}^{(1)} v_{-1,1}^{\dagger 11} \Big), \tag{4.3.56}$$

$$\int d\eta d\theta \left(e_0^3 \bar{e}_1 v_{0,-1}^{11} + e_0^3 e_1 v_{-1,0}^{\dagger 11} \right) \;\propto\; \int d\eta d\theta \Big(\bar{\Phi}_{2,1}^{(2)} v_{0,-1}^{11} + \Phi_{2,1}^{(2)} v_{-1,0}^{\dagger 11} \Big). \tag{4.3.57}$$

Since the $\Phi_{2,1}^{(1)}$ and $\Phi_{2,1}^{(2)}$ do not correspond to the zero modes of the operators $O^{(1)}$ and $O^{(2)}$ one can eliminate these terms by performing a simple shift in v, $v^\dagger$. This

Table 4.3: Quantum fluctuations on the Gaussian level. The $\delta S^{(0)}$, $\delta S^{(1)}$ and $\delta S^{(2)}_{\rm linear}$ contributions.

$$\begin{aligned}
\delta S^{(0)} = & - \tfrac{\sigma_{xx}}{4}\int d\eta d\theta \sum_{\alpha,\beta=2}^{N_r}\sum_{n_1\cdots n_4}\delta_{n_{12},n_{34}} \\
& \times v^{\alpha\beta}_{n_1n_2}\Big[(O^{(0)}+\kappa^2 z n_{12})\delta_{n_1n_3} - \kappa^2 zc\delta^{\alpha\beta}\Big]v^{\dagger\beta\alpha}_{n_4n_3}
\end{aligned}$$

$$\begin{aligned}
\delta S^{(1)} = & - \tfrac{\sigma_{xx}}{4}\int d\eta d\theta \sum_{\alpha=2}^{N_r}\Bigg\{\sum_{n_1n_2}{}'' v^{1\alpha}_{n_1n_2}(O^{(0)}+\kappa^2 z n_{12})v^{\dagger\alpha 1}_{n_2n_1} \\
& + \sum_{n_1n_2}{}'' v^{\alpha 1}_{n_1n_2}(O^{(0)}+\kappa^2 z n_{12})v^{\dagger 1\alpha}_{n_2n_1} \\
& + \sum_{n_1}{}' v^{1\alpha}_{n_1,-1}(O^{(0)}+\kappa^2 z(n_1+1))v^{\dagger\alpha 1}_{-1,n_1} + \sum_{n_2}{}' v^{\alpha 1}_{0n_2}(O^{(0)}-\kappa^2 z n_2)v^{\dagger 1\alpha}_{n_20} \\
& + \sum_{n_1} v^{\alpha 1}_{n_1,-1}\Big(O^{(1)}+\kappa^2 z(n_1+1)+\kappa^2 zce_0^2\Big(2|e_1|^2-\tfrac{1}{c}\Big)\Big)v^{\dagger 1\alpha}_{-1,n_1} \\
& + \sum_{n_2} v^{1\alpha}_{0n_2}\Big(O^{(1)}-\kappa^2 z n_2+\kappa^2 zce_0^2\Big(2|e_1|^2-\tfrac{1}{c}\Big)\Big)v^{\dagger\alpha 1}_{n_20}\Bigg\} \\
& - \tfrac{\sigma_{xx}}{4}\kappa^2 zc\int d\eta d\theta \sum_{\alpha=2}^{N_r}\Bigg\{\sum_{n_1,n_2}{}'' e_0\Big[\bar{e}_1 v^{1\alpha}_{n_1+1,n_2}v^{\dagger\alpha 1}_{n_2n_1} + e_1 v^{1\alpha}_{n_1n_2}v^{\dagger\alpha 1}_{n_2,n_1+1} \\
& - \bar{e}_1 v^{\alpha 1}_{n_1,n_2-1}v^{\dagger 1\alpha}_{n_2n_1} - e_1 v^{\alpha 1}_{n_1n_2}v^{\dagger 1\alpha}_{n_2-1,n_1}\Big] \\
& + \sum_{n_1}{}'\Big[e_0\bar{e}_1 v^{1\alpha}_{n_1+1,-1}v^{\dagger\alpha 1}_{-1,n_1} + e_0 e_1 v^{1\alpha}_{n_1,-1}v^{\dagger\alpha 1}_{-1,n_1+1}\Big] \\
& - \sum_{n_2}{}'\Big[e_0\bar{e}_1 v^{\alpha 1}_{0,n_2-1}v^{\dagger 1\alpha}_{n_2,0} +_0 e_1 v^{\alpha 1}_{0,n_2}v^{\dagger 1\alpha}_{n_2-1,0}\Big] \\
& + \sum_{n_2}{}'\Big[e_0 e_1^2 v^{1\alpha}_{1,n_2}v^{\dagger\alpha 1}_{n_2,0} + e_0\bar{e}_1^2 v^{1\alpha}_{0,n_2}v^{\dagger\alpha 1}_{n_2,1}\Big] \\
& - \sum_{n_1}{}'\Big[e_0 e_1^2 v^{\alpha 1}_{n_1,-2}v^{\dagger 1\alpha}_{-1,n_1} + e_0\bar{e}_1^2 v^{\alpha 1}_{n_1,-1}v^{\dagger 1\alpha}_{-2,n_1}\Big]\Bigg\}
\end{aligned}$$

$$\begin{aligned}
\delta S^{(2)}_{\rm linear} = & \tfrac{\sigma_{xx}}{2}\kappa^2 zc\int d\eta d\theta\Bigg\{e_0^2(\bar{e}_1 v^{11}_{0,-2} + e_1 v^{\dagger 11}_{-2,0} - \bar{e}_1 v^{11}_{1,-1} - e_1 v^{\dagger 11}_{-1,1}) \\
& + e_0(1-2e_0^2)(\bar{e}_1 v^{11}_{0,-1} + e_1 v^{\dagger 11}_{-1,0})\Bigg\}
\end{aligned}$$

Table 4.4: Quantum fluctuations on the Gaussian level. The $\delta S^{(2)}$ contributions.

$$
\begin{aligned}
\delta S^{(2)} = \quad & - \quad \tfrac{\sigma_{xx}}{4}\int d\eta d\theta \\
& \times \quad \Big\{ \sum_{n_1\cdots n_4}{}'''' v^{11}_{n_1n_2}\Big((O^{(0)}+\kappa^2 z n_{12})\delta_{n_1n_3}\delta_{n_2n_4} - \kappa^2 zc\delta_{n_{12},n_{34}}\Big)v^{\dagger 11}_{n_4n_3} \\
& + \quad \sum_{n_1}{}' v^{11}_{n_1,-1}\Big(O^{(1)} + \kappa^2 z(n_1+1) - \kappa^2 zc + \kappa^2 zce_0^2\Big(2|e_1|^2 - \tfrac{1}{c}\Big)\Big)v^{\dagger 11}_{-1,n_1} \\
& + \quad \sum_{n_2}{}' v^{11}_{0n_2}\Big(O^{(1)} - \kappa^2 z n_2 - \kappa^2 zc + \kappa^2 zce_0^2\Big(2|e_1|^2 - \tfrac{1}{c}\Big)\Big)v^{\dagger 11}_{n_20} \\
& + \quad v^{11}_{0,-1}\Big(O^{(2)} + \kappa^2 z(1-c) + 2\kappa^2 zce_0^2\Big(3|e_1|^2 - \tfrac{1}{c}\Big)\Big)v^{\dagger 11}_{-1,0}\Big\} \\
& - \quad \tfrac{\sigma_{xx}}{4}\kappa^2 zc\int d\eta d\theta \Big\{ \sum_{n_1,n_2}{}'' \Big[e_0\bar{e}_1 v^{11}_{n_1+1,n_2}v^{\dagger 11}_{n_2n_1} + e_0 e_1 v^{11}_{n_1n_2}v^{\dagger 11}_{n_2,n_1+1} \\
& - \quad e_0\bar{e}_1 v^{11}_{n_1,n_2-1}v^{\dagger 11}_{n_2n_1} - e_0 e_1 v^{11}_{n_1n_2}v^{\dagger 11}_{n_2-1,n_1}\Big] \\
& - \quad \sum_{n_1\cdots n_3}{}''' v^{11}_{n_1n_2}\Big[\bar{e}_1\delta_{n_{12},n_3+1} - e_0\delta_{n_{12},n_3}\Big]v^{\dagger 11}_{-1,n_3} \\
& - \quad \sum_{n_1\cdots n_3}{}''' v^{11}_{n_3,-1}\Big[e_1\delta_{n_{12},n_3+1} - e_0\delta_{n_{12},n_3}\Big]v^{\dagger 11}_{n_2,n_1} \\
& - \quad \sum_{n_2\cdots n_4}{}''' v^{11}_{n_3n_2}\Big[\bar{e}_1\delta_{n_{32},-n_4} + e_0\delta_{n_{32},1-n_4}\Big]v^{\dagger 11}_{n_4,0} \\
& - \quad \sum_{n_2\cdots n_4}{}''' v^{11}_{0,n_4}\Big[e_1\delta_{n_{32},-n_4} + e_0\delta_{n_{32},1-n_4}\Big]v^{\dagger 11}_{n_2n_3} \\
& - \quad e_0\sum_{n_1}{}'\Big[\bar{e}_1^2 v^{11}_{n_1,-2}v^{\dagger 11}_{-1,n_1} + e_1^2 v^{11}_{n_1,-1}v^{\dagger 11}_{-2,n_1}\Big] \\
& + \quad e_0\sum_{n_2}{}'\Big[\bar{e}_1^2 v^{11}_{1,n_2}v^{\dagger 11}_{n_2,0} + e_1^2 v^{11}_{0,n_2}v^{\dagger 11}_{n_2,1}\Big] \\
& + \quad 2\sum_{n_1}{}'\Big[e_0\bar{e}_1 v^{11}_{n_1+1,-1}v^{\dagger 11}_{-1,n_1} + e_0 e_1 v^{11}_{n_1,-1}v^{\dagger 11}_{-1,n_1+1}\Big] \\
& - \quad 2\sum_{n_2}{}'\Big[e_0\bar{e}_1 v^{11}_{0,n_2-1}v^{\dagger 11}_{n_2,0} + e_0 e_1 v^{11}_{0,n_2}v^{\dagger 11}_{n_2-1,0}\Big] \\
& - \quad (1-2e_0^2)\sum_{n_1}{}'\Big[v^{11}_{n_1,-1}v^{\dagger 11}_{-n_1-1,0} + v^{11}_{0,-n_1-1}v^{\dagger 11}_{-1,n_1}\Big] \\
& + \quad e_0(\bar{e}_1 - e_1)\sum_{n_1}{}'\Big[v^{11}_{n_1,-1}v^{\dagger 11}_{-n_1,0} - v^{11}_{0,-n_1}v^{\dagger 11}_{-1,n_1}\Big] \\
& + \quad e_0(\bar{e}_1^2 v^{11}_{1,-1}v^{11}_{0,-1} + e_1^2 v^{\dagger 11}_{-1,1}v^{\dagger 11}_{-1,0} - \bar{e}_1^2 v^{11}_{0,-2}v^{11}_{0,-1} - e_1^2 v^{\dagger 11}_{-2,0}v^{\dagger 11}_{-1,0}) \\
& + \quad e_0^2\Big[e_1^2 v^{\dagger 11}_{-1,0}v^{\dagger 11}_{-1,0} + \bar{e}_1^2 v^{11}_{0,-1}v^{11}_{0,-1}\Big]\Big\}
\end{aligned}
$$

leads to an insignificant contribution to the classical action of the order $O(T^2)$. Next,

$$\int d\eta d\theta \Big(e_0 \bar{e}_1 v^{11}_{0,-1} + e_0 e_1 v^{\dagger 11}_{-1,0} \Big) \propto \int d\eta d\theta \Big(\bar{\Phi}^{(2)}_{1,-1} v^{11}_{0,-1} + \Phi^{(2)}_{1,-1} v^{\dagger 11}_{-1,0} \Big) \propto \frac{\delta\lambda}{\lambda}. \quad (4.3.58)$$

This means that the fluctuations tangential to the instanton parameter λ are the only unstable fluctuations in the problem. As will be discussed further below, these fluctuations will be treated separately and we will proceed by formally evaluating the quantum theory to first order in the temperature T only.

Trivial vacuum

For completeness we give the expression for the quantum fluctuations about the trivial vacuum. The result can be written as follows

$$S_0 = S_F[\Lambda] + \delta S_0, \quad (4.3.59)$$

where

$$\begin{aligned} \delta S_0 = & - \frac{\sigma_{xx}}{4} \int d\eta d\theta \sum_{\alpha,\beta=1}^{N_r} \sum_{n_1 \cdots n_4} \delta_{n_{12},n_{34}} v^{\alpha\beta}_{n_1 n_2} \\ & \times \Big[(O^{(0)} + \kappa^2 z n_{12}) \delta_{n_1 n_3} - \kappa^2 z c \delta^{\alpha\beta} \Big] v^{\dagger \beta\alpha}_{n_4 n_3}. \end{aligned} \quad (4.3.60)$$

4.4 Details of computations

In this Section we present the detailed computations of the harmonic oscillator problem. In the first part we address the thermodynamic potential which is in many ways standard. The complications primarily arise from the infinite sums over Matsubara frequencies which fundamentally alter the ultraviolet singularity structure of the theory. We set up a systematic series expansion of the thermodynamic potential in powers of the temperature T. To perform the algebra we make use of the complete set of eigenvalues and eigenfunctions obtained in the previous Section as well as certain mathematical identities that are all listed in Appendix 2.A. In the second part of this Section we show that the ultraviolet singularity structure of the small oscillator problem is identically the same as the one computed on the basis of the theory of observable parameters. These important computations and results, which are briefly summarized in Appendix 4.A, permit one to proceed in an unambiguous manner and develop - in the remaining part of this paper - a non-perturbative analysis of the observable quantities of the theory.

4.4.1 Pauli-Villars regulators

Introduction

Recall that after integration over the quantum fluctuations one is in general left with two sources of divergences. First, there are the ultraviolet singularities which eventually result in a renormalization of the coupling constant or σ_{xx}. At present we wish to extend the analysis to include the renormalization of the z and zc fields. The ultraviolet of the theory can be dealt with in a standard manner by employing Pauli-Villars regulator fields with masses $\mathcal{M}_f$ $(f = 0, 1, \ldots, K)$ and an alternating metric e_f. [94, 44, 72, 20] We assume $e_0 = 1$, $\mathcal{M}_0 = 0$ and large masses $\mathcal{M}_f \gg 1$ for $f > 1$. The following constraints are imposed

$$\sum_{f=0}^{K} e_f \mathcal{M}_f^k = 0, \qquad 0 \leqslant k < K, \tag{4.4.1}$$

$$\sum_{f=1}^{K} e_f \ln \mathcal{M}_f = -\ln \mathcal{M}. \tag{4.4.2}$$

The regularized theory is then defined as

$$\delta S_{\text{reg}} = \delta S_0 + \sum_{f=1}^{K} e_f \delta S_f. \tag{4.4.3}$$

Here, action δS_f is the same as δS except that the operators $O^{(a)}$ are all replaced by $O^{(a)} + \mathcal{M}_f^2$. Our task is to evaluate Eq. (4.4.3) to lowest orders in a series expansion in powers of T. This expansion still formally diverges due to the zero modes of the operators $O^{(a)}$. These zero modes, however, shall be treated separately by employing the collective coordinate formalism introduced in Ref. [20].

To simplify the notation we will next present the results while omitting the alternating metric and the Pauli-Villars masses. This can be done since in each case we consider one easily recognizes how the metric and masses should be included. Consider the ratio

$$\frac{Z_{\text{inst}}}{Z_0} = \frac{\int \mathcal{D}[v, v^\dagger] \exp S}{\int \mathcal{D}[v, v^\dagger] \exp S_0} = \exp\Big[-2\pi\sigma_{xx} + i\theta + S_F^{\text{inst}} + \Delta S_\sigma + \Delta S_F\Big]. \tag{4.4.4}$$

Here, the quantum corrections denoted by ΔS_σ and ΔS_F can be expressed in terms of the propagators

$$\mathcal{G}_a(\omega) = \frac{1}{O^{(a)} + \omega} = \sum_{JM} \frac{|JM\rangle_{(a)(a)}\langle JM|}{E_J^{(a)} + \omega}, \tag{4.4.5}$$

$$\mathcal{G}_a^c(\omega) = \frac{1}{O^{(a)} + \alpha\omega} = \sum_{JM} \frac{|JM\rangle_{(a)(a)}\langle JM|}{E_J^{(a)} + \alpha\omega}, \tag{4.4.6}$$

where $a = 0, 1, 2$. These expressions are directly analogous to those that appear in flat space (see Eq. (3.4.4)). It is important to emphasize that even at a Gaussian level the integration over the field variables $v, v^\dagger$ in Eq. (4.4.4) is not simple and straight forward. The main reason is that some of the frequency sums can be written as an integral in the limit $T \to 0$ and, along with that, they absorb a factor of T. It is therefore not always obvious how the series expansion in powers of T should be evaluated. The simplest way to proceed is to expand the functional integrals of Eq. (4.4.4) in non-diagonal elements which are proportional to $\kappa^2 \sim T$. By inspection one can then convince oneself that in the replica limit $N_r \to 0$, the expansion in the non-diagonal terms can be truncated beyond third order only. We shall next summarize the various different contributions to ΔS_σ as well as ΔS_F.

ΔS_σ

The quantum correction ΔS_σ is obtained by expanding the non-diagonal terms of Tables 4.3 and 4.4 up to the second order. The results in the limit $T \to 0$, $N_r \to 0$ and $N_m \to \infty$ can be written as follows

$$\begin{aligned}\Delta S_\sigma &= 2\,\mathrm{tr}[\ln \mathcal{G}_1(0) - \ln \mathcal{G}_0(0)] - \mathrm{tr}[\ln \mathcal{G}_2(0) - \ln \mathcal{G}_0(0)] \\ &+ 2c \int_0^\infty d\omega\, \mathrm{tr}[\mathcal{G}_1(\omega) - \mathcal{G}_0(\omega)] \qquad (4.4.7) \\ &+ 2c^2 \int_0^\infty d\omega\, \omega\, \mathrm{tr}[\bar{e}_1 \mathcal{G}_0^c(\omega) e_1 \mathcal{G}_1(\omega) + e_0 \mathcal{G}_0^c(\omega) e_0 \mathcal{G}_1(\omega) - \mathcal{G}_0^c(\omega)\mathcal{G}_0(\omega)]. \qquad (4.4.8)\end{aligned}$$

In these expressions the trace is taken with respect to the complete set of eigenfunctions of the operators $O^{(a)}$. To evaluate these expressions we need the help of the identities (2.A.4) and (2.A.7) of Appendix 2.A. After elementary algebra we obtain

$$\Delta S_\sigma = 2\alpha D^{(1)} - D^{(2)} - 2c\Big(H^{(1)} \ln \alpha - H^{(2)} - cH^{(3)}\Big). \qquad (4.4.9)$$

Here the quantities

$$D^{(r)} = \sum_{J=1}^{\infty} (2J + r - 1) \ln E_J^{(r)} - \sum_{J=0}^{\infty} (2J+1) \ln E_J^{(0)}, \qquad (4.4.10)$$

with $r = 1, 2$ originate from Eq. (4.4.7). The quantities $H^{(i)}$ originate from Eq. (4.4.8) and are defined by

$$H^{(1)} = \sum_{J=0}^{\infty} \sum_{J_1=J}^{J+1} J_1 \frac{E_J^{(0)} - E_{J_1}^{(1)}}{E_J^{(0)} - \alpha E_{J_1}^{(1)}}, \qquad (4.4.11)$$

$$H^{(2)} = \sum_{J=0}^{\infty} \ln E_J^{(0)} \sum_{J_1=J}^{J+1} J_1 \frac{E_J^{(0)} - E_{J_1}^{(1)}}{E_J^{(0)} - \alpha E_{J_1}^{(1)}}, \qquad (4.4.12)$$

$$H^{(3)} = \sum_{J=0}^{\infty} \sum_{J_1=J}^{J+1} J_1 \frac{E_{J_1}^{(1)} \ln E_{J_1}^{(1)}}{E_J^{(0)} - \alpha E_{J_1}^{(1)}}. \qquad (4.4.13)$$

ΔS_F

To obtain the quantum correction ΔS_F we need to carry out the expansion in the non-diagonal terms of Tables 4.3 and 4.4 up to the third order. By taking the appropriate limits as discussed earlier we find the following results

$$\Delta S_F = 2\kappa^2 z\Big\{\mathrm{tr}\Big[(\alpha+c^2)(\mathcal{G}_1(0)-\mathcal{G}_0(0)) - \frac{\alpha}{2}(\mathcal{G}_2(0)-\mathcal{G}_0(0))\Big] \tag{4.4.14.1}$$

$$+ \mathrm{tr}\Big[\alpha(2c|e_1|^2-1)e_0^2\,\mathcal{G}_1(0)] - (3c|e_1|^2-1)e_0^2\mathcal{G}_2(0) - 2c^2e_0^2\mathcal{G}_1(0)\Big] \tag{4.4.14.2}$$

$$+ c^3\int_0^\infty d\omega\,\omega\,\mathrm{tr}\Big[\bar{e}_1\mathcal{G}_0^c(\omega)e_1\mathcal{G}_1^2(\omega) + e_0\mathcal{G}_0^c(\omega)e_0\mathcal{G}_1^2(\omega) - \mathcal{G}_0^c(\omega)\mathcal{G}_0^2(\omega)\Big] \tag{4.4.14.3}$$

$$- c^2\int_0^\infty d\omega\,\omega\,\mathrm{tr}\Big[\bar{e}_1\mathcal{G}_0^c(\omega)e_1\mathcal{G}_1(\omega)e_0^2(2c|e_1|^2-1)\mathcal{G}_1(\omega) + e_0\mathcal{G}_0^c(\omega)e_0\mathcal{G}_1(\omega)e_0^2(2c|e_1|^2-1)\mathcal{G}_1(\omega)\Big] \tag{4.4.14.4}$$

$$+ 2c^2\int_0^\infty d\omega \quad \mathrm{tr}\Big[e_0^2\mathcal{G}_1(\omega)e_0^2\mathcal{G}_1(\omega)\Big] \tag{4.4.14.5}$$

$$+ 5c^2\int_0^\infty d\omega \quad \mathrm{tr}\Big[e_0e_1\mathcal{G}_1(\omega)e_0\bar{e}_1\mathcal{G}_1(\omega)\Big] \tag{4.4.14.6}$$

$$- c^2\int_0^\infty d\omega \quad \mathrm{tr}\Big[e_0e_1\mathcal{G}_0(\omega)e_0\bar{e}_1\mathcal{G}_0(\omega)\Big] \tag{4.4.14.7}$$

$$- c^2\int_0^\infty d\omega\,\omega\,\mathrm{tr}\Big[e_0\mathcal{G}_0^c(\omega)e_0\mathcal{G}_1^2(\omega)\Big] \tag{4.4.14.8}$$

$$- 4c^3\int_0^\infty d\omega\,\omega\,\mathrm{tr}\Big[e_0e_1\mathcal{G}_1(\omega)e_0\mathcal{G}_0^c(\omega)\bar{e}_1\mathcal{G}_1(\omega)\Big]\Big\}. \tag{4.4.14.9}$$

To evaluate these expressions we use the identities (2.A.8)-(2.A.17) (see Appendix 2.A). After some algebra we find

$$\Delta S_F = 2\kappa^2 z^2\sum_{i=1}^{9} B^{(i)}. \tag{4.4.15}$$

Here, the nine contributions $B^{(i)}$, $i=1,\ldots,9$ correspond to the nine equations (4.4.14.1)-(4.4.14.9). The first two of them are given by

$$B^{(1)} = (\alpha+c^2)(Y^{(1)}-Y^{(0)}) - \frac{\alpha}{2}(Y^{(2)}-Y^{(0)}), \tag{4.4.16}$$

$$B^{(2)} = \frac{\alpha}{2}\left(\frac{2c}{3}-1\right)Y^{(1)} + \frac{\alpha}{2}Y^{(2)} - c^2Y^{(1)}, \tag{4.4.17}$$

where we have introduced

$$Y^{(s)} = \sum_{J=1}^{\infty}\frac{2J+(s-1)^2}{E_J^{(r)}}, \qquad s=0,1,2. \tag{4.4.18}$$

The next two terms can be written as

$$B^{(3)} = c^3\sum_{J=0}^{\infty}\sum_{J_1=J}^{J+1} J_1K_\alpha(E_J^{(0)},E_{J_1}^{(1)}) - c^3\sum_{J=0}^{\infty}(2J+1)K_\alpha(E_J^{(0)},E_J^{(0)}) \tag{4.4.19}$$

and

$$B^{(4)} = -\frac{c^2}{2}\left(\frac{2c}{3} - 1\right)\sum_{J=0}^{\infty}\sum_{J_1=J}^{J+1} J_1 K_\alpha(E_J^{(0)}, E_{J_1}^{(1)}). \qquad (4.4.20)$$

The function $K_\alpha(x, y)$ is defined as

$$K_\alpha(x, y) = \frac{x}{(x - \alpha y)^2}\ln\frac{x}{\alpha y} - \frac{1}{x - \alpha y}. \qquad (4.4.21)$$

Notice that $K_\alpha(x, x) = -\ln(\alpha + c)/(c^2 x)$. The next three terms are given by

$$\begin{aligned}
B^{(5)} &= 2c^2\sum_{J=1}^{\infty}\left[\frac{J(6J^2-1)}{3(4J^2-1)}\frac{1}{E_J^{(1)}} + \frac{J(J+1)}{3(2J+1)}L(E_J^{(1)}, E_{J+1}^{(1)})\right], && (4.4.22)\\
B^{(6)} &= \frac{5c^2}{3}\sum_{J=1}^{\infty}\left[\frac{J}{4J^2-1}\frac{1}{E_J^{(1)}} + 2\frac{J(J+1)}{2J+1}L(E_J^{(1)}, E_{J+1}^{(1)})\right], && (4.4.23)\\
B^{(7)} &= -\frac{c^2}{3}\sum_{J=0}^{\infty}(J+1)L(E_J^{(1)}, E_{J+1}^{(1)}). && (4.4.24)
\end{aligned}$$

We have introduced the function

$$L(x, y) = \frac{\ln x - \ln y}{x - y}, \qquad (4.4.25)$$

such that $L(x, x) = 1/x$. Finally, the last to terms are

$$B^{(8)} = -\frac{c^2}{2}\sum_{J=0}^{\infty}\sum_{J_1=J}^{J+1} J_1 K_\alpha(E_J^{(0)}, E_{J_1}^{(1)}) \qquad (4.4.26)$$

and

$$B^{(9)} = \frac{c^3}{3}\sum_{J=0}^{\infty}\frac{(2J+1)^2+2}{2J+1}F_\alpha(E_J^{(0)}, E_J^{(1)}, E_{J+1}^{1}). \qquad (4.4.27)$$

Here the function $F_\alpha(x, y, z)$ is defined as follows

$$F_\alpha(x, y, z) = \frac{1}{y - z}\left[\frac{y}{x - \alpha y}\ln\frac{x}{\alpha y} - \frac{z}{x - \alpha z}\ln\frac{x}{\alpha z}\right]. \qquad (4.4.28)$$

such that $F_\alpha(x, y, y) = K_\alpha(x, y)$.

4.4.2 Regularized expressions

To obtain the regularized theory one has to include the alternating metric e_f and add the masses $\mathcal{M}_f$ to the energies $E_J^{(a)}$ in the expressions for $D^{(r)}$, $H^{(i)}$, $Y^{(s)}$ and $B^{(i)}$ respectively. We will proceed by discussing the regularization of ΔS_σ and ΔS_F separately.

ΔS_σ

To start let us define the function

$$\Phi^{(\Lambda)}(p) = \sum_{J=p}^{\Lambda} 2J \ln(J^2 - p^2). \tag{4.4.29}$$

According to Eq. (4.4.3) the regularized function $\Phi^{(\Lambda)}_{\rm reg}(p)$ is

$$\Phi^{(\Lambda)}_{\rm reg}(p) = \sum_{f=1}^{K} e_f \sum_{J=p}^{\Lambda} 2J \ln(J^2 - p^2 + \mathcal{M}_f^2) \sum_{J=p+1}^{\Lambda} 2J \ln(J^2 - p^2). \tag{4.4.30}$$

where we assume that the cut-off Λ is much larger than $\mathcal{M}_f$. In the presence of a large mass $\mathcal{M}_f$ we may consider the logarithm to be a slowly varying function of the discrete variable J. We may therefore approximate the summation by means of the Euler-Maclaurin formula

$$\sum_{J=p+1}^{\Lambda} g(J) = \int_p^{\Lambda} g(x)dx + \frac{1}{2} g(x)\Big|_p^{\Lambda} + \frac{1}{12} g'(x)\Big|_p^{\Lambda}. \tag{4.4.31}$$

After some algebra we find that Eq.(4.4.30) can be written as follows [20]

$$\begin{aligned} \Phi^{(\Lambda)}_{\rm reg}(p) &= -2\Lambda(\Lambda+1)\ln\Lambda + \Lambda^2 - \frac{\ln e\Lambda}{3} + 4\sum_{J=1}^{\Lambda} J \ln J \\ &+ \frac{1-6p}{3} \ln \mathcal{M} + 2p^2 - 2\sum_{J=1}^{2p} (J-p)\ln J. \end{aligned} \tag{4.4.32}$$

The regularized expression for $D^{(r)}$ can now be obtained as

$$D^{(r)}_{\rm reg} = \lim_{\Lambda\to\infty} \left[\Phi^{(\Lambda)}_{\rm reg}\left(\frac{1+r}{2}\right) - \Phi^{(\Lambda)}_{\rm reg}\left(\frac{1}{2}\right) \right]. \tag{4.4.33}$$

The final results are obtained as follows

$$D^{(1)}_{\rm reg} = -\ln \mathcal{M} + \frac{3}{2} - 2\ln 2 \tag{4.4.34}$$

$$D^{(2)}_{\rm reg} = -2\ln \mathcal{M} + 4 - 3\ln 3 - \ln 2. \tag{4.4.35}$$

The evaluation of $H^{(i)}_{\rm reg}$ is somewhat more subtle but proceeds along similar lines. The results can be written as follows

$$H^{(1)}_{\rm reg} = -\frac{\alpha}{c^2} \left[2\ln \mathcal{M} + 1 - \psi\left(\frac{3c-1}{c}\right) - \psi\left(\frac{1}{c}\right) \right]. \tag{4.4.36}$$

The Euler digamma function $\psi(z)$ appears as a result of the following summation $\sum_{J=0}^{\infty}\left[(J+1)^{-1}-(J+z)^{-1}\right]=\psi(z)-\psi(1)$. Similarly we find

$$\begin{aligned} H_{\rm reg}^{(2)} &= -\lim_{\Lambda\to\infty}\Phi_{\rm reg}^{(\Lambda)}\left(\frac{1}{2}\right)-\frac{\alpha}{c^2}\Big[2\ln\mathcal{M}+1+2\ln^2\mathcal{M} \\ &+ 4\gamma_S+f\left(\frac{\alpha}{c}\right)+f\left(1-\frac{\alpha}{c}\right)\Big], \qquad (4.4.37)\\ H_{\rm reg}^{(3)} &= \frac{1}{c}\lim_{\Lambda\to\infty}\Phi_{\rm reg}^{(\Lambda)}(1)+\frac{1}{c^3}\left(2\ln\mathcal{M}+1\right)+\frac{\alpha}{c^3}\Big[2\ln^2\mathcal{M} \\ &+ 4\gamma_S+f\left(\frac{1}{c}\right)+f\left(1-\frac{\alpha}{c}\right)-2c^2\frac{\ln 2}{1-2\alpha}\Big]. \qquad (4.4.38)\end{aligned}$$

where $\gamma_S\approx -0.0728$ is the Stieltjes constant and

$$f(z)=2z^2\sum_{J=2}^{\infty}\frac{\ln J}{J(J^2-z^2)}. \qquad (4.4.39)$$

We finally have the following total result for the quantum correction ΔS_σ

$$\exp\Delta S_\sigma^{\rm reg}=\frac{27}{8}\tilde{\mathcal{D}}(c)\exp\left[4\left(1+\frac{\alpha\ln\alpha}{c}\right)\ln\mathcal{M}+1\right], \qquad (4.4.40)$$

where

$$\begin{aligned}\ln\tilde{\mathcal{D}}(c) &= -2\frac{\alpha}{c}\Bigg\{\left[\psi\left(\frac{3c-1}{c}\right)+\psi\left(\frac{1}{c}\right)-1\right]\ln\alpha \\ &- f\left(\frac{1-c}{c}\right)+f\left(\frac{1}{c}\right)-2\ln 2\frac{c^2}{2c-1}\Bigg\}. \qquad (4.4.41)\end{aligned}$$

Notice that according to Eq. (4.4.41) the quantity $\tilde{\mathcal{D}}(c)$ depends on the crossover parameter c in a highly non-trivial fashion. Some of contributions diverge at the points $c_k=1/k$ with $k=1,2,3,\dots$ but the final total answer remains finite for all values of c in the interval $0\leqslant c\leqslant 1$ ranging from $\tilde{\mathcal{D}}(0)=e^{-2}$ to $\tilde{\mathcal{D}}(1)=1$. A plot of the function $\tilde{\mathcal{D}}(c)$ with varying c is shown in Fig. 4.4.

ΔS_F

Notice that in contrast to the expression for $\Delta S_\sigma^{\rm reg}$ where the numerical constants play an important role, the expression for $\Delta S_F^{\rm reg}$ can only be determined up to the logarithmic singularity in the Pauli-Villars mass $\mathcal{M}$. In the latter case the constant terms should actually be considered to be of order $1/\sigma_{xx}$ which is beyond the level of approximation as considered in this paper. Keeping this in mind we proceed and define the following function

$$Y^{(\Lambda)}(p)=\sum_{J=p}^{\Lambda}\frac{2J}{J^2-p^2}. \qquad (4.4.42)$$

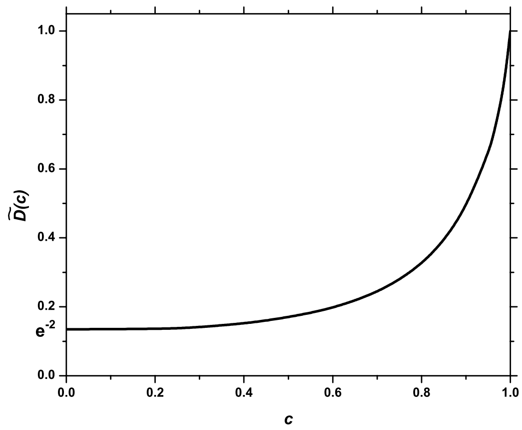


Figure 4.4: The plot of the function $\tilde{\mathcal{D}}(c)$

According to Eq. (4.4.3) the regularized function $Y_{\rm reg}^{(\Lambda)}(p)$ is given by

$$Y_{\rm reg}^{(\Lambda)}(p) = \sum_{f=1}^{K} e_f \sum_{J=p}^{\Lambda} \frac{2J}{J^2 - p^2 + \mathcal{M}_f^2} + \sum_{J=p+1}^{\Lambda} \frac{2J}{J^2 - p^2} \tag{4.4.43}$$

where as before we assume that $\Lambda \gg \mathcal{M}_f$. Proceeding along similar lines as discussed earlier we now find

$$Y_{\rm reg}^{(\Lambda)}(p) = 2\ln\mathcal{M} + 2\gamma_E - \sum_{J=1}^{2p} \frac{1}{J} \tag{4.4.44}$$

where, we remind, $\gamma_E \approx 0.577$ is the Euler constant. The regularized expressions for $Y^{(s)}$ can be written as

$$Y_{\rm reg}^{(s)} = \lim_{\Lambda\to\infty} Y_{\rm reg}^{(\Lambda)}\left(\frac{1+s}{2}\right). \tag{4.4.45}$$

We finally obtain

$$Y_{\rm reg}^{(s)} = 2\ln\mathcal{M} + 2\gamma_E - 1 - \sum_{J=2}^{s+1} \frac{1}{J}. \tag{4.4.46}$$

Within the same logarithmic accuracy we can substitute $K_\alpha(x,x)$ for the functions $K_\alpha(x,y)$ and $F_\alpha(x,y,z)$ in Eqs. (4.4.19)-(4.4.22) and (4.4.27). Similarly we write $L(x,x)$ for $L(x,y)$ in Eqs. (4.4.23)-(4.4.26). With the help of Eq. (4.4.46) we then find

$$B^{(1)} = 0 \ln \mathcal{M}, \tag{4.4.47}$$

$$B^{(2)} = \left(\frac{2c(1-c)}{3} - 2c^2\right) \ln \mathcal{M}, \tag{4.4.48}$$

$$B^{(3)} = 0 \ln \mathcal{M}, \tag{4.4.49}$$

$$B^{(4)} = \left(\frac{2c}{3} - 1\right) (\ln \alpha + c) \ln \mathcal{M}, \tag{4.4.50}$$

$$B^{(5)} = \frac{4c^2}{3} \ln \mathcal{M}, \tag{4.4.51}$$

$$B^{(6)} = \frac{5c^2}{3} \ln \mathcal{M}, \tag{4.4.52}$$

$$B^{(7)} = -\frac{c^2}{3} \ln \mathcal{M}, \tag{4.4.53}$$

$$B^{(8)} = (\ln \alpha + c) \ln \mathcal{M}, \tag{4.4.54}$$

$$B^{(9)} = -\frac{2c}{3} (\ln \alpha + c) \ln \mathcal{M}. \tag{4.4.55}$$

The final total result for ΔS_F^{reg} can now be written as follows

$$\Delta S_F^{\text{reg}} = \frac{32}{3\sigma_{xx}} \pi T z c (\ln \mathcal{M} e^{\gamma_E - 1/2} + \text{const}). \tag{4.4.56}$$

Regularized Z_{inst}/Z_0

We next collect the various different contributions together and obtain the following result for the instanton contribution to the thermodynamic potential

$$\ln \left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} = 3 \ln 3 - 7 \ln 2 - \ln \pi + \ln \mathcal{D}(c) + i\theta \tag{4.4.57}$$

$$- \quad 2\pi\sigma_{xx} \Big[1 - \frac{2}{\pi\sigma_{xx}} \Big(1 + \frac{\alpha}{c} \ln \alpha\Big) \ln \mathcal{M} e^{\gamma_E}\Big] \tag{4.4.58}$$

$$+ \quad \frac{16\pi^2}{3} T z c \Big[1 - \frac{1}{\pi\sigma_{xx}} \ln \mathcal{M} e^{\gamma_E - 1/2}\Big] \tag{4.4.59}$$

$$- \quad 16\pi^2 T z \Big[1 - \frac{c}{\pi\sigma_{xx}} \ln \mathcal{M} e^{\gamma_E - 1/2}\Big]. \tag{4.4.60}$$

We have introduced new function $\mathcal{D}(c)$ which is defined as

$$\mathcal{D}(c) = 16\pi \tilde{\mathcal{D}}(c) \exp\Big[1 - 4\Big(1 + \frac{\alpha}{c} \ln \alpha\Big) \gamma_E\Big]. \tag{4.4.61}$$

A plot of $\mathcal{D}(c)$ with varying c is shown in Fig. 4.5.

4.4.3 Observable theory in Pauli-Villars regularization

The most important result next is that the quantum corrections to the parameters σ_{xx}, zc, and z in Eqs. (4.4.58)-(4.4.60) are all identically the same as those obtained from

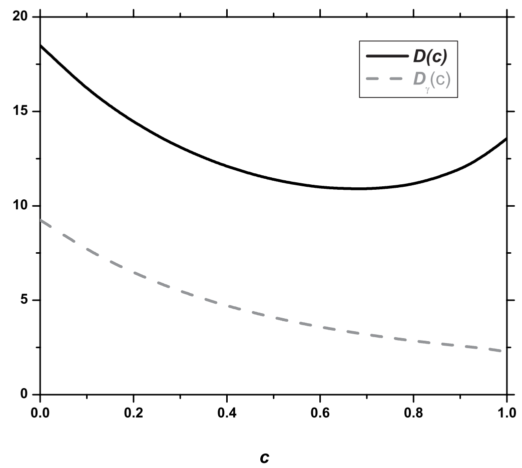


Figure 4.5: The functions $D(c)$ and $D_\gamma(c)$.

a perturbative expansion of the observable parameters σ_{xx}, $z'c'$, and z' introduced in Section 4.2.3. In Appendix 4.A we give the details of the computation. Denoting the results for σ'_{xx}, z' and c' by $\sigma_{xx}(\mathcal{M})$, $z(\mathcal{M})$ and $c(\mathcal{M})$ respectively then we have

$$\sigma_{xx}(\mathcal{M}) = \sigma_{xx}\left[1 - \frac{\beta_0(c)}{\sigma_{xx}}\ln \mathcal{M}e^{\gamma_E}\right], \tag{4.4.62}$$

$$z(\mathcal{M})c(\mathcal{M}) = zc\left[1 + \frac{\gamma_0}{\sigma_{xx}}\ln \mathcal{M}e^{\gamma_E - 1/2}\right], \tag{4.4.63}$$

$$z(\mathcal{M}) = z\left[1 + \frac{c\gamma_0}{\sigma_{xx}}\ln \mathcal{M}e^{\gamma_E - 1/2}\right]. \tag{4.4.64}$$

The results of Eqs. (4.4.57)-(4.4.60) can therefore be written as follows

$$\left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} = \frac{27\mathcal{D}(c)}{128\pi}\exp\left(-2\pi\sigma_{xx}(\mathcal{M}) + i\theta + \check{S}_F^{\text{inst}}[\rho]\right), \tag{4.4.65}$$

where

$$\check{S}_F^{\text{inst}}[\rho] = \pi T z(\mathcal{M})\int d\mathbf{r}\mu^2(\mathbf{r})\Big(c(\mathcal{M})\sum_{\alpha n}\operatorname{tr} I_n^\alpha\rho \operatorname{tr} I_{-n}^\alpha\rho + 4\operatorname{tr}\eta\rho\Big). \tag{4.4.66}$$

Notice that the expression in the exponent is similar to the classical action with the rotation matrix $\mathcal{T}_0$ put equal to unity. The main difference is in the expressions for $\sigma_{xx}(\mathcal{M})$, $z(\mathcal{M})c(\mathcal{M})$ as well as $z(\mathcal{M})$ which are all precisely the radiative corrections as obtained from the observable theory.

At this stage of the analysis several remarks are in order. First of all, it is important to stress that our result for the observable theory, Eq. (4.4.62), uniquely fixes the amplitude $\mathcal{D}(c)$ of the thermodynamic potential which is left unresolved otherwise. This aspect of the problem is going to play a significant role when extracting the non-perturbative renormalization behavior of the theory. In fact, we shall see later on, in Section 4.5.4, that the most important features of the theory, notably the values of $\mathcal{D}(c)$ at $c = 0$ and $c = 1$ respectively, are universal in the sense independent of they are independent of the specific regularization scheme that one uses to define the renormalized theory. Secondly, our results demonstrate that the idea of spatially varying masses does not alter the ultraviolet singularity structure of the instanton theory. In particular, Eqs. (4.4.57)-(4.4.66) display exactly the same logarithms as found previously in flat space and by employing dimensional regularization. [33] The detailed computations of Appendix 4.A provide a deeper understanding of this aspect of the problem, especially where it says that the Pauli-Villars regularization scheme generally retains translational invariance of the electron gas. These extremely important results permit one to proceed and decipher the consequences of our theory of curved space in terms of flat space.

4.5 Transformation from curved space to flat space

4.5.1 Instanton manifold

Integration over zero frequency modes

We are now in a position to extend the results for the thermodynamic potential to include the integration over the zero modes. The complete expression for Z_{inst}/Z_0 can be written as follows [20]

$$\left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} \rightarrow \frac{A_{\text{inst}}}{A_0}\frac{\int \mathcal{D}[Q_{\text{inst}}]}{\int \mathcal{D}[Q_0]}\left[\frac{Z_{\text{inst}}[Q_{\text{inst}}]}{Z_0[Q_0]}\right]^{\text{reg}}. \tag{4.5.1}$$

The meaning of the symbols is as follows.

$$\begin{aligned}\left[\frac{Z_{\text{inst}}[Q_{\text{inst}}]}{Z_0[Q_0]}\right]^{\text{reg}} &= \frac{27\mathcal{D}(c)}{128\pi}e^{-2\pi\sigma_{xx}(\mathcal{M})+i\theta} \\ &+ \frac{e^{z(\mathcal{M})\int d\eta d\theta\left(c(\mathcal{M})\sum_{\alpha n}\operatorname{tr} I_n^\alpha Q_{\text{inst}}\operatorname{tr} I_{-n}^\alpha Q_{\text{inst}}+4\operatorname{tr}\eta Q_{\text{inst}}\right)}}{e^{z(\mathcal{M})\int d\eta d\theta\left(c(\mathcal{M})\sum_{\alpha n}\operatorname{tr} I_n^\alpha Q_0\operatorname{tr} I_{-n}^\alpha Q_0+4\operatorname{tr}\eta Q_0\right)}}.\end{aligned} \tag{4.5.2}$$

Here, Q_{inst} denotes the manifold of the instanton parameters as is illustrated in Fig. 4.3

$$\int \mathcal{D}[Q_{\text{inst}}] \;\; = \;\; \int d\mathbf{r}_0 \int \frac{d\lambda}{\lambda^3}\int \mathcal{D}[\mathcal{T}_0]. \tag{4.5.3}$$

Here, the integral over $\mathcal{T}_0$ can be decomposed according to

$$\int \mathcal{D}[\mathcal{T}_0] = \int \mathcal{D}[t_0]\int \mathcal{D}[U], \tag{4.5.4}$$

where schematically we can write

$$\int \mathcal{D}[t_0] = \int \mathcal{D}\left[\frac{U(2N)}{U(N)\times U(N)}\right], \tag{4.5.5}$$

$$\begin{aligned} \int \mathcal{D}[U] &= \int \mathcal{D}\left[\frac{U(N)}{U(1)\times U(N-1)}\right] \\ &\times \int \mathcal{D}\left[\frac{U(N)}{U(1)\times U(N-1)}\right] \\ &\times \int \mathcal{D}\ [U(1)]. \end{aligned} \tag{4.5.6}$$

On the other hand, the Q_0 are the zero modes associated with the trivial vacuum

$$\int \mathcal{D}[Q_0] = \int \mathcal{D}\left[\frac{U(2N)}{U(N)\times U(N)}\right]. \tag{4.5.7}$$

The numerical factors A_{inst} and A_0 are given by

$$A_{\text{inst}} = \langle e_0^4\rangle\langle|e_1|^4\rangle\langle e_0^2|e_1|^2\rangle\left(\langle e_0^2\rangle\langle|e_1|^2\rangle\right)^{2N-2}\langle 1\rangle^{(N-1)(N-1)}\pi^{-N^2-2N} \tag{4.5.8}$$

$$A_0 = \langle 1\rangle^{N^2}\pi^{-N^2} \tag{4.5.9}$$

where the average $\langle\ldots\rangle$ is with respect to the surface of a sphere

$$\langle f\rangle = \sigma_{xx}\int_{-1}^{1} d\eta \int_0^{2\pi} d\theta f(\eta,\theta). \tag{4.5.10}$$

U rotation

We have already mentioned earlier that the fluctuations in the Goldstone modes $t_0, Q_0 \in U(2N)/U(N)\times U(N)$ have an infinite action in flat space and eventually drop out. Keeping this in mind we can write the result of Eq. (4.5.1) as follows

$$\left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} = \frac{27}{128\pi}\int d\mathbf{r}_0 \int \frac{d\lambda}{\lambda^3}\int \mathcal{D}[U]\frac{A_{\text{inst}}\mathcal{D}(c)}{A_0}e^{S'_{\text{inst}}}. \tag{4.5.11}$$

Here,

$$S'_{\text{inst}} = -2\pi\sigma_{xx}(\mathcal{M}) \pm i\theta + \check{S}_F^{\text{inst}}[U^{-1}\rho U] \tag{4.5.12}$$

with $\check{S}_F$ defined by Eq. (4.4.66). Next, by making use of the identity [20]

$$\int \mathcal{D}\left[\frac{U(k)}{U(1)\times U(k-1)}\right] = \frac{\pi^{k-1}}{\Gamma(k)}, \tag{4.5.13}$$

we can write the result for the thermodynamic potential in the limit $N_r \to 0$ in a more compact fashion as follows

$$\left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} = \frac{N^2}{8\pi^2}\int d\mathbf{r}_0 \int \frac{d\lambda}{\lambda^3}\mathcal{D}(c)\langle e^{S'_{\text{inst}}}\rangle_U, \tag{4.5.14}$$

where the average is defined according to

$$\langle X\rangle_U = \frac{\int \mathcal{D}[U]X}{\int \mathcal{D}[U]}. \tag{4.5.15}$$

Curved space versus flat space

Our final result of Eq. (4.5.14) still involves a spatially varying momentum scale $\mu(\mathbf{r})$ and our task next is to express the final answer in quantities that are defined in *flat* space, rather than *curved* space. The first step is to rewrite the integral $\int d\eta d\theta$ in $\tilde{S}_F^{\rm inst}$ as an integral over *flat* space following the substitution

$$\int d\eta d\theta = \int d\mathbf{r}\mu^2(\mathbf{r}) \quad \rightarrow \quad \int d\mathbf{r}. \tag{4.5.16}$$

The expression for $\tilde{S}_F^{\rm inst}$ now reads

$$\tilde{S}_F^{\rm inst}[U^{-1}\rho U] = \pi T \int' d\mathbf{r} z(\mathcal{M})\Big[c(\mathcal{M})\sum_{\alpha n} \operatorname{tr} I_n^\alpha U^{-1}\rho U \operatorname{tr} I_{-n}^\alpha U^{-1}\rho U + 4\operatorname{tr}\eta U^{-1}\rho U\Big], \tag{4.5.17}$$

where the "prime" on the integral sign reminds us of the fact that the expression for $\tilde{S}_F^{\rm inst}$, as it now stands, still diverges logarithmically in the sample size. What remains, however, is to perform the next step which is to express the Pauli-Villars masses $\mathcal{M}$ in terms of the appropriate quantities that are defined in *flat* space. Notice hereto that $\mathcal{M}$ actually describes a *spatially varying* momentum scale $\mu(r)\mathcal{M}$. In Section 4.5.3 below as well as in the remainder of this paper we will embark on the general problem of how to translate a momentum scale in *curved space* into a quantity μ_0 that is defined in *flat space*. As an extremely important consequence of this procedure we shall show in what follows that the final expression for the interaction term $\tilde{S}_F^{\rm inst}$ is finite in the infrared. This remarkable result is the primary reason as to why one can proceed and obtain the non-perturbative corrections to the renormalization of the quantities z and c.

4.5.2 Physical observables

Linear response

Our results for the thermodynamic potential are easily extended to include the quantities σ'_{xx} and θ' defined by Eqs. (4.2.31) and (4.2.32). To leading order in σ_{xx} we obtain the following result (see also Ref. [21])

$$\sigma'_{xx} = \sigma_{xx}(\mathcal{M}) + \int \frac{d\lambda}{\lambda}\mathcal{D}(c)\Big\langle \big(J_{xx}[Q_{\rm inst}]e^{i\theta} + c.c.\big)\, e^{-2\pi\sigma_{xx}(\mathcal{M})+\tilde{S}_F^{\rm inst}}\Big\rangle_U, \tag{4.5.18}$$

$$\frac{\theta'}{2\pi} = \frac{\theta}{2\pi} + \int \frac{d\lambda}{\lambda}\mathcal{D}(c)\Big\langle \big(J_{xy}[Q_{\rm inst}]e^{i\theta} + c.c.\big)\, e^{-2\pi\sigma_{xx}(\mathcal{M})+\tilde{S}_F^{\rm inst}}\Big\rangle_U. \tag{4.5.19}$$

Here, we have introduced the quantity $J_{ab}[Q_{\rm inst}]$ which is given as

$$J_{ab}[Q_{\rm inst}] = N^2\frac{\sigma_{xx}^2}{32\pi^2 n\lambda^2}\int d\mathbf{r}\operatorname{tr} I_n^\alpha U\rho\nabla_a\rho U^{-1}\int d\mathbf{r}'\operatorname{tr} I_{-n}^\alpha U\rho\nabla_b\rho U^{-1}. \tag{4.5.20}$$

The interaction term $\tilde{S}_F^{\text{inst}}$ in Eqs. (4.5.18) and (4.5.19) does not contribute in the limit $T \to 0$ and can be dropped. By using normalization conditions

$$\sum_{n_1,\alpha}(U^{-1})^{1\alpha}_{0,n_1}(U)^{\alpha 1}_{n_1,0} = 1, \tag{4.5.21}$$

$$\sum_{n_2,\alpha}(U^{-1})^{1\alpha}_{-1,n_2}(U)^{\alpha 1}_{n_2,-1} = 1, \tag{4.5.22}$$

we find the following results for the expressions bilinear in the U

$$\left\langle (U)^{\alpha 1}_{n_1,0}(U^{-1})^{1\beta}_{0,n_3}\right\rangle = \frac{1}{N}\delta_{n_1 n_3}\delta^{\alpha\beta}, \tag{4.5.23}$$

$$\left\langle (U)^{\alpha 1}_{n_2,-1}(U^{-1})^{1\beta}_{-1,n_4}\right\rangle = \frac{1}{N}\delta_{n_3 n_4}\delta^{\alpha\beta}. \tag{4.5.24}$$

For the quartic combinations we find

$$\left\langle (U)^{\alpha 1}_{n_1,0}(U^{-1})^{1\beta}_{0,n_3}(U)^{\gamma 1}_{n_5,0}(U^{-1})^{1\delta}_{0,n_7}\right\rangle = \frac{\delta^{\alpha\beta}_{n_1n_3}\delta^{\gamma\delta}_{n_5n_7}+\delta^{\alpha\delta}_{n_1n_7}\delta^{\gamma\beta}_{n_5n_3}}{N(1+N)}, \tag{4.5.25}$$

$$\left\langle (U)^{\alpha 1}_{n_2,-1}(U^{-1})^{1\beta}_{-1,n_4}(U)^{\gamma 1}_{n_6,-1}(U^{-1})^{1\delta}_{-1,n_8}\right\rangle = \frac{\delta^{\alpha\beta}_{n_2n_4}\delta^{\gamma\delta}_{n_6n_8}+\delta^{\alpha\delta}_{n_2n_8}\delta^{\gamma\beta}_{n_4n_6}}{N(1+N)}. \tag{4.5.26}$$

We have used the shorthand notation $\delta^{\alpha\beta}_{n_1n_3} \equiv \delta_{n_1n_3}\delta^{\alpha\beta}$. In the limit where $N_r \to 0$ we obtain

$$\langle J_{ab}[Q_{\text{inst}}]\rangle_U = \frac{\sigma_{xx}^2}{32\pi^2\lambda^2}\int d\mathbf{r}(\rho\nabla_a\rho)^{11}_{-1,0}\int d\mathbf{r}'(\rho\nabla_b\rho)^{11}_{0,-1} = \frac{\sigma_{xx}^2}{2}\left(\delta_{ab} - i\varepsilon_{ab}\right). \tag{4.5.27}$$

The expressions for σ'_{xx} and θ' can now be written as follows

$$\sigma'_{xx} = \sigma_{xx}(\mathcal{M}) - \int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\sigma_{xx}^2 e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta, \tag{4.5.28}$$

$$\frac{\theta'}{2\pi} = \frac{\theta}{2\pi} - \int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\sigma_{xx}^2 e^{-2\pi\sigma_{xx}(\mathcal{M})}\sin\theta. \tag{4.5.29}$$

Specific heat

By using definitions in Section 4.2.3 we obtain the following results for the parameters z' and $z'c'$

$$z' = z(\mathcal{M}) - \int' \frac{d\lambda}{\lambda}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta\frac{N^2}{8\pi^3\lambda^2 T\,\mathrm{tr}\,\eta\Lambda}\langle\tilde{S}_F[U]\rangle_U, \tag{4.5.30}$$

$$z'c' = z(\mathcal{M})c(\mathcal{M}) - \int' \frac{d\lambda}{\lambda}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta\frac{N^2}{8\pi^3\lambda^2 T\,\mathrm{tr}\,\eta\Lambda}\langle\tilde{S}_F[U]\rangle_U. \tag{4.5.31}$$

The expectations can be evaluated along the same lines as was done in the previous Section and the result is

$$\left\langle \tilde{S}_F[U]\right\rangle_U = \frac{2\pi T}{N^2}\int d\mathbf{r}z(\mathcal{M})c(\mathcal{M})|\rho^{11}_{00}(r)|\,\mathrm{tr}\,\eta\Lambda. \tag{4.5.32}$$

Eqs. (4.5.31) and (4.5.30) therefore greatly simplify and we obtain

$$z' = z(\mathcal{M}) \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta\int'\frac{d\mathbf{r}}{\lambda}z(\mathcal{M})c(\mathcal{M})\mu(\mathbf{r}), \tag{4.5.33}$$

$$z'c' = z(\mathcal{M})c(\mathcal{M}) \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\mathcal{M})}\cos\theta\int'\frac{d\mathbf{r}}{\lambda}z(\mathcal{M})c(\mathcal{M})\mu(\mathbf{r}). \tag{4.5.34}$$

The most important feature of these results is that the non-perturbative contributions to the observable parameters σ'_{xx}, θ', c' and z' are all unambiguously expressed in terms of the perturbative quantities $\sigma_{xx}(\mathcal{M})$, $\theta(\nu_f)$, $c(\mathcal{M})$ and $z(\mathcal{M})$.

4.5.3 Transformation $\mu^2(\mathbf{r})\mathcal{M} \to \mu_0$

As a last step in the development of a quantum theory we next wish to express the Pauli-Villars masses which carry the metric of a sphere $\mu^2(\mathbf{r})\mathcal{M}^2$ in terms of a mass or momentum scale in flat space, say μ_0^2. By changing the momentum scale from $\mu(\mathbf{r})\mathcal{M}$ to μ_0 one changes the renormalized theory according to

$$\begin{aligned}\sigma_{xx}(\mathcal{M}) \to \sigma_{xx}(\mathcal{M})&\left[1+\frac{\beta_0(c)}{\sigma_{xx}}\ln\frac{\mu(\mathbf{r})\mathcal{M}}{\mu_0}\right]\\ &= \sigma_{xx}\left[1-\frac{\beta_0(c)}{\sigma_{xx}}\ln\frac{\mu_0}{\mu(\mathbf{r})}e^{\gamma_E}\right] = \sigma_{xx}(\mu(\mathbf{r})),\end{aligned} \tag{4.5.35}$$

$$\begin{aligned}c(\mathcal{M}) \to c(\mathcal{M})&\left[1+\alpha\frac{\gamma_0}{\sigma_{xx}}\ln\frac{\mu(\mathbf{r})\mathcal{M}}{\mu_0}\right]\\ &= c\left[1-\alpha\frac{\gamma_0}{\sigma_{xx}}\ln\frac{\mu_0}{\mu(\mathbf{r})}e^{\gamma_E-1/2}\right] = c(\mu(\mathbf{r})),\end{aligned} \tag{4.5.36}$$

$$\begin{aligned}z(\mathcal{M}) \to z(\mathcal{M})&\left[1+c\frac{\gamma_0}{\sigma_{xx}}\ln\frac{\mu(\mathbf{r})\mathcal{M}}{\mu_0}\right]\\ &= z\left[1-c\frac{\gamma_0}{\sigma_{xx}}\ln\frac{\mu_0}{\mu(\mathbf{r})}e^{\gamma_E-1/2}\right] = z(\mu(\mathbf{r})).\end{aligned} \tag{4.5.37}$$

The introduction of spatially varying parameters $\sigma_{xx}(\mu(\mathbf{r}))$, $c(\mu(\mathbf{r}))$ and $z(\mu(\mathbf{r}))$ means that the action S'_{inst} gets modified according to the prescription

$$S'_{\text{inst}} \to -\int d\mathbf{r}\,\sigma_{xx}(\mu(\mathbf{r}))\,\text{tr}(\nabla Q_{\text{inst}}(\mathbf{r}))^2 \pm i\theta + \hat{S}_F[U], \tag{4.5.38}$$

where

$$\hat{S}_F[U] = \pi T\int' d\mathbf{r}z(\mu(\mathbf{r}))\Big[c(\mu(\mathbf{r}))\sum_{\alpha n}\text{tr}\,I_n^\alpha U^{-1}\rho U\,\text{tr}\,I_{-n}^\alpha U^{-1}\rho U + 4\,\text{tr}\,\eta U^{-1}\rho U\Big]. \tag{4.5.39}$$

Notice that in these expressions the instanton quantity ρ depends explicitly on $\mathbf{r}$ and should be read as $\rho = \rho(\mathbf{r})$.

4.5.4 The quantities σ_{xx} and σ'_{xx} in flat space

Transformation

Let us first evaluate the first spatial integral in Eq. (4.5.38) which can be written as

$$\int d\mathbf{r}\, \sigma_{xx}(\mu(\mathbf{r}))\, \mathrm{tr}(\nabla Q_{\mathrm{inst}}(\mathbf{r}))^2 = \int d\mathbf{r} \mu^2(\mathbf{r}) \sigma_{xx}(\mu(\mathbf{r})) = 2\pi \sigma_{xx}(\zeta\lambda), \tag{4.5.40}$$

where

$$\sigma_{xx}(\zeta\lambda) = \sigma_{xx} - \beta_0(c) \ln \zeta\lambda\mu_0 e^{\gamma_E}, \qquad \zeta = e^2/4. \tag{4.5.41}$$

Notice that the expression for $\sigma_{xx}(\zeta\lambda)$ can simply be obtained from $\sigma_{xx}(\mathcal{M})$ by replacing the Pauli-Villars mass $\mathcal{M}$ according to

$$\mathcal{M} \to \zeta\lambda\mu_0. \tag{4.5.42}$$

By using this result we can write the expression for σ'_{xx} as follows

$$\sigma'_{xx} = \sigma_{xx}(\mathcal{M}) - \int' \frac{d\lambda}{\lambda} \mathcal{D}(c) \sigma^2_{xx} e^{-2\pi\sigma_{xx}(\zeta\lambda)} \cos\theta. \tag{4.5.43}$$

This expression for σ'_{xx} still contains the Pauli-Villars mass $\mathcal{M}$. To complete the transformation from curved space to flat space we first change the momentum scale of the *observable* theory from $\mu'(\mathbf{r})\mathcal{M}$ to μ_0. Write

$$\sigma'_{xx}(\mathcal{M}) \to \sigma'_{xx}(\mu'(\mathbf{r})),$$

then completely analogous to the definition of Eq. (4.5.40) we obtain the observable parameter σ'_{xx} in flat space according to the prescription

$$\sigma'_{xx}(\zeta\lambda') = \frac{1}{2\pi} \int d\mathbf{r} (\mu'(\mathbf{r}))^2 \sigma'_{xx}(\mu'(\mathbf{r})). \tag{4.5.44}$$

One can think of the $\mu'(\mathbf{r}) = 2\lambda'/(r^2 + \lambda'^2)$ as being a background instanton with a large scale size λ'. The expressions for σ'_{xx} and θ' in flat space can now be written as follows

$$\begin{aligned} \sigma'_{xx}(\zeta\lambda') &= \sigma_{xx}(\zeta\lambda') - \int' \frac{d[\zeta\lambda]}{\zeta\lambda} \mathcal{D}(c) \sigma^2_{xx} e^{-2\pi\sigma_{xx}(\zeta\lambda)} \cos\theta, && (4.5.45) \\ \theta'(\zeta\lambda') &= \theta \quad - 2\pi \int' \frac{d[\zeta\lambda]}{\zeta\lambda} \mathcal{D}(c) \sigma^2_{xx} e^{-2\pi\sigma_{xx}(\zeta\lambda)} \sin\theta. && (4.5.46) \end{aligned}$$

Integration over scale sizes λ

Notice that the expression for $\sigma_{xx}(\zeta\lambda')$ has precisely the same meaning as Eq. (4.5.41) with λ replaced by λ'. Next, by writing $\sigma_{xx}(\zeta\lambda')$ as an integral over scale sizes

$$\sigma_{xx}(\zeta\lambda') = \sigma^0_{xx} - \int'_{1/\mu_0 e^{\gamma_E}} \frac{d[\zeta\lambda]}{\zeta\lambda} \beta_0(c) \tag{4.5.47}$$

where $\sigma^0_{xx} = \sigma_{xx}(1/\mu_0 e^{\gamma_E})$ we obtain a natural general expression for the observable theory

$$\sigma'_{xx}(\zeta\lambda') = \sigma^0_{xx} - \int_{1/\mu_0 e^{\gamma_E}}^{\prime} \frac{d[\zeta\lambda]}{\zeta\lambda} \left[\beta_0(c) + \mathcal{D}(c)\sigma^2_{xx} e^{-2\pi\sigma_{xx}(\zeta\lambda)} \cos\theta\right] \quad (4.5.48)$$

$$\theta'(\zeta\lambda') = \theta - 2\pi \int_{1/\mu_0 e^{\gamma_E}}^{\prime} \frac{d[\zeta\lambda]}{\zeta\lambda} \left[0 \quad + \mathcal{D}(c)\sigma^2_{xx} e^{-2\pi\sigma_{xx}(\zeta\lambda)} \sin\theta\right]. \quad (4.5.49)$$

Notice that the contributions from instantons are finite in the ultraviolet and the limit $\mu_0 \to \infty$ was taken implicitly in the computation of the original expressions of Eqs. (4.5.45) and (4.5.46). Comparison of Eqs. (4.5.47)-(4.5.49) with the results obtained from the theory in dimensional regularization, Eq. (3.2.19), clearly shows that the integral over scale sizes λ has exactly the same significance as the integral over momentum scales that generally defines the relation between to the observable theory and the renormalization group β functions. Hence, we have found the natural meaning of the instanton scale size λ. This meaning can only be found on the basis of the theory of observable parameters and clearly does not arise from the usual free energy considerations that historically has left the subject matter largely obscure. [115, 116]

In the absence of the instanton contributions the observable theory of Eqs. (4.5.48) and (4.5.49) and the renormalized theory of Eq. (4.5.47) are identically the same. At a perturbative level we have $\theta' = \theta$ which are arbitrary length scale independent parameters. Eqs. (4.5.48) and (4.5.49) indicate, however, that instantons are important in order to correctly describe the infrared of the system. The observable parameter θ', unlike θ, now becomes length scale dependent and following the discussion in Section 4.2.2 we have found the fundamental mechanism by which the system *dynamically* generates the quantization of the Hall conductance. Before embarking on the renormalization of the theory we shall first address the various difficulties associated with the observable parameters z' and $z'c'$. This will be done in the Sections below and we will come back to the β and γ functions of the theory in Section 4.6.

4.5.5 The quantities z, zc and z', $z'c'$ in flat space

In this Section we wish to extend the various steps of Eqs. (4.5.40)-(4.5.44) and translate the parameters $z(\mathcal{M})$ and $z'(\mathcal{M})$ as well as $z(\mathcal{M})c(\mathcal{M})$ and $z'(\mathcal{M})c'(\mathcal{M})$ into the appropriate quantities in *flat* space. As an important check on the procedure we compare the results with those previously obtained from the theory in dimensional regularization. At the same time we will have to make sure that at each step of the analysis the relation $z'\alpha' = z\alpha$ is satisfied. For the main part, however, the analysis of the present Section proceeds along the same lines as the one in Chapter 2 for the ordinary Grassmannian theory.

Transformation

Let us first introduce the spatially varying momentum scales $\mu(\mathbf{r})$ and $\mu'(\mathbf{r})$ according to Eqs. (4.5.36) and (4.5.37)

$$z' = z(\mu'(\mathbf{r})) \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\mathcal{A}_1 e^{-2\pi\sigma_{xx}(\zeta\lambda')}\cos\theta, \qquad (4.5.50)$$

$$z'c' = z(\mu'(\mathbf{r}))c(\mu'(\mathbf{r})) \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\mathcal{A}_1 e^{-2\pi\sigma_{xx}(\zeta\lambda')}\cos\theta, \qquad (4.5.51)$$

where the amplitude $\mathcal{A}_1$ is given as

$$\mathcal{A}_1 = \int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda} z(\mu(\mathbf{r}))c(\mu(\mathbf{r})). \qquad (4.5.52)$$

Next, by using exactly the same procedure as in Eqs. (4.5.40) and (4.5.44) we wish to express the quantities $z(\mu(\mathbf{r}))$ and $z(\mu(\mathbf{r}))c(\mu(\mathbf{r}))$ in the amplitude $\mathcal{A}_1$ in terms of $z(\zeta\lambda)$ and $z(\zeta\lambda)c(\zeta\lambda)$ respectively according to

$$z(\zeta\lambda) \quad = \quad \frac{1}{2\pi}\int d\mathbf{r}\mu^2(\mathbf{r})z(\mu(\mathbf{r})), \qquad (4.5.53)$$

$$z(\zeta\lambda)c(\zeta\lambda) \quad = \quad \frac{1}{2\pi}\int d\mathbf{r}\mu^2(\mathbf{r})z(\mu(\mathbf{r}))c(\mu(\mathbf{r})). \qquad (4.5.54)$$

From this one obtains

$$z(\zeta\lambda) \quad = \quad z\left[1 - \frac{c\gamma_0}{\sigma_{xx}}\ln\zeta\lambda\mu_0 e^{\gamma_E - 1/2}\right], \qquad (4.5.55)$$

$$z(\zeta\lambda)c(\zeta\lambda) \quad = \quad zc\left[1 - \frac{\gamma_0}{\sigma_{xx}}\ln\zeta\lambda\mu_0 e^{\gamma_E - 1/2}\right]. \qquad (4.5.56)$$

These expressions for the renormalized parameters as well as the $\sigma_{xx}(\zeta\lambda)$ obtained in the previous Section are precisely consistent with those of the theory in dimensional regularization. Although these quantities are clearly the ones of interest to us, it is not quite obvious in what manner the amplitude $\mathcal{A}_1$ can be expressed in terms of Eqs. (4.5.55) and (4.5.56). Recall that the classical expression for Eq. (4.5.52) diverges logarithmically in the sample size. In the Sections below, however, we will show that the amplitude $\mathcal{A}_1$ is finite at a quantum level. On the other hand, completely analogous to Eq. (4.5.44) we write the observable theory as follows

$$z'(\zeta\lambda') = z(\zeta\lambda') \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\mathcal{A}_1 e^{-2\pi\sigma_{xx}(\zeta\lambda')}\cos\theta, \qquad (4.5.57)$$

$$z'c'(\zeta\lambda') = z(\zeta\lambda')c(\zeta\lambda') \quad + \quad \frac{\gamma_0}{4\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\mathcal{A}_1 e^{-2\pi\sigma_{xx}(\zeta\lambda')}\cos\theta, \qquad (4.5.58)$$

where $z(\zeta\lambda')$ and $z(\zeta\lambda')c(\zeta\lambda')$ are defined by Eqs. (4.5.55) and (4.5.56). Notice that the results so far are consistent with the statements made on $z'\alpha'$ and $z\alpha$.

Amplitude $\mathcal{A}_1$

To evaluate $\mathcal{A}_1$ further it is convenient to introduce the quantity $M_1(\mathbf{r})$ according to

$$\mathcal{A}_1 = z(\mu(0))c(\mu(0))\mathcal{M}_1, \tag{4.5.59}$$

$$\mathcal{M}_1 = -2\pi \int_{\mu(0)}^{\mu(L')} d[\ln\mu(\mathbf{r})] M_1(\mathbf{r}), \tag{4.5.60}$$

$$M_1(\mathbf{r}) = \frac{z(\mu(\mathbf{r}))c(\mu(\mathbf{r}))}{z(\mu(0))c(\mu(0))}. \tag{4.5.61}$$

Since the anomalous dimension $\gamma_{zc} = \gamma_z/c$ is negative the quantity $M_1(\mathbf{r})$ is in all respects like a spatially varying *spontaneous magnetization* in the classical Heisenberg ferromagnet. The associated momentum scale $\mu(\mathbf{r})$ strongly varies from *large* values $O(\lambda^{-1})$ at short distances ($|\mathbf{r}| \ll \lambda$) to *small* values $O(\lambda/|\mathbf{r}|^2)$ at very large distances ($|\mathbf{r}| \gg \lambda$). This means that at distances sufficiently far from the center of the instanton the system is effectively in the *symmetric* or *strong coupling* phase where $M_1(\mathbf{r})$ vanishes. Hence we expect the amplitude $\mathcal{M}_1$ to remain finite as $|\mathbf{r}| \to \infty$. This is in spite of the fact that the amplitude $\mathcal{A}_1$ diverges at a classical level.

Details of computation

The expression for $\mathcal{M}_1$ can be written in terms of the γ_{zc} function as follows

$$\mathcal{M}_1 = -2\pi \int_{\ln\mu(0)}^{\ln\mu(L')} d[\ln\mu(\mathbf{r})] \exp\Big\{-\int_{\ln\mu(0)}^{\ln\mu(\mathbf{r})} d[\ln\mu]\gamma_{zc}\Big\}. \tag{4.5.62}$$

Taking the derivative with respect to $\ln\lambda$ we find that $\mathcal{M}_1$ obeys the following differential equation

$$\left(-\frac{d}{d\ln\lambda} + \gamma_{zc}\right)\mathcal{M}_1 = 2\pi\left(1 + M_1(L')\right). \tag{4.5.63}$$

We can safely take the limit $L' = \infty$ and put $M_1(L') = 0$ from now onward. At the same time one can solve Eq. (4.5.63) in the weak coupling limit where $\lambda \to 0$, $\mu(0) \to \infty$ and $\sigma_{xx}(\mu(0)) \to \infty$. Under these circumstances it suffices to insert the perturbative expressions of Eqs. (3.2.22), (3.2.23) and (3.6.2) for the γ_{zc}, β_σ and β_c functions such that the quantity $\mathcal{M}_1 = \mathcal{M}_1(\sigma_{xx}(\mu(0)), c(\mu(0)))$ is obtained as the solution of the differential equation

$$\left(\beta_\sigma \frac{\partial}{\partial\sigma_{xx}} + \beta_c\frac{\partial}{\partial c} + \gamma_{zc}\right)\mathcal{M}_1 = 2\pi, \tag{4.5.64}$$

where to leading order $\beta_\sigma = \beta_0(c)$, $\beta_c = c(1-c)\gamma_0/\sigma_{xx}$ and $\gamma_{zc} = -\gamma_0/\sigma_{xx}$. The result for $\mathcal{M}_1$ can generally be expanded in powers of $\sigma_{xx}^{-1}(\mu(0))$

$$\mathcal{M}_1 = 2\pi^2\sigma_{xx} m_1^{(1)}(c) + m_0^{(1)}(c) + \sigma_{xx}^{-1} m_{-1}^{(1)}(c) + \ldots \tag{4.5.65}$$

We are interested in the leading order quantity $m_1^{(1)}(c)$ which obeys the following differential equation

$$\left(-\gamma_0 c(1-c)\frac{d}{dc} + (\beta_0(c) - \gamma_0)\right) m_1^{(1)}(c) = \frac{1}{\pi}. \tag{4.5.66}$$

The solution can be written as

$$m_1^{(1)}(c) = \frac{\alpha}{c} \exp\left[\frac{2}{c}\ln\alpha\right] \int_0^c \mathrm{d}s(1-s)^{-2-2/s}. \qquad (4.5.67)$$

The quantity $m_1^{(1)}(c)$ varies between the Fermi liquid value $m_1^{(1)}(0)$ and the Coulomb interaction value $m_1^{(1)}(1)$ which are obtained as

$$m_1^{(1)}(0) = 1, \qquad m_1^{(1)}(1) = 1/3. \qquad (4.5.68)$$

The result for $\mathcal{A}_1$ becomes

$$\mathcal{A}_1 = -2\pi^2 z(\mu(0))c(\mu(0))\sigma_{xx}(\mu(0))m_1^{(1)}(c(\mu(0))). \qquad (4.5.69)$$

As a final step we wish to express $\sigma_{xx}(\mu(0))$, $c(\mu(0))$ and $z(\mu(0))$ in terms of the spatially flat quantities $\sigma_{xx}(\zeta\lambda)$, $c(\zeta\lambda)$ and $z(\zeta\lambda)$ respectively. The following relations are obtained

$$\sigma_{xx}(\mu(0)) = \sigma_{xx}(\zeta\lambda)\left[1 + \frac{\beta_0(c)}{\sigma_{xx}(\zeta\lambda)}\ln 2\zeta\right], \qquad (4.5.70)$$

$$c(\mu(0)) = c(\zeta\lambda)\left[1 + \frac{\alpha\gamma_0}{\sigma_{xx}(\zeta\lambda)}\ln 2\zeta\right], \qquad (4.5.71)$$

$$z(\mu(0)) = z(\zeta\lambda)\left[1 + \frac{c\gamma_0}{\sigma_{xx}(\zeta\lambda)}\ln 2\zeta\right]. \qquad (4.5.72)$$

For our purposes the correction terms $O(\sigma_{xx}^{-1})$ are unimportant. Using these results we have finally solved the problem stated at the outset which is to express the amplitude $\mathcal{A}_1$ in terms of the quantities $\sigma_{xx}(\zeta\lambda)$, $c(\zeta\lambda)$ and $z(\zeta\lambda)$, i.e.

$$\mathcal{A}_1 = 2\pi^2 z(\zeta\lambda)c(\zeta\lambda)\sigma_{xx}(\zeta\lambda)m_1^{(1)}(c(\zeta\lambda)) \qquad (4.5.73)$$

with the function $m_1(c)$ given by Eq. (4.5.67). The complete expressions for the quantities z' and $z'c'$ are as follows

$$z'(\zeta\lambda') = z(\zeta\lambda') - \int' \frac{d[\zeta\lambda]}{\zeta\lambda} zc\mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta, \qquad (4.5.74)$$

$$z'(\zeta\lambda')c'(\zeta\lambda') = z(\zeta\lambda')c(\zeta\lambda') - \int' \frac{d[\zeta\lambda]}{\zeta\lambda} zc\mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta. \qquad (4.5.75)$$

where the quantities σ_{xx}, z and c under the integral sign are defined for length scale $\zeta\lambda$ and

$$\mathcal{D}_\gamma(c) = -\frac{\gamma_0\pi}{2}\mathcal{D}(c)m_1(c). \qquad (4.5.76)$$

In Fig. 4.5 we plot the function $\mathcal{D}_\gamma(c)$ with varying c. It has the Fermi-liquid value $\mathcal{D}_\gamma(0) = \mathcal{D}(0)/2$ and the Coulomb interaction value $\mathcal{D}_\gamma(1) = \mathcal{D}(1)/6$.

Table 4.5: Regularization schemes

	Pauli-Villars regularization (*curved space*)	Pauli-Villars regularization (*flat space*)	Dimensional regularization
σ'_{xx}	$\sigma_{xx} - \beta_0(c)\ln X e^{1/2}$	$\sigma_{xx} - \beta_0(c)\ln X e^{1/2}$	$\sigma_{xx} - \beta_0(c)\ln X e^{1/2}$
z'	$z\left(1 - \frac{c\gamma_0}{\sigma_{xx}}\ln X\right)$	$z\left(1 - \frac{c\gamma_0}{\sigma_{xx}}\ln X\right)$	$z\left(1 - \frac{c\gamma_0}{\sigma_{xx}}\ln X\right)$
$z'c'$	$zc\left(1 - \frac{\gamma_0}{\sigma_{xx}}\ln X\right)$	$zc\left(1 - \frac{\gamma_0}{\sigma_{xx}}\ln X\right)$	$zc\left(1 - \frac{\gamma_0}{\sigma_{xx}}\ln X\right)$
X	$\mathcal{M}e^{\gamma_E - 1/2}$	$\zeta\lambda\mu_0 e^{\gamma_E-1/2}$	$\mu_0\mu^{-1}e^{-1/2}$

Integration over scale sizes λ

As before we can write the renormalized parameters $z(\zeta\lambda')$ and $z(\zeta\lambda')c(\zeta\lambda')$ as an integral over scale sizes. This leads to the more general expression for the observable theory

$$z'(\zeta\lambda') = z_0 + \int_{1/\mu_0 e^{\gamma_E}}^{\prime} \frac{d[\zeta\lambda]}{\zeta\lambda} zc\left(\frac{\gamma_0}{\sigma_{xx}} - \mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta\right), \quad (4.5.77)$$

$$z'(\zeta\lambda')c'(\zeta\lambda') = z_0 c_0 + \int_{1/\mu_0 e^{\gamma_E}}^{\prime} \frac{d[\zeta\lambda]}{\zeta\lambda} zc\left(\frac{\gamma_0}{\sigma_{xx}} - \mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta\right), \quad (4.5.78)$$

where the parameters z_0 and $z_0 c_0$ are defined for a fixed microscopic length scale $1/\mu_0 e^{\gamma_E}$. Again we compare the results with those obtained from the theory in dimensional regularization, Eqs. (3.2.20) and (3.2.21). This, then, completes the statement made earlier which says that the significance of the instanton scale size λ should primarily be found in the fundamental relation that exists between the observable theory on the one hand, and the renormalization group β and γ functions on the other. We will elaborate further on this point in Section 4.6.

4.6 The β' and γ' functions

4.6.1 Observable and renormalized theories

Next, introducing an arbitrary scale size λ_0 we can write Eqs. (4.5.45), (4.5.46), (4.5.74) and (4.5.75) as follows

$$\sigma'_{xx}(\zeta\lambda') = \sigma'_{xx}(\zeta\lambda_0) - \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda}\beta'_\sigma(\sigma_{xx},\theta,c), \tag{4.6.1}$$

$$\theta'(\zeta\lambda') = \theta'(\zeta\lambda_0) - \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda}\beta'_\theta(\sigma_{xx},\theta,c), \tag{4.6.2}$$

$$z'(\zeta\lambda') = z'(\zeta\lambda_0) + \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda} z'\gamma'_z(\sigma_{xx},\theta,c,c'), \tag{4.6.3}$$

$$z'(\zeta\lambda')c'(\zeta\lambda') = z'(\zeta\lambda_0)c'(\zeta\lambda_0) + \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda} z'c'\gamma'_{zc}(\sigma_{xx},\theta,c,c'), \tag{4.6.4}$$

where

$$\beta'_\sigma(\sigma_{xx},\theta,c) = -\frac{d\sigma'_{xx}}{d\ln\lambda} = \beta_0(c) + \mathcal{D}(c)\sigma_{xx}^2 e^{-2\pi\sigma_{xx}}\cos\theta, \tag{4.6.5}$$

$$\beta'_\theta(\sigma_{xx},\theta,c) = -\frac{d\theta'}{d\ln\lambda} = 2\pi\mathcal{D}(c)\sigma_{xx}^2 e^{-2\pi\sigma_{xx}}\sin\theta, \tag{4.6.6}$$

$$\gamma'_z(\sigma_{xx},\theta,c,c') = \frac{d\ln z'}{d\ln\lambda} = \frac{(1-c')c}{1-c}\Big[\frac{\gamma_0}{\sigma_{xx}} - \mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta\Big], \tag{4.6.7}$$

$$\gamma'_{zc}(\sigma_{xx},\theta,c,c') = \frac{d\ln z'c'}{d\ln\lambda} = \frac{(1-c')c}{(1-c)c'}\Big[\frac{\gamma_0}{\sigma_{xx}} - \mathcal{D}_\gamma(c)\sigma_{xx}e^{-2\pi\sigma_{xx}}\cos\theta\Big]. \tag{4.6.8}$$

The difference between the *observable* theory σ'_{xx}, θ', c' and z' and the *renormalized* theory σ_{xx}, θ, c and z can be expressed in terms of the renormalization group functions as follows

$$\beta_\sigma(\sigma_{xx},c) \Leftrightarrow \beta'_\sigma(\sigma_{xx},\theta,c), \tag{4.6.9}$$

$$\beta_\theta = 0 \Leftrightarrow \beta'_\theta(\sigma_{xx},\theta,c), \tag{4.6.10}$$

$$\gamma_z(\sigma_{xx},c) \Leftrightarrow \gamma'_z(\sigma_{xx},\theta,c,c'), \tag{4.6.11}$$

$$\gamma_{zc}(\sigma_{xx},c) \Leftrightarrow \gamma'_{zc}(\sigma_{xx},\theta,c,c'). \tag{4.6.12}$$

What clearly remains is to express the β'_σ, β'_θ, γ'_{zc} and γ'_z functions in terms of the observable parameters σ'_{xx}, θ' and c' alone, rather than the renormalized quantities σ_{xx}, θ and c. To ensure that this can be done without introducing any unwanted singularities into the problem we proceed as follows. It is important to notice, first of all, that the following general relations hold

$$\gamma_z(\sigma_{xx},c) = c\gamma_z(\sigma_{xx},c), \tag{4.6.13}$$

$$\gamma'_z(\sigma_{xx},\theta,c,c') = c'\gamma'_{zc}(\sigma_{xx},\theta,c,c'). \tag{4.6.14}$$

This means that both quantities $z\alpha$ and $z'\alpha'$ are unrenormalized as it should be. Next, we compare the renormalization behavior of the quantities c and c'

$$\beta_c(\sigma_{xx}, c) = -\frac{dc}{d\ln\lambda} = (1-c)c\gamma_{zc}(\sigma_{xx}, c), \quad (4.6.15)$$

$$\beta'_c(\sigma_{xx}, \theta, c, c') = -\frac{dc'}{d\ln\lambda} = (1-c')c'\gamma'_{zc}(\sigma_{xx}, \theta, c, c'). \quad (4.6.16)$$

We see that the Fermi liquid plane $c = c' = 0$ and the Coulomb interaction plane $c = c' = 1$ correspond to zero's of both the β_c and β'_c functions provided the γ'_{zc} is well behaved.

4.6.2 The β' and γ' functions

The relation between the observable and renormalized theories can be obtained by solving the following differential equations

$$\beta_\sigma(\sigma_{xx}, c)\frac{\partial\sigma'_{xx}}{\partial\sigma_{xx}} + \beta_c(\sigma_{xx}, c)\frac{\partial\sigma'_{xx}}{\partial c} = \beta'_\sigma(\sigma_{xx}, \theta, c), \quad (4.6.17)$$

$$\beta_\sigma(\sigma_{xx}, c)\frac{\partial\theta'}{\partial\sigma_{xx}} + \beta_c(\sigma_{xx}, c)\frac{\partial\theta'}{\partial c} = \beta'_\theta(\sigma_{xx}, \theta, c), \quad (4.6.18)$$

$$\beta_\sigma(\sigma_{xx}, c)\frac{\partial c'}{\partial\sigma_{xx}} + \beta_c(\sigma_{xx}, c)\frac{\partial c'}{\partial c} = \beta'_c(\sigma_{xx}, \theta, c, c'). \quad (4.6.19)$$

To obtain solutions that are meaningful in the entire range $0 \leqslant c \leqslant 1$ we must work with the two loop results for the β_σ function as in Eq. (3.2.22). It is next a matter of simple algebra to show that the results can be expressed in terms of an infinite double series in powers of $\exp(-2\pi\sigma'_{xx})$ and the trigonometric functions of θ'. The first few terms in the series can be written as follows

$$\beta'_\sigma(\sigma'_{xx}, \theta', c') = \left\{\beta_\sigma(\sigma'_{xx}, c') + F'_0 e^{-4\pi\sigma'_{xx}}\right\} + \left\{\mathcal{D}(c')\left(\sigma'_{xx}\right)^2 e^{-2\pi\sigma'_{xx}}\right\}\cos\theta' + \left\{F'_2 e^{-4\pi\sigma'_{xx}}\right\}\cos 2\theta' + \ldots, \quad (4.6.20)$$

$$\beta'_\theta(\sigma'_{xx}, \theta', c') = \left\{\mathcal{D}(c')\left(\sigma'_{xx}\right)^2 e^{-2\pi\sigma'_{xx}}\right\}\sin\theta' + \left\{F'_2 e^{-4\pi\sigma'_{xx}}\right\}\sin 2\theta' + \ldots, \quad (4.6.21)$$

$$\gamma'_{zc}(\sigma'_{xx}, \theta', c') = \left\{\gamma_{zc}(\sigma'_{xx}, c') + H'_0 e^{-4\pi\sigma'_{xx}}\right\} + \left\{\mathcal{D}_\gamma(c')\sigma'_{xx} e^{-2\pi\sigma'_{xx}}\right\}\cos\theta' + \left\{H'_2 e^{-4\pi\sigma'_{xx}}\right\}\cos 2\theta' + \ldots, \quad (4.6.22)$$

where the F', and H' are rational functions in σ'_{xx} and to leading order are given by

$$F'_0 = -\frac{(\sigma'_{xx})^4}{\beta_\sigma(\sigma'_{xx}, c')}\mathcal{D}^2(c'), \tag{4.6.23}$$

$$F'_2 = \frac{(\sigma'_{xx})^3}{2\pi\beta_\sigma(\sigma'_{xx}, c')}\mathcal{D}(c')\left[\mathcal{D}(c') + \frac{1}{2}c'(1-c')\mathcal{D}_\gamma(c')\partial_{c'}\mathcal{D}_\gamma(c')\right], \tag{4.6.24}$$

$$H'_0 = -\frac{(\sigma'_{xx})^3}{\beta_\sigma(\sigma'_{xx}, c')}\mathcal{D}_\gamma(c')\mathcal{D}(c'), \tag{4.6.25}$$

$$H'_2 = \frac{(\sigma'_{xx})^2}{4\pi\beta_\sigma(\sigma'_{xx}, c')}\mathcal{D}_\gamma(c')\left[\mathcal{D}(c') + \mathcal{D}_\gamma(c') + c'(1-c')\partial_{c'}\mathcal{D}_\gamma(c')\right]. \tag{4.6.26}$$

We see that the renormalization group β' and γ' functions are formally given as a sum over all topological sectors of the theory. This is in spite of the fact that we started out the computation with single instanton only. The final answer smoothly interpolates between the Coulomb interaction problem at $c = 1$ and the free electron theory $c = 0$ where the one-loop result $\beta_0(c = 0)$ of β_σ vanishes but the two-loop result $\beta_1(c = 0)$ is finite. We end this Section with several remarks. First of all, on the basis of Eq. (4.6.22) we obtain the desired result for the remaining renormalization group functions of Eqs. (4.6.14) and (4.6.16)

$$\gamma'_z(\sigma'_{xx}, \theta', c') = c'\gamma'_{zc}(\sigma'_{xx}, \theta', c'), \tag{4.6.27}$$

$$\beta'_c(\sigma'_{xx}, \theta', c') = (1-c')c'\gamma'_{zc}(\sigma'_{xx}, \theta', c'). \tag{4.6.28}$$

Since the γ'_{zc} is generally well behaved we have verified the statement which says that the combination $z'\alpha'$ is unrenormalized. At the same time we conclude that the planes $c' = 0$ and $c' = 1$ map unto themselves under the action of the renormalization group as it should be.

Secondly, it is clear that the terms with the exponential factors $\exp(-4\pi\sigma'_{xx})$ in Eqs. (4.6.20) and (4.6.22) correspond to higher order contributions that become important when *multi - instanton* configurations are taken into account. For example, the correction $O(\exp(-4\pi\sigma'_{xx}))$ in the first term of Eq. (4.6.20), i.e. $\{\beta_\sigma(\sigma'_{xx}, c') \quad + F'_0\exp(-4\pi\sigma'_{xx})\}$, clearly indicates that the topologically trivial vacuum is generally affected by *instanton* and *anti instanton* combinations. It is easy to see that the quantity $F'_0\exp(-4\pi\sigma'_{xx})$ actually arises from instantons and anti instantons that either are *widely separated* or have vastly *different scale sizes*. It therefore is natural to interpret the perturbative $\beta_\sigma(\sigma'_{xx}, c')$ function in terms of *tightly bound* instanton - anti instanton pairs with *equal sizes*. Similarly, the terms proportional to F'_2, G'_2 and H'_2 in Eqs. (4.6.20) and (4.6.22) are recognized as the *disconnected* pieces that appear in the contributions from instantons of topological charge ± 2. A complete multi-instanton analysis is beyond the scope of the present paper and likely involves the effects of merons. [117] Keeping these remarks in mind we conclude that the observable theory can be generally expressed as follows

$$
\begin{aligned}
\sigma'_{xx}(\zeta\lambda') &= \sigma'_{xx}(\zeta\lambda_0) - \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda}\beta'_\sigma(\sigma'_{xx}, \theta', c'), && (4.6.29)\\
\theta'(\zeta\lambda') &= \theta'(\zeta\lambda_0) - \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda}\beta'_\theta(\sigma'_{xx}, \theta', c'), && (4.6.30)\\
c'(\zeta\lambda') &= c'(\zeta\lambda_0) - \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda}\beta'_c(\sigma'_{xx}, \theta', c'), && (4.6.31)\\
z'(\zeta\lambda') &= z'(\zeta\lambda_0) + \int_{\zeta\lambda_0}^{\zeta\lambda'} \frac{d[\zeta\lambda]}{\zeta\lambda} z'\gamma'_z(\sigma'_{xx}, \theta', c'), && (4.6.32)
\end{aligned}
$$

$$
z'(\zeta\lambda')\alpha'(\zeta\lambda') = z'(\zeta\lambda_0)\alpha'(\zeta\lambda_0). \qquad (4.6.33)
$$

Here, β'_σ, β'_θ, β'_c and γ'_z are given to the appropriate order by Eqs. (4.6.20)-(4.6.22). For convenience of a reader we summarize the results for β'_σ, β'_θ, β'_c and γ'_z as follows

$$
\begin{aligned}
\beta'_\sigma &= \beta_0(c') + \frac{\beta_1(c')}{\sigma'_{xx}} + \mathcal{D}(c')\sigma'^2_{xx}e^{-2\pi\sigma'_{xx}}\cos\theta', && (4.6.34)\\
\beta'_\theta &= 2\pi\mathcal{D}(c')\sigma'^2_{xx}e^{-2\pi\sigma'_{xx}}\sin\theta', && (4.6.35)\\
\beta'_c &= (1-c')c'\left(-\frac{1}{\pi\sigma'_{xx}} - \frac{\beta^c_1(c')}{\sigma'^2_{xx}} - \mathcal{D}_\gamma(c')\sigma'_{xx}e^{-2\pi\sigma'_{xx}}\cos\theta'\right), && (4.6.36)\\
\gamma'_z &= c'\left(-\frac{1}{\pi\sigma'_{xx}} - \frac{\beta^c_1(c')}{\sigma'^2_{xx}} - \mathcal{D}_\gamma(c')\sigma'_{xx}e^{-2\pi\sigma'_{xx}}\cos\theta'\right). && (4.6.37)
\end{aligned}
$$

Here, we remind, one and two loop results are given as

$$
\begin{aligned}
\beta_0(c') &= \frac{2}{\pi}\left[1 + \frac{1-c'}{c'}\ln(1-c')\right],\\
\beta_1(c') &= \begin{cases} \dfrac{1}{2\pi^2} & c' = 0,\\ 4\pi^{-2}\Big[50 + \dfrac{1}{6} - 3\pi^2 + \dfrac{19}{2}\zeta(3) + 16\ln^2 2 & \\ \quad -44\ln 2 + \dfrac{\pi^2}{2}\ln 2 + 16G + \dfrac{\pi^4}{12} + \dfrac{\pi^2}{3}\ln^2 2 & c' = 1,\\ \quad -\dfrac{1}{3}\ln^4 2 - 7\zeta(3)\ln 2 - 8\,\mathrm{li}_4(1/2)\Big] & \end{cases}\\
\beta^c_1(c') &= \begin{cases} 0 & c' = 0,\\ \dfrac{3}{\pi^2} + \dfrac{1}{6} & c' = 1. \end{cases} && (4.6.38)
\end{aligned}
$$

The amplitudes of instanton contributions are given as

$$\begin{aligned}\mathcal{D}(c') &= 16\pi e^{1-2\pi\gamma_E\beta_0(c')} \exp\left\{-2\frac{1-c'}{c'}\left[\left[\psi\left(3-\frac{1}{c'}\right)+\psi\left(\frac{1}{c'}\right)-1\right]\ln(1-c')\right.\right.\\ &\qquad \left.\left. -f\left(\frac{1}{c'}-1\right)+f\left(\frac{1}{c'}\right)-\frac{2c'^2\ln 2}{2c'-1}\right]\right\}, \qquad (4.6.39)\\ \mathcal{D}_\gamma(c') &= \frac{\mathcal{D}(c')}{2c'}(1-c')^{1+2/c'}\int_0^{c'} ds\,(1-s)^{-2-2/s}, \qquad (4.6.40)\end{aligned}$$

where $f(z) = 2z^2\sum_{J=2}^{\infty} J^{-1}(J^2-z^2)^{-1}\ln J$. These final results generalize those previously obtained on the basis of perturbative expansions, Eqs. (3.2.19)-(3.2.21).

4.7 Effective action for the edge

Sofar we have completed the weak coupling analysis of the observable parameters of the theory that were introduced in Section 3.2.3. We have shown how the theory of ordinary radiative corrections can be extended to include the non-perturbatve effects of instantons. As an important general check on the consistency of the results we next focus the attention on the thermodynamic potential Ω of the electron gas. A major objective of this Section is to explicitly verify the statement $z'\alpha' = z\alpha$ at a non-perturbative level that until now we have assumed to be valid. This will be done in Sections 4.7.1 - 4.7.3 below where we address the various different steps that eventually lead to a systematic expansions of Ω in powers of T. A second major goal of this Section is to extend the effective action procedure of the massless chiral edge excitations to include the lowest order terms in T. The extended action can then be used to investigate whether and how the theory of topological excitations (instantons) facilitate a microscopic understanding of the quantum Hall regime. This will be the main topic of Sections 4.7.3 and 4.7.4 where we embark on the strong coupling consequences of our results.

4.7.1 Thermodynamic potential

We start out from the instanton contribution Ω_{inst} as given by Eq. (4.5.14)

$$\begin{aligned}\Omega_{\text{inst}} &= \left[\frac{Z_{\text{inst}}}{Z_0}\right]^{\text{reg}} = \frac{N^2}{4\pi^2}\,\text{Re}\int d\mathbf{r}_0 \int \frac{d\lambda}{\lambda^3}\mathcal{D}(c)\langle e^{S'_{\text{inst}}}\rangle_U\\ &= \frac{N^2}{4\pi^2}\int d\mathbf{r}_0\int\frac{d\lambda}{\lambda^3}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta\\ &\qquad \times\langle e^{S_F^{\text{inst}}[U]}\rangle_U. \qquad (4.7.1)\end{aligned}$$

Here, the sum over the instanton and anti-instanton sectors is understood. We have made use of the results for S'_{inst} as given by Eqs (4.5.38) and (4.5.40). The expression for the singlet interaction term $\hat{S}_F^{\text{inst}}[U]$ as given by Eq. (4.5.12) still contains the *spatially varying* momentum scale $\mu(\mathbf{r})$. Introducing the matrices $\hat{\Lambda}$ and $\hat{\mathbf{1}}$

$$\hat{\Lambda}^{\alpha\beta}_{nm} = \delta^{\alpha 1}\delta^{\beta 1}\delta_{nm}[\delta_{n0} - \delta_{n,-1}] \tag{4.7.2}$$
$$\hat{\mathbf{1}}^{\alpha\beta}_{nm} = \delta^{\alpha 1}\delta^{\beta 1}\delta_{nm}[\delta_{n0} + \delta_{n,-1}]. \tag{4.7.3}$$

then the expression for $\hat{S}_F^{\rm inst}[U]$ (Eq. (4.5.12)) can be written in a more compact fashion as follows

$$\hat{S}_F(U) = \hat{S}_i(U) + \hat{S}_\eta(U) \tag{4.7.4}$$

where

$$\begin{aligned}\hat{S}_i[U] &= -\frac{\pi}{2}T\lambda^2\left(\mathcal{A}_1 - \frac{5}{2}\mathcal{A}_2\right)\sum_{\alpha n}\operatorname{tr} I^\alpha_n U^{-1}\hat{\Lambda}U I^\alpha_{-n}U^{-1}\hat{\Lambda}U \\ &+ \frac{\pi}{2}T\lambda^2\left(\mathcal{A}_1 - \frac{1}{2}\mathcal{A}_2\right)\sum_{\alpha n}\operatorname{tr} I^\alpha_n U^{-1}\hat{\mathbf{1}}U I^\alpha_{-n}U^{-1}\hat{\mathbf{1}}U \end{aligned}\tag{4.7.5}$$

$$\hat{S}_\eta[U] = -4\pi T\lambda^2\mathcal{A}_3\operatorname{tr}\eta U^{-1}\hat{\Lambda}U. \tag{4.7.6}$$

The spatial integrals are all absorbed in the three different amplitudes $\mathcal{A}_i$ which are given by

$$\mathcal{A}_1 = \int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda}z(\mu(\mathbf{r}))c(\mu(\mathbf{r})) \tag{4.7.7}$$
$$\mathcal{A}_2 = \int' d\mathbf{r}\mu^2(\mathbf{r})z(\mu(\mathbf{r}))c(\mu(\mathbf{r})) \tag{4.7.8}$$
$$\mathcal{A}_3 = \int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda}z(\mu(\mathbf{r})). \tag{4.7.9}$$

Notice that $\mathcal{A}_1$ is identically the same as in Eq. (4.5.52). When written in this form it is of interest to compare the results of the quantum theory with those obtained in Section 4.3 for the classical theory. In this latter case one finds the same expressions as in Eqs (4.7.4) - (4.7.6) except that the $\mathcal{A}_i$ are replaced by the classical amplitudes $\mathcal{A}_i^{\rm inst}$

$$\mathcal{A}_1^{\rm inst} = zc\int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda} \tag{4.7.10}$$
$$\mathcal{A}_2^{\rm inst} = zc\int' d\mathbf{r}\mu^2(\mathbf{r}) \tag{4.7.11}$$
$$\mathcal{A}_3^{\rm inst} = z\int' d\mathbf{r}\frac{\mu(\mathbf{r})}{\lambda}. \tag{4.7.12}$$

We have already noticed in Section 4.3 that the classical results for $S_F^{\rm inst}[U]$, in particular the amplitudes $\mathcal{A}_1^{\rm inst}$ and $\mathcal{A}_3^{\rm inst}$, are logarithmically divergent in the sample size. Remarkably, however, these infrared troubles have disappeared in the quantum theory. By following a similar procedure as the one discussed in Section 4.5.5 we find that the amplitudes of Eqs (4.7.7) - (4.7.9) are generally finite. Whereas the result

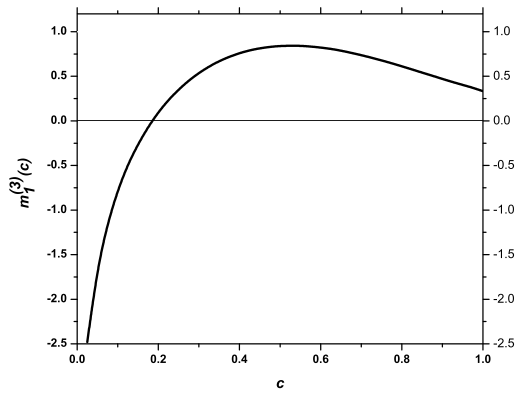


Figure 4.6: $m^{(3)}(c)$ versus c, see text.

for $\mathcal{A}_1$ is given by Eq. (4.5.73) the final expressions for the amplitudes $\mathcal{A}_2$ and $\mathcal{A}_3$ are obtained as follows

$$\mathcal{A}_2 = -2\pi z(\zeta\lambda) \tag{4.7.13}$$

$$\mathcal{A}_3 = -2\pi^2 z(\zeta\lambda)\sigma_{xx}(\zeta\lambda) m_1^{(3)}(c(\zeta\lambda)) \tag{4.7.14}$$

where

$$m_1^{(3)}(c) = \alpha \exp\Big[\frac{2}{c}\ln\alpha\Big] \int_0^c \frac{ds}{s(1-s)^2} \exp\Big[-\frac{2}{s}\ln(1-s)\Big]. \tag{4.7.15}$$

In Fig. 4.6 we plot of the function $m_1^{(3)}(c)$ with varying c. We see that $m_1^{(3)}(c)$ diverges as c tends to 0. This means that for $c = 0$ the leading term in $\mathcal{A}_3$ is proportional to σ_{xx}^2 rather than σ_{xx}. Furthermore, the results indicate that we can neglect the amplitude $\mathcal{A}_2$ which is one order in σ_{xx} smaller than $\mathcal{A}_1$ as well as $\mathcal{A}_3$. Keeping these remarks in mind we can write the instanton contribution to the thermodynamic potential as follows

$$\Omega_{\text{inst}} = \frac{N^2}{4\pi^2} \int d\mathbf{r}_0 \int' \frac{d\lambda}{\lambda^3} \mathcal{D}(c(\zeta\lambda)) e^{-2\pi\sigma_{xx}(\zeta\lambda)} \cos\theta \langle e^{\hat{S}_i(U)+\hat{S}_\eta(U)} \rangle_U \tag{4.7.16}$$

where $\hat{S}_\eta[U]$ is given by Eq. (4.7.6) and $\hat{S}_i[U]$ can be written as

$$\hat{S}_i[U] = -\frac{\pi}{4} T\lambda^2 \mathcal{A}_1 \sum_{\alpha n} \text{tr}\left([I_n^\alpha, U^{-1}\hat{\Lambda}U][I_{-n}^\alpha, U^{-1}\hat{\Lambda}U] - [I_n^\alpha, U^{-1}\hat{\mathbf{1}}U][I_{-n}^\alpha, U^{-1}\hat{\mathbf{1}}U] \right) \tag{4.7.17}$$

4.7.2 Expansion in T

Next, in a naive expansion of the thermodynamic potential Ω to lowest orders T one would proceed by computing the expectation of $\hat{S}_i(U)$ and $\hat{S}_\eta(U)$ with respect to the matrix U. In the limit where $N = N_r N_m \to 0$ this expectation is given by

$$\langle \hat{S}_i(U) \rangle_U = \frac{2\pi\lambda^2 T \mathcal{A}_1}{N^2} \operatorname{tr} \eta\Lambda \tag{4.7.18}$$

$$\langle \hat{S}_\eta(U) \rangle_U = -\frac{4\pi\lambda^2 T \mathcal{A}_3}{N} \operatorname{tr} \eta\Lambda. \tag{4.7.19}$$

Hence, only the quantity $\langle \hat{S}_i(U) \rangle_U$ survives in Eq. (4.7.16) whereas the term $\langle \hat{S}_\eta(U) \rangle_U$ vanishes in the limit where $N \to 0$. We have already mentioned, however, that the expectation of S_F, in particular Eq. (4.7.18), is complicated and cut-off dependent. These as well as other complications disappear once it is recognized that the frequency term $\hat{S}_\eta(U)$ in the action is actually not a perturbative quantity at all and should generally be retained in the exponential of Ω_{inst}. The correct series expansion in powers of T therefore has the following general form

$$\Omega_{\text{inst}} = \frac{N^2}{4\pi^2} \int d\mathbf{r}_0 \int' \frac{d\lambda}{\lambda^3} \mathcal{D}(c(\zeta\lambda)) e^{-2\pi\sigma_{xx}(\zeta\lambda)} \cos\theta \times \langle e^{\hat{S}_\eta(U)} (1 + \hat{S}_i(U) + \ldots) \rangle_U. \tag{4.7.20}$$

The problem that remains is to evaluate expectations of the type

$$\langle X \rangle_\epsilon = \langle X e^{-\epsilon \operatorname{tr} \eta U^{-1} \hat{\Lambda} U} \rangle_U \tag{4.7.21}$$

where we have written $\epsilon = 4\pi\lambda^2 T \mathcal{A}_3$. For our purposes the only expectations that we shall need are the following results which are valid in the limit $N \to 0$

$$\left\langle (U^{-1}\hat{\Lambda} U)^{\alpha\beta}_{nm} \right\rangle_\epsilon = \frac{1}{N} \Lambda^{\alpha\beta}_{nm} e^{-\epsilon |n|} \tag{4.7.22}$$

$$\left\langle (U^{-1}\hat{\mathbf{1}} U)^{\alpha\beta}_{nm} \right\rangle_\epsilon = \frac{1}{N} \mathbf{1}^{\alpha\beta}_{nm} e^{-\epsilon |n|}. \tag{4.7.23}$$

We see that the main effect of ϵ is to exponentially suppress the large Matsubara frequency components. To justify Eqs. (4.7.22) and (4.7.23) we proceed as follows. Since the averaging over positive and negative frequency blocks is independent of one another we first introduce for brevity the symbol $P^\alpha_n = (U^{-1}\hat{\mathbf{1}} U)^{\alpha\alpha}_{nn}$ where n is limited to, say, positive frequency indices only. Eqs (4.7.22) and (4.7.23) can then be expressed in terms of an infinite series expansion in powers of ϵ with coefficients of the type

$$\langle P^{\beta_1}_{m_1} \ldots P^{\beta_k}_{m_k} \rangle_U. \tag{4.7.24}$$

The lowest order coefficients we already have, in particular

$$\langle P^\alpha_{n_1} \rangle_U = \frac{1}{N}, \qquad \langle P^\alpha_{n_1} P^\beta_{n_3} \rangle_U = \frac{1}{N} \frac{\delta^{\alpha\beta}_{n_1 n_3} + 1}{1 + N}. \tag{4.7.25}$$

The second of these equations simplifies in the limit $N_r \to 0$ and can be replaced by the following expression

$$\langle P^{\alpha}_{n_1} P^{\beta}_{n_3} \rangle_U = \frac{1}{N} \delta^{\alpha\beta}_{n_1 n_3}. \tag{4.7.26}$$

Here we have neglected unity in comparison with $\delta^{\alpha\beta}_{n_1 n_3}$ due to the following reasons. In expansion of Eq. (4.7.23) the term of the first order in ϵ is given as

$$-\epsilon \sum_{m_1,\beta_1} m_1 \langle P^{\alpha}_{n_1} P^{\beta_1}_{m_1} \rangle_U = -\frac{\epsilon}{N} \left(n_1 + N_r \sum_{m_1} m_1 \right) \to -\frac{\epsilon}{N} n_1 \tag{4.7.27}$$

where the last term in the brackets in the first line of Eq. (4.7.27) have been left out because of being higher order in N_r and therefore insignificant. The result (4.7.27) corresponds exactly to the expression (4.7.26) for the average $\langle P^{\alpha}_{n_1} P^{\beta}_{n_3} \rangle_U$.

Proceeding along the same lines one can prove by induction that the general expression can be written as

$$\langle P^{\beta_1}_{m_1} \dots P^{\beta_k}_{m_k} \rangle_U = \frac{1}{N} \delta^{\beta_1 \dots \beta_k} \delta_{m_1 \dots m_k}. \tag{4.7.28}$$

Using this result one can re-exponentiate the series in powers of ϵ and the result can be written as follows

$$\left\langle P^{\alpha}_n \exp \left(-\epsilon \sum_{\beta, m>0} m P^{\beta}_m \right) \right\rangle = \frac{1}{N} \exp(-\epsilon |n|). \tag{4.7.29}$$

This, then, directly leads to the result of Eqs. (4.7.22) and (4.7.23).

On the basis of Eqs. (4.7.22) and (4.7.23) one can write the expectation $\langle \hat{S}_i(U) \rangle_\epsilon$ as follows

$$\langle \hat{S}_i(U) \rangle_\epsilon = -\frac{\pi}{2} T \lambda^2 \mathcal{A}_1 \sum_{\alpha n} \mathrm{tr} \left\langle [I^{\alpha}_n, U^{-1} \hat{\Lambda} U] \right\rangle_\epsilon \left\langle [I^{\alpha}_{-n}, U^{-1} \hat{\Lambda} U] \right\rangle_\epsilon. \tag{4.7.30}$$

Explicitly written out we now have

$$\langle \hat{S}_i(U) \rangle_\epsilon = \frac{2\pi \lambda^2 T \mathcal{A}_1}{N^2} \sum_{\alpha n} |n| e^{-\epsilon |n|} = \frac{2\pi \lambda^2 T \mathcal{A}_1}{N^2} \mathrm{tr} \left(\eta \Lambda e^{-\epsilon \eta \Lambda} \right). \tag{4.7.31}$$

It is important to emphasize that a finite value of ϵ not only suppresses the large frequency components in the sum over n but also makes the result of Eq. (4.7.31) independent of the arbitrary frequency cut-off N_m that one generally imposes on the *size* of matrices appearing in Eq. (4.7.30). Eq. (4.7.31) therefore resolves the aforementioned ambiguities in Eq. (4.7.18) that were obtained by putting ϵ strictly equal to zero.

4.7.3 Background field q_0

We are now in a position to carry the analysis one step further and evaluate the thermodynamic potential Ω_{inst} in the presence of a global *background matrix field*

$t_0 \in U(2N)/U(N)\times U(N)$. It is not difficult to see that this amounts to a replacement of Eq. (4.7.20) by the following expression

$$\Omega_{\rm inst}[t_0] = \frac{N^2}{4\pi^2}\int d\mathbf{r}_0 \int' \frac{d\lambda}{\lambda^3}\mathcal{D}(c(\zeta\lambda))e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta \times(1+\left\langle \hat{S}_i(Ut_0)\right\rangle_\epsilon + \dots) \quad (4.7.32)$$

where instead of Eq. (4.7.30) we now have

$$\langle \hat{S}_i(Ut_0)\rangle_\epsilon = -\frac{\pi}{2}T\lambda^2\mathcal{A}_1\sum_{\alpha n}\mathrm{tr}\left\langle [I_n^\alpha, t_0^{-1}U^{-1}\hat{\Lambda}Ut_0]\right\rangle_\epsilon \times\left\langle [I_{-n}^\alpha, t_0^{-1}U^{-1}\hat{\Lambda}Ut_0]\right\rangle_\epsilon . \quad (4.7.33)$$

The expectations are defined by Eq. (4.7.21) as before. According to the rules of $\mathcal{F}$ algebra [32] one should think of the matrix t_0 as being a "small" unitary rotation that is embedded in the space of much "larger" matrices of size $2N \times 2N$. More specifically, we take $t_0 \in U(2n)/U(n)\times U(n)$ with $n = N_r n_m$ being much "smaller" than $N = N_r N_m$. One is ultimately interested in the theory where the various different frequency cut-offs n_m, ϵ^{-1} and N_m are being sent off to infinity. This, however, should be done according to the general prescription $n_m \ll \epsilon^{-1} \ll N_m \to \infty$. [33]

Using Eqs. (4.7.22) and (4.7.23) we obtain Eq. (4.7.33) explicitly as follows

$$\langle \hat{S}_i(Ut_0)\rangle_\epsilon = -\frac{\pi\lambda^2 T\mathcal{A}_1}{2N^2}\Gamma[t_0] \quad (4.7.34)$$

where

$$\Gamma[t_0] = \sum_{\alpha n}\mathrm{tr}[I_n^\alpha, t_0^{-1}\Lambda e^{-\epsilon\eta\Lambda}t_0][I_{-n}^\alpha, t_0^{-1}\Lambda e^{-\epsilon\eta\Lambda}t_0]. \quad (4.7.35)$$

Next, by keeping the constraints on the different frequency cut-offs in mind and after some subtle but elementary $\mathcal{F}$ algebra [32] one finds the following simple expression for $\Gamma[t_0]$

$$\Gamma[t_0] = \Gamma[q_0] = \sum_{\alpha n} e^{-\epsilon|n|}\,\mathrm{tr}[I_n^\alpha, q_0][I_{-n}^\alpha, q_0] \quad (4.7.36)$$

where

$$q_0 = t_0^{-1}\Lambda t_0. \quad (4.7.37)$$

This final result is manifestly $U(N)\times U(N)$ invariant as it should be. Moreover, Eq. (4.7.36) is $\mathcal{F}$ *invariant* which means that it is invariant under electrodynamic $U(1)$ gauge transformations. [32] For all practical purposes we can replace the result of Eq. (4.7.36) by the more conventional expression [32]

$$\Gamma[q_0] = {\sum_{\alpha n}}' \mathrm{tr}[I_n^\alpha, q_0][I_{-n}^\alpha, q_0]. \quad (4.7.38)$$

Here, the $q_0 \in U(2n)/U(n)\times U(n)$ is embedded in a "large" matrix space of *fixed* size, say, $2\tilde{N}\times 2\tilde{N}$ with $\tilde{N} = N_r\tilde{N}_m$ and $\tilde{N}_m = \mathcal{O}(\epsilon^{-1})$. The prime on the summation sign indicates that the sum over n is restricted to $-\tilde{N}_m \lesssim n \lesssim \tilde{N}_m$. Both Eqs (4.7.36) and (4.7.38) can be re-expressed as a statement made on the space of "small" matrices alone according to [32]

$$\Gamma[q_0] \quad = \quad 2\left(\sum_{\alpha n} \operatorname{tr} I_n^\alpha q_0 \operatorname{tr} I_{-n}^\alpha q_0 + 4\operatorname{tr}\eta q_0 - 6\operatorname{tr}\eta\Lambda\right). \tag{4.7.39}$$

To summarize the results of this Section we write thermodynamic potential, Eq. (4.7.32), in the following manner

$$\Omega_{\text{inst}}[q_0] \quad = \quad \Omega_{\text{inst}} + S_{\text{eff}}^{\text{inst}}[q_0]. \tag{4.7.40}$$

Here, Ω_{inst} denotes the $T=0$ piece

$$\Omega_{\text{inst}} \quad = \quad \frac{N^2}{4\pi^2}\int d\mathbf{r}_0 \int' \frac{d\lambda}{\lambda^3}\mathcal{D}(c)e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta. \tag{4.7.41}$$

The $S_{\text{eff}}^{\text{inst}}$ denotes the $T \neq 0$ piece that contains the background matrix field q_0

$$S_{\text{eff}}^{\text{inst}}[q_0] \quad = \quad T(z'c')_{\text{inst}}\int d\mathbf{r}_0\Gamma[q_0] + \mathcal{O}(T^2) \tag{4.7.42}$$

where

$$(z'c')_{\text{inst}} \quad = \quad \frac{1}{2\pi}\int' \frac{d\lambda}{\lambda}\mathcal{D}(c)\mathcal{A}_1 e^{-2\pi\sigma_{xx}(\zeta\lambda)}\cos\theta \qquad . \tag{4.7.43}$$

Finally, to obtain the complete expression for the effective action in q_0 we have to add to the instanton expression of Eq. (4.7.42) the result $S_{\text{eff}}^0[q_0]$ corresponding to the topologically trivial vacuum. This is written in Eq. (4.2.30) and the final answer can be expressed as follows

$$S_{\text{eff}}[q_0] \quad = \quad S_{\text{eff}}^0[q_0] + S_{\text{eff}}^{\text{inst}}[q_0] \tag{4.7.44}$$

$$= \quad T(z'c')\int d\mathbf{r}_0 {\sum_{\alpha n}}' \operatorname{tr}[I_n^\alpha, q_0][I_{-n}^\alpha, q_0]$$

$$+ Tz\alpha\int d\mathbf{r}_0\left(4\operatorname{tr}\eta q_0 - 6\operatorname{tr}\eta\Lambda\right). \tag{4.7.45}$$

Here, $(z'c')$ contains the contributions from both perturbation theory and instantons whereas the quantity $z\alpha = z(1-c)$ is un-renormalized. We have therefore explicitly verified the T dependent pieces of the effective action for massless chiral edge excitations and the final answer retains the general form as obtained using the back field methodology, Eq. (4.2.30). Most importantly, we have verified the general statement made in the beginning of this paper which says that the quantity $z\alpha$ does not acquire any quantum corrections, neither from perturbative expansions nor from instantons.

4.7.4 Effective action

We next generalize the result of the previous Section and obtain the complete effective action for the *massless chiral edge excitations* $S_{\text{eff}}[q]$ for *spatially varying* matrix fields $q(\mathbf{r})$ rather than a *global* one q_0. By making use of Eq. (4.7.45) as well as Eq. (3.3.18) of Appendix 3.3 we obtain the final total expression for $S_{\text{eff}}[q]$ valid for $0 \leqslant c, c' \leqslant 1$ (cf. Eq. (4.2.44))

$$S_{\text{eff}}[q] = S_\sigma^{\text{edge}}[q] + S_\sigma^{\text{bulk}}[q] + S_F^{\text{bulk}}[q] \qquad (4.7.46)$$

where

$$S_\sigma^{\text{edge}}[q] = \frac{k(\nu_f)}{2} \oint dx \, \text{tr}\, t \nabla_x t^{-1} \Lambda \qquad (4.7.47)$$

$$\begin{aligned} S_\sigma^{\text{bulk}}[q] &= -\frac{\sigma'_{xx}}{8} \int d\mathbf{r}\, \text{tr}(\nabla q)^2 \\ &\quad + \frac{\theta'}{16\pi} \int d\mathbf{r}\, \text{tr}\, \epsilon_{jk} q \nabla_j q \nabla_k q. \end{aligned} \qquad (4.7.48)$$

On the other hand, we obtain directly from Eq. (4.7.45)

$$S_F^{\text{bulk}}[q] = S_i[q] + S_\eta[q] \qquad (4.7.49)$$

$$S_i[q] = \frac{1}{2}\pi T z' c' \int d\mathbf{r} \sum_{\alpha n}{}' \text{tr}[I_n^\alpha, q][I_{-n}^\alpha, q] \qquad (4.7.50)$$

$$S_\eta[q] = \pi T z' \alpha' \int d\mathbf{r}\, (4\, \text{tr}\, \eta q - 6\, \text{tr}\, \eta \Lambda) \qquad (4.7.51)$$

with $z'\alpha' = z\alpha$. Eqs (4.7.49) - (4.7.51) are the terms of order $\mathcal{O}(T)$ for arbitrary values of c that have been discarded in Eq. (4.2.25).

Let us next see how these results can be reconciled with the fixed point action for the quantum Hall state as formulated in Section 4.2.2. As illustrated in Fig. 5.1, our renormalization group results indicate that the strong coupling phases are controlled by the stable fixed points $\sigma'_{xx} = \theta' = 0$ in the $c = 1$ and $c = 0$ planes respectively. Our results indicate furthermore that for the $\mathcal{F}$ invariant theory $c = 1$ the singlet interaction amplitude z' scales to zero as well. Since the z' is completely analogous to the spontaneous magnetization in the classical Heisenberg ferromagnet, we conclude that the Coulomb interaction problem behaves in many ways like a conventional problem with a continuous symmetry. Our results are therefore completely consistent with the statements made in Section 4.2.2 which say that in the strong coupling quantum Hall phase the Coulomb interaction system renders insensitive to changes in the boundary conditions.

The situation is slightly different for the problem with finite ranged interactions $0 < c < 1$. The Fermi liquid fixed point $\sigma'_{xx} = \theta' = 0$ with $c = 0$ is stable in the infrared but the quantity $z\alpha$ in S_η is not renormalized and, hence, generally does not vanish. Eq. (4.7.51) therefore seems to be in conflict with the strong coupling phase where the action S_{eff} should generally become invariant under the replacement

$$q = t^{-1} \Lambda t \to t^{-1} T_0^{-1} \Lambda T_0 t \qquad (4.7.52)$$

for an arbitrary matrix field T_0 that reduces to a $U(N) \times U(N)$ gauge at the edge of the system. What happens in this case, however, is that the classical equations of motion for the matrix field variable q are going to change as one approaches the quantum Hall fixed point. For example, for $c = 0$ these equations of motion are

$$\frac{\sigma'_{xx}}{2} \nabla_j (q \nabla_j q) + \pi T z [\eta, q] = 0 \tag{4.7.53}$$

except at the edge of the system. As long as σ'_{xx} is large we can treat the term linear in T as a perturbation and the solution to Eq. (4.7.53) is then determined by value of the matrix field q at the edge of the system. This value is precisely the independent degree of freedom of the edge of the electron gas. Next, upon increasing the linear dimension of the system the value of σ'_{xx} eventually renders exponentially small and the term linear in T in Eq. (4.7.53) now becomes the dominating one. The solution to Eq. (4.7.53) is now given by $q = \Lambda$ in the bulk of the system. This, however, does not affect the matrix field q at the edge and Eq. (4.7.51) in the strong coupling phase therefore replaced by

$$S_\eta[q] \quad \rightarrow \quad \pi T z \alpha \left(-2 \int d\mathbf{r} \, \mathrm{tr} \, \eta \Lambda + \Delta \oint dx \, \mathrm{tr} \, \eta q \right) \tag{4.7.54}$$

with Δ denoting a phenomenological quantity that depends on the microscopic details of the edge.

4.8 Conclusions

In this chapter we have extended the perturbative theory of localization and interaction effects presented in chapter 3 to include the highly non-trivial effects of the θ term. The analysis that we have presented is an important technical as well as conceptual advance since it permits us to answer, for the first time, some of the long standing problems of the interacting electron gas on the strong coupling side.

We have seen, first of all, that the appearance of *massless chiral edge excitations* has important consequences for the low energy dynamics of the instanton vacuum and can be used, amongst many other things, to formulate a Thouless-like criterion for the quantum Hall effect. The introduction of an effective action for the edge excitations completely resolves the previously encountered ambiguities in the Kubo formulae and renormalization group such as the choice of boundary conditions, quantization of topological charge etc. The effective action procedure for edge excitations uniquely defines the *response parameters* or *physical observables* σ'_{xx} and θ'. Moreover, the differences between the *edge* excitations and *bulk* excitations fundamentally explain the various different aspects of symmetry in the problem such as *particle-hole* symmetry, *periodicity* in σ'_{xy} etc. Furthermore, the conditions for the quantum Hall effect can now quite generally be expressed by saying that $\sigma'_{xx} = \theta' = 0$ which means that the bulk of the system renders insensitive to changes in the boundary conditions. This generally happens when the bulk excitations of the system generate a mass gap.

These general statements have motivated us to develop, in the main part of this Chapter, a unified microscopic theory for the physical observables σ'_{xx} and θ' of the electron gas in the presence of electron-electron interactions. The complete list of observable parameters includes also the parameter c' which distinguishes between *finite range* electron-electron interactions ($0 < c' < 1$) and *infinite range* interactions ($c' = 1$), as well as the parameter z' which controls the temperature and frequency dependence of the system. The most important results of this Chapter are given by Eqs. (4.6.29) -(4.6.31) expressing how the observable parameters are related to the renormalization group β' and γ' functions of the theory (Eqs (4.6.34)-(4.6.37)). The closed set of renormalization group functions β'_σ, β'_θ and β'_c that we have obtained (Eqs. (4.6.20)-(4.6.22)) controls the low energy dynamics of the electron gas at $T = 0$ and zero external frequency.

4.A Quantum corrections of trivial vacuum in Pauli-Villars regularization

The σ_{xx} renormalization

In order to find the renormalization of the σ_{xx} conductivity we should compute the average in Eq. (4.2.31). Using the parameterization $Q = U_0^{-1}\mathcal{V}U_0$ with the global unitary matrix $U_0 \in U(N) \times U(N)$ and expanding the $\mathcal{V}$ to the second order in w, we obtain

$$\sigma'_{xx} = \sigma_{xx} + \frac{\sigma_{xx}^2}{2n}\int d\mathbf{r}\nabla^2\langle \operatorname{tr} I_n^\alpha v(\mathbf{r})\nabla v^\dagger(\mathbf{r}) \operatorname{tr} I_{-n}^\alpha v(\mathbf{r}')\nabla v^\dagger(\mathbf{r}')\rangle, \tag{4.A.1}$$

where a point $\mathbf{r}'$ can be chosen arbitrary since the averages are depend only on the difference of the coordinates. Now by going from (x, y) to (η, θ) coordinates and performing the averages, we find

$$\sigma'_{xx} = \sigma_{xx} - 4c\int\limits_{\eta\theta} O^{(0)}\int\limits_0^\infty d\omega\mathcal{G}_0(\omega;\eta\theta;\eta'\theta')\mathcal{G}_0^\alpha(\omega;\eta'\theta';\eta\theta), \tag{4.A.2}$$

where

$$\mathcal{G}_a^\alpha(\omega) = \frac{1}{(O^{(a)}+\omega)(O^{(a)}+\alpha\omega)} = \sum_{JM}\frac{|JM\rangle_{(a)(a)}\langle JM|}{(E_J^{(a)}+\omega)(E_J^{(a)}+\alpha\omega)}, \tag{4.A.3}$$

Integrating over η, θ and ω and introducing the Pauli-Villars masses as above, we leads to the following result

$$\begin{aligned}\sigma'_{xx} \;&=\; \sigma_{xx} - 2\pi\beta_0(c)\Big[\sum_{f=1}^K \hat{e}_f\sum_{J=0}^\infty \frac{E_J^{(0)}}{(E_J^{(0)}+\mathcal{M}_f^2)^2} + \sum_{J=1}^\infty \frac{1}{E_J^{(0)}}\Big] \\ &\quad\times\sum_{M=-J}^J \Phi_{JM}^{(0)}(\eta',\theta')\bar{\Phi}_{JM}^{(0)}(\eta',\theta').\end{aligned} \tag{4.A.4}$$

It is worth mentioning that the Jacobi polynomial $P_{J-M}^{M,M}(\eta)$ is proportional to the Gegenbauer polynomial $C_{J-M}^{M+1/2}(\eta)$. By using the summation theorem [98]

$$C_J^\lambda(\cos\phi\cos\phi' + z\sin\phi\sin\phi') = \frac{\Gamma(2\lambda-1)}{\Gamma^2(\lambda)}\sum_{M=0}^{J}\frac{2^{2M}\Gamma(J-M+1)}{\Gamma(J+M+2\lambda)}\Gamma^2(M+\lambda)$$
$$\times(2M+2\lambda-1)\sin^M\phi\sin^M\phi' C_{J-M}^{M+\lambda}(\cos\phi)C_{J-M}^{M+\lambda}(\cos\phi')C_M^{\lambda-1/2}(z) \quad (4.A.5)$$

with $z=1$ and $\lambda=1/2$, we find that the projection operator

$$\sum_{M=-J}^{J}\Phi_{JM}^{(0)}(\cos\phi,\theta)\bar{\Phi}_{JM}^{(0)}(\cos\phi',\theta) = \frac{2J+1}{4\pi}C_J^{1/2}\left(\cos(\phi-\phi')\right). \quad (4.A.6)$$

Since $C_J^{1/2}(1)=1$, we obtain

$$\sigma'_{xx} = \sigma_{xx} - \frac{\beta_0(c)}{2}\lim_{\Lambda\to\infty}\left[\sum_{J=3/2}^{\Lambda}\frac{2J(J^2-\frac{1}{4})}{(J^2-\frac{1}{4})^2} + \sum_{f=1}^{K}\hat{e}_f\sum_{J=1/2}^{\Lambda}\frac{2J(J^2-\frac{1}{4})}{(J^2-\frac{1}{4}+\mathcal{M}_f^2)^2}\right]. \quad (4.A.7)$$

Finally, evaluation of the sums above yields

$$\sigma'_{xx} = \sigma_{xx} - \frac{\beta_0(c)}{2}\left(Y_{\text{reg}}^{(0)}+1\right) = \sigma_{xx}\left(1-\frac{\beta_0(c)}{\sigma_{xx}}\ln\mathcal{M}e^{\gamma_E}\right). \quad (4.A.8)$$

The zc renormalization

Using the parameterization $Q = U_0^{-1}\mathcal{V}U_0$ with the global unitary matrix $U_0\in U(N)\times U(N)$ and expanding $\mathcal{V}$ to the second order in w, we find

$$z'c' = zc\left(1 - \frac{1}{\operatorname{tr}\eta\Lambda}\sum_{\alpha,n>0}\langle\operatorname{tr}I_n^\alpha v\operatorname{tr}I_{-n}^\alpha v^\dagger\rangle\right). \quad (4.A.9)$$

The averages yield

$$z'c' = zc\left(1+\frac{2\pi\gamma_0}{\sigma_{xx}}\mathcal{G}_0^c(0;\eta\theta;\eta\theta)\right) \quad (4.A.10)$$

By using Eq.(4.A.6), we find

$$z'c' = zc\left(1+\frac{\gamma_0}{2\sigma_{xx}}Y_{\text{reg}}^{(0)}\right). \quad (4.A.11)$$

With a help of the result (4.4.46) for the $Y_{\text{reg}}^{(0)}$ we finally obtain

$$z'c' = zc\left(1+\frac{\gamma_0}{\sigma_{xx}}\ln\mathcal{M}e^{\gamma_E-1/2}\right). \quad (4.A.12)$$

Chapter 5

Non-Fermi liquid criticality and super universality

5.1 Introduction

The results (4.6.34)-(4.6.37) indicates that the electron-electron interaction enhances the perturbative corrections to the β_σ function and suppresses the instanton contributions. Contrary to the case of non-interacting electrons, in the presence of the Coulomb interaction ($c' = 1$) one instanton approximation is not enough to produce the zero of the renormalization group β_σ function at $\theta' = \pi$. By adjusting the value of $\mathcal{D}(1)$ we can conjecture the following three dimensional renormalization group flow diagram as sketched in Fig. 5.1. We emphasize that this three dimensional renormalization group flow diagram is in agreement with experiments on plateau-plateau transitions in the quantum Hall regime. [54, 53] The regime of finite range electron-electron interactions $0 < c < 1$, like the theory in $2 + \epsilon$ dimensions, lies the domain of attraction of the Fermi liquid plane $c = 0$ which is stable in the infrared. These results are in accordance with the principle of $\mathcal{F}$ invariance which states the distinctly different problems of the Coulomb interaction $c = 1$ and finite range electron-electrons interactions $0 \leqslant c < 1$ are preserved separately under the action of the renormalization group.

5.2 Robust quantization of Hall conductance

We are now in a position to elaborate on the quantum Hall effect which is represented in Fig. 5.1 by the infrared fixed points located at precise values of $\sigma'_{xy} = k(\nu_f)$ or $\theta' = 0$ and $\sigma'_{xx} = 0$. For this purpose let us consider the renormalization group equations along the lines $\sigma'_{xy} \approx k(\nu_f)$ or $\theta' \approx 0$. Specializing to the most interesting

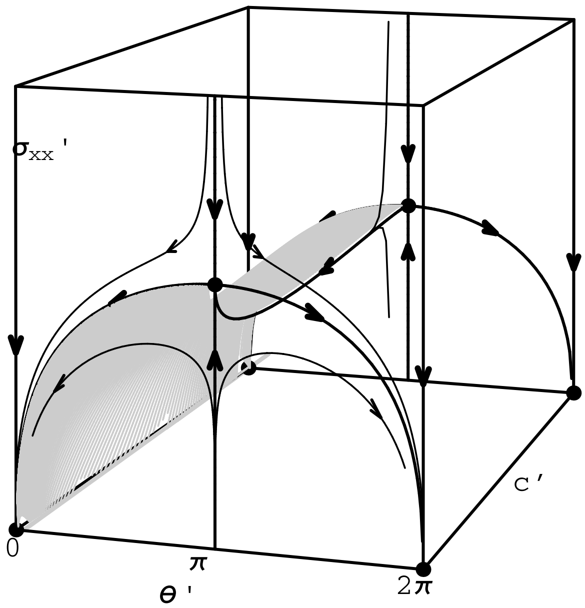


Figure 5.1: Sketch of renormalization group flow diagram in the parameter space σ'_{xx}, θ' and c'. The arrows indicate the direction toward the infrared

case $c = 1$ then we can write

$$\frac{d\ln\sigma'_{xx}}{d\ln\lambda} = \tilde{\beta}_\sigma(\sigma'_{xx}) = -\frac{2}{\pi\sigma'_{xx}} - \frac{\beta_1(1)}{(\sigma'_{xx})^2} - \mathcal{D}(1)\sigma'_{xx}e^{-2\pi\sigma'_{xx}}, \tag{5.2.1}$$

$$\frac{d\ln|\theta'|}{d\ln\lambda} = \tilde{\beta}_\theta(\sigma'_{xx}) = -2\pi\mathcal{D}(1)\sigma'^2_{xx}e^{-2\pi\sigma'_{xx}}. \tag{5.2.2}$$

These results are clearly consistent with the Thouless-like criterion presented in Section 4.2.2 of the previous Chapter which tells us that along the lines $\theta' \approx 0$ both quantities σ'_{xx} and θ' should vanish for large scale sizes λ. To see the meaning of the various different terms we recall from the discussion in Section 3.6 (see Chapter 3) that the perturbative $\tilde{\beta}_\sigma$ function naively indicates that the response parameter σ'_{xx} scales from $-(2/\pi)\ln(\lambda/\xi)$ for small values of λ to $\exp(-\lambda/\xi)$ for large values of λ. Here, ξ is the *dynamically generated* correlation length (localization length), see Eqs. (3.6.6) and (3.6.7). From Eq. (5.2.1) we see that the instanton contribution generally enhances the tendency of the electron gas to localize at large distances. In Fig. 5.2 we sketch the overall behavior of the $\tilde{\beta}_\sigma$ function which is given by the weak coupling result of Eq. (5.2.1) for large values of σ'_{xx} and the strong coupling result

$$\tilde{\beta}_\sigma = \ln\sigma'_{xx} \tag{5.2.3}$$

as σ'_{xx} goes to zero. These results for the dissipative conductance give rise to the well known scaling scenario of localization in two spatial dimensions. [12] However, Eq. (5.2.2) indicates that $|\theta'|$ decreases at a much slower (exponential) rate with

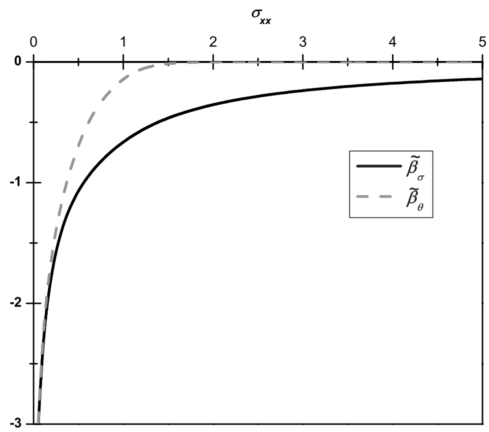


Figure 5.2: Sketches of the renormalization group functions $\tilde{\beta}_\sigma$ and $\tilde{\beta}_\theta/a$ for $\theta = 0$.

increasing values of λ indicating that the quantum Hall effect is generally confined to the regime of "bad conductors" $\sigma'_{xx} \lesssim 1$ only. The most important feature of the asymptotic expressions of Eqs. (5.2.1) and (5.2.2) is that they are universal in the sense that they are independent of the specific regularization scheme that defines the theory at short distances. On the other hand, the exact exponential form with which $|\theta'|$ vanishes in the strong coupling regime generally depends on the specific application of the instanton vacuum that one is interested in. The experiments on the quantum Hall effect, [53] for example, indicate that $\theta' \propto (\sigma'_{xx})^a$ with some positive value for the exponent a which is presumably equal to two. The same behavior has been found in closely related two dimensional models of the instanton vacuum. [48] Analogous to Eq. (5.2.3) one therefore expects that

$$\tilde{\beta}_\theta = a \ln \sigma'_{xx} \tag{5.2.4}$$

in the limit where σ'_{xx} goes to zero. In Fig. 5.2 we compare the scaling results for the Hall conductance $\tilde{\beta}_\theta$ with those for the longitudinal conductance $\tilde{\beta}_\sigma$. These scaling results are very similar for the free electron case $c = 0$ which means that the quantization phenomenon is a *(super) universal strong coupling* feature of the instanton vacuum concept, independent of the presence of electron-electron interactions.

5.3 Scaling results

5.3.1 Scaling equations for the observables

The physical idea that complicates the comparison between the theory presented in the previous Chapters and the experiment is that the *macroscopic* conductances measured at finite T can in general be very different from *ensemble averaged* conductances at $T = 0$ (σ'_{xx} and σ'_{xy}) that we have used for theoretical purposes. The concept of *conductance fluctuations* implies, for example, that a complete knowledge of the macroscopic conductances at finite T necessarily involves a detailed knowledge of the complete conductance *distributions* at $T = 0$. [119] The general scaling behavior of macroscopic quantities is obtained by using well known methods of quantum field theory (method of *characteristics*). [42, 93] To distinguish the *measured* conductances at finite T from the *ensemble averaged* quantities σ'_{xx} and θ' we denote the former as $G_{jk} = G_{xx}, G_{xy}$. Compared to the case of non-interacting electrons we have also to introduce the quantity U, the *measured* crossover parameter at finite T. Finally we introduce the *measured* "magnetization" M. They all obey the following differential equations

$$\left[-\beta_\sigma \frac{\partial}{\partial \sigma_{xx}} - \beta_\theta \frac{\partial}{\partial \theta} + (2+\gamma_z) Tz \frac{\partial}{\partial Tz} + (1-c)\gamma_z \frac{\partial}{\partial c}\right] G_{jk} = 0, \quad (5.3.1)$$

$$\left[-\beta_\sigma \frac{\partial}{\partial \sigma_{xx}} - \beta_\theta \frac{\partial}{\partial \theta} + (2+\gamma_z) Tz \frac{\partial}{\partial Tz} + (1-c)\gamma_z \frac{\partial}{\partial c}\right] U = 0, \quad (5.3.2)$$

$$\left[-\beta_\sigma \frac{\partial}{\partial \sigma_{xx}} - \beta_\theta \frac{\partial}{\partial \theta} + (2+\gamma_z) Tz \frac{\partial}{\partial Tz} + (1-c)\gamma_z \frac{\partial}{\partial c} + \gamma_z\right] M = 0. \quad (5.3.3)$$

It is worthwhile to mention that if we computed the temperature dependence of the *observable* parameters σ'_{jk}, c' and z' then the results would satisfy Eqs (5.3.1)-(5.3.3). Below we shall use the Eqs. (5.3.1)-(5.3.3) to make some predictions and comparisons with the experiments.

5.3.2 Scaling results in weak coupling regime, $\sigma_{xx} \gtrsim 1$

Recently, it became possible to study the instanton effects experimentally. In Ref. [55] the magneto resistance data taken from samples with heavily Si-doped GaAs layers at low temperatures have been investigated. However, the comparison of the experimental data with the theory developed above has been performed in a wrong way. The main physical idea that has been overlooked in Ref. [55] is that the *macroscopic* conductances G_{jk} are in general very different from zero temperature *ensemble averaged* conductances σ'_{jk}. In what follows we shall specialize to the Coulomb interaction case $c' = 1$ which was realized in the experiment.

In the weak coupling regime $\sigma_{xx} \gg 1$ we can find the solutions of Eq. (5.3.1) in the following form

$$G_{xx} \approx g_0 - g_c \cos\theta, \quad (5.3.4)$$

$$G_{xy} \approx \frac{\theta}{2\pi} - g_s \sin\theta, \quad (5.3.5)$$

$$M \approx m_0(h_0 - h_c \cos\theta), \quad (5.3.6)$$

where the functions $g_{0,c,s}$ and $h_{0,c}$ depend on T only through the scaling variable X

$$X = zT\xi^2 m_0. \quad (5.3.7)$$

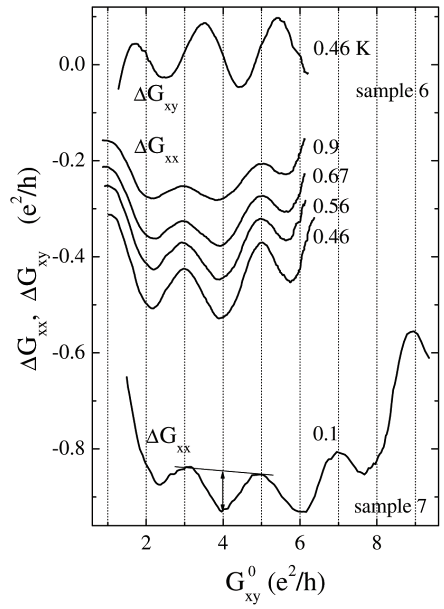


Figure 5.3: Variations $\Delta G_{jk} = G_{jk}(T) - G_{jk}(T_0)$ for $T_0 = 4.2K$ as a functions of $G^0_{xy} \equiv G_{xy}(T_0)$. The digits near curves indicate temperature T. After Ref. [55].

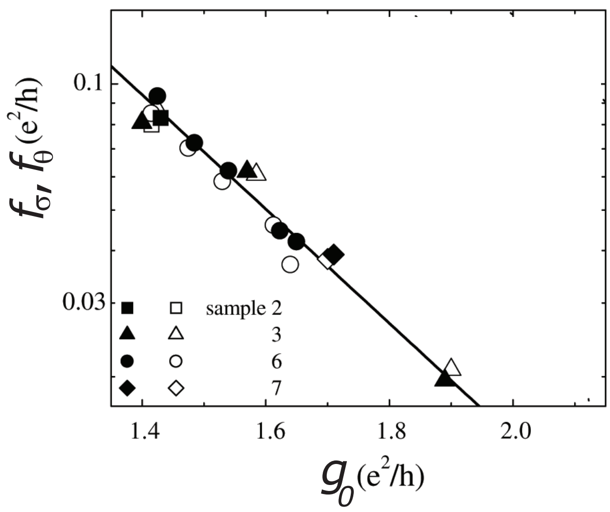


Figure 5.4: The dependence of functions $f_{\sigma,\theta}$ on g_0 for four samples. The solid curve is $7.6 \exp(-2\pi g_0)$. After Ref. [55].

Here, the correlation length ξ and magnetization m_0 are defined in Eqs. (3.6.6) and (3.6.11) respectively. The functions $g_0 = g_0(X)$ and $h_0 = h_0(X)$ are arbitrary regular functions of X whereas

$$g_c = [f_\sigma(g_0(X)) - \pi f_\infty(\sigma_{xx})] X \frac{dg_0(X)}{dX}, \quad (5.3.8)$$

$$g_s = f_\theta(g_0(X)) - f_\infty(\sigma_{xx}), \quad (5.3.9)$$

$$h_c = [f_M(g_0(X)) - \pi f_\infty(\sigma_{xx})] X \frac{dh_0(X)}{dX} \quad (5.3.10)$$

with $f_{\sigma,\theta,M}$ being unspecified regular functions of the variable X and

$$f_\infty(\sigma_{xx}) = \left[\frac{\mathcal{D}(1)}{4}\sigma_{xx}^2 + \mathcal{O}(\sigma_{xx})\right] e^{-2\pi\sigma_{xx}}. \quad (5.3.11)$$

In the limit $\sigma_{xx} \gg 1$ one expects that (see Eqs. (3.6.7) and (3.6.12))

$$g_0(X) \approx \frac{1}{\pi}\ln X, \quad (5.3.12)$$

$$h_0(X) \approx \left(\frac{1}{\pi}\ln X\right)^{1/2}, \quad (5.3.13)$$

but no first principle computation of $f_{\sigma,\theta,M}(g_0)$ exists now. Experimentally [55] $G_{xx} \approx g_0 = \mathcal{O}(1)$ such that both the convergence of the series in σ_{xx}^{-1} (see Chapter 3 and the occurrence of broad conductance distributions [119] now complicate the problem. These complications are clearly reflected by the fact that the $g_0(X)$ data in the range of experimental temperatures [55] do not truly display the asymptotic behavior in X given by Eq. (5.3.12). Subtracting Eq. (5.3.5) for two different temperatures T and T_0 we find for the Hall conductance

$$G_{xy}(T) \approx G_{xy}(T_0) - [f_\theta(g_0(T)) - f_\theta(g_0(T_0))] \sin[2\pi G_{xy}(T_0)]. \quad (5.3.14)$$

Assuming that the function $g_0(X)$ is determined by the asymptotic expression (5.3.12) and applying the same trick as above for Eq. (5.3.4) we obtain

$$G_{xx}(T) \approx G_{xx}(T_0) - [f_\sigma(g_0(T)) - f_\sigma(g_0(T_0))] \cos[2\pi G_{xy}(T_0)]. \quad (5.3.15)$$

Now if we choose temperature T_0 be relatively high such that the oscillations in $G_{xx}(T_0)$ are absent then $g_0(T_0) \approx G_{xx}(T_0)$. Eqs (5.3.14) and (5.3.15) are very general results that is predicted by the theory.

In Ref. [55] the magneto resistance has been measured in relatively narrow range of $g_0 = 1.4 - 1.9$. The results (5.3.14) and (5.3.15) have been confirmed and it was found that (see Figures 5.3 and 5.4)

$$f_\sigma(g_0) \approx 7.6 e^{-2\pi g_0}, \qquad f_\theta(g_0) \approx 7.6 e^{-2\pi g_0}. \quad (5.3.16)$$

It is worthwhile to mention that the exponential dependence $\exp(-2\pi g_0)$ observed in experiment [55] is typical for the instanton effects.

We mention that Eq. (5.3.6) can be used to extract the temperature behavior of the specific heat of the electron system. We use the expression (see Eq. (3.5.9))

$$c_v = \frac{\partial}{\partial T} \int_0^\infty d\omega \frac{\omega \rho_{qp}(\omega)}{e^{\omega/T} - 1} \tag{5.3.17}$$

where the density of states of bosonic quasiparticles given as

$$\rho_{qp}(\omega) \propto \mathrm{Re}[h_0(i\omega X/T) - h_c(i\omega X/T)\cos\theta]. \tag{5.3.18}$$

Hence, we obtain that

$$c_v \propto T\,[H_0(X) - H_c(X)\cos\theta] \tag{5.3.19}$$

where

$$H_{0,c}(X) = \mathrm{Re} \int_0^\infty du \frac{u^2}{\sinh^2(u/2)} h_{0,c}(iuX) \tag{5.3.20}$$

We emphasize that the specific heat of electrons should *oscillate* as a function of θ.

At $\sigma_{xx} \gg 1$ we find $H_0(X) \propto (\ln X)^{1/2}$ that results in the following non-Fermi liquid behavior of the specific heat

$$c_v \propto T\sqrt{|\ln T|}. \tag{5.3.21}$$

5.3.3 Plateau transitions in the quantum Hall regime. Short-ranged interaction ($c' < 1$)

The most important features of the strong coupling regime ($\sigma'_{xx} \sim 1$) are the *quantum critical* fixed points that are located at $\theta = \pi$ or half-integer values of σ_{xy}. Fig. 5.1 indicates that the *Fermi liquid* fixed point located at $c' = 0$ is distinctly different from the *Coulomb interaction* fixed point at $c' = 1$. Like the *mobility edge* problem in $2+\epsilon$ dimensions considered in Chapter 3, the quantum critical behavior of the *transitions* between adjacent quantum Hall plateaus is very different for finite range electron-electron interactions and the Coulomb potential, each involving different exponent values as well as a fundamentally different dynamical behavior.

We start the scaling analysis from the Fermi liquid fixed point at $c' = 0$, $\theta' = \pi$ and $\sigma'_{xx} = \sigma^\star_{xx} \approx 0.88$ (see Fig. 5.1). Near the Fermi-liquid fixed point ($c' \ll 1$) it is convenient to introduce the following variables

$$\Delta\theta = \theta - \pi, \qquad \Delta\sigma = \frac{\sigma_{xx} - \sigma^\star_{xx}}{\sigma^\star_{xx}} + \frac{x_c}{y_\sigma - \gamma^\star_{zc}} c \tag{5.3.22}$$

where the critical exponents y_σ, $\gamma^\star_{zc}$ and quantity x_c can be formally obtained from the β_σ, β_θ and γ_{zc} functions at the Fermi liquid fixed point as

$$y_\sigma = -\left.\frac{\partial \beta_\sigma}{\partial \sigma_{xx}}\right|_{FP} \approx -0.17, \tag{5.3.23}$$

$$\gamma^\star_{zc} = \left.\gamma_{zc}\right|_{FP} \approx -0.36, \tag{5.3.24}$$

$$x_c = -\left.\frac{\partial \beta_\sigma}{\partial c}\right|_{FP} \approx -0.19. \tag{5.3.25}$$

Thence the scaling equations (5.3.1)-(5.3.3) becomes

$$\left[y_\sigma \Delta\sigma \frac{\partial}{\partial \Delta\sigma} + \nu^{-1}\Delta\theta \frac{\partial}{\partial \Delta\theta} + 2Tz\frac{\partial}{\partial Tz} + \left(\frac{2}{p} - 2\right) c\frac{\partial}{\partial c}\right] G_{jk} = 0, \quad (5.3.26)$$

$$\left[y_\sigma \Delta\sigma \frac{\partial}{\partial \Delta\sigma} + \nu^{-1}\Delta\theta \frac{\partial}{\partial \Delta\theta} + 2Tz\frac{\partial}{\partial Tz} + \left(\frac{2}{p} - 2\right) c\frac{\partial}{\partial c}\right] U = 0, \quad (5.3.27)$$

$$\left[y_\sigma \Delta\sigma \frac{\partial}{\partial \Delta\sigma} + \nu^{-1}\Delta\theta \frac{\partial}{\partial \Delta\theta} + 2Tz\frac{\partial}{\partial Tz} + \left(\frac{2}{p} - 2\right) c\frac{\partial}{\partial c}\right] M = 0. \quad (5.3.28)$$

Here we have introduce two more exponents

$$\nu^{-1} = -\left.\frac{\partial \beta_\theta}{\partial \theta}\right|_{FP} \approx 0.36, \quad (5.3.29)$$

$$p = \frac{2}{2+\gamma^\star_{zc}} \approx 1.22. \quad (5.3.30)$$

Next from the method of characteristics we obtain the scaling behavior as follows

$$\begin{aligned} G_{jk} &= F_{jk}\left(X, Y, Z\right), \\ U &= F_U(X, Y, Z), \\ M &= F_M(X, Y, Z). \end{aligned} \quad (5.3.31)$$

Here F_{jk}, F_U and F_M are unspecified regular functions, whereas

$$X = (zcT)^{-\kappa}\Delta\theta, \qquad \kappa = \frac{p}{2\nu} \approx 0.22, \quad (5.3.32)$$

denotes the *relevant* scaling variable and

$$Y = (zcT)^{\mu_\sigma}\Delta\sigma, \qquad \mu_\sigma = -\frac{p y_\sigma}{2} \approx 0.10, \quad (5.3.33)$$

$$Z = (zcT)^{\mu_c} c, \qquad \mu_c = p - 1 \approx 0.22, \quad (5.3.34)$$

are the *irrelevant* ones. It is worthwhile to mention that temperature involves in the scaling relations (5.3.31) only in the combination zcT. This *substantial* fact can be obtained on the basis of the microscopic theory (Finkelstein non-linear sigma model) only. On the basis of a phenomenological approach alone [120] one would choose the combination X/Z^β with arbitrary exponent β for the *relevant* variable. The main fault of Ref. [120] was that the value $\beta = (1-\gamma^\star_{zc})/2\nu$ was taken on the basis of plausible phenomenological arguments whereas the microscopic theory implies that $\beta = 0$!

The *relevant* scaling variable X can be written in the following way

$$X = \left(\frac{L_\phi}{\xi}\right)^{1/\nu} \quad (5.3.35)$$

where $\xi \propto |\Delta\theta|^{-\nu}$ is the divergent correlation length and

$$L_\phi \propto (zcT)^{-p/2} \quad (5.3.36)$$

is the dephasing length introduced by the electron-electron interaction.

Eqs.(5.3.31)-(5.3.34) imply that the "magnetization" is constant at the Fermi-liquid fixed point, M = const. Thence, the quasiparticle density of states $\rho_{qp}(\omega) =$ const and we obtain the Fermi-liquid scaling behavior for the specific heat

$$c_v \propto T. \tag{5.3.37}$$

The best estimates for the exponent values known from numerical works [76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89] lie in the range $\nu = 2.30 - 2.38$, $p = 1.22 - 1.48$ and $y_\sigma = -(0.34 - 0.42)$. It yields the following estimates for the Fermi liquid exponents

$$\kappa = 0.29 \pm 0.04, \tag{5.3.38}$$
$$\mu_\sigma = 0.26 \pm 0.05, \tag{5.3.39}$$
$$\mu_c = 0.35 \pm 0.15 \tag{5.3.40}$$

that are *clearly in conflict* with the experimental results obtained recently for the critical exponents in the plateau-plateau and plateau-insulator transitions [53, 54]

$$\kappa = 0.42 \pm 0.01, \qquad \mu_\sigma = 2.5 \pm 0.5. \tag{5.3.41}$$

Therefore we can conclude that the present experimental results on the scaling [53, 54] fall into the non-Fermi liquid universality class. The results obtained above completely invalidate the attempt made in Ref. [54] to explain the experimentally observed exponent value $\kappa = 0.42 \pm 0.01$ on the basis of phenomenological Fermi liquid type of ideas introduced in the field long ago. [22]

5.3.4 Plateau transitions in the quantum Hall regime. Coulomb interaction ($c' = 1$)

We next focus on the consequences of the non-Fermi liquid fixed point located at $c' = 1$, $\theta' = \pi$ and $\sigma'_{xx} = \sigma^\star_{xx}$ in Fig. 5.1. Unfortunately, the instantons analysis carried out in the previous chapter are not enough to produce the quantitative estimates for the critical exponents in the non-Fermi liquid fixed point. We perform therefore the scaling analysis based on the general grounds.

The non-perturbative renormalization group equations (4.6.34)-(4.6.37) suggest to introduce the following variables near non-Fermi liquid fixed point at $c' = 1$

$$\Delta\theta = \theta - \pi, \qquad \Delta\sigma = \frac{\sigma_{xx} - \sigma^\star_{xx}}{\sigma^\star_{xx}} + \frac{x_c}{y_\sigma + \gamma^\star_{zc}}\zeta, \qquad \zeta = (1-c)\ln(1-c) \tag{5.3.42}$$

where the critical exponents y_σ, $\gamma^\star_{zc}$ and quantity x_c can be formally obtained from the β_σ, β_θ and γ_{zc} functions at the Fermi liquid fixed point as

$$y_\sigma = -\frac{\partial\beta_\sigma}{\partial\sigma_{xx}}\bigg|_{FP}, \tag{5.3.43}$$

$$\gamma_{zc}^\star = \gamma_{zc}\bigg|_{FP}, \tag{5.3.44}$$

$$x_c = -\frac{\partial\beta_\sigma}{\partial\zeta}\bigg|_{FP}. \tag{5.3.45}$$

Thence the scaling equations (5.3.1)-(5.3.3) becomes

$$\left[y_\sigma\Delta\sigma\frac{\partial}{\partial\Delta\sigma}+\nu^{-1}\Delta\theta\frac{\partial}{\partial\Delta\theta}+\frac{2}{p}Tz\frac{\partial}{\partial Tz}+\left(2-\frac{2}{p}\right)\zeta\frac{\partial}{\partial\zeta}\right]G_{jk} = 0, \tag{5.3.46}$$

$$\left[y_\sigma\Delta\sigma\frac{\partial}{\partial\Delta\sigma}+\nu^{-1}\Delta\theta\frac{\partial}{\partial\Delta\theta}+\frac{2}{p}Tz\frac{\partial}{\partial Tz}+\left(2-\frac{2}{p}\right)\zeta\frac{\partial}{\partial\zeta}\right]U = 0, \tag{5.3.47}$$

$$\left[y_\sigma\Delta\sigma\frac{\partial}{\partial\Delta\sigma}+\nu^{-1}\Delta\theta\frac{\partial}{\partial\Delta\theta}+\frac{2}{p}Tz\frac{\partial}{\partial Tz}+\left(2-\frac{2}{p}\right)\zeta\frac{\partial}{\partial\zeta}\right]M = \left(2-\frac{2}{p}\right)M. \tag{5.3.48}$$

Here we have introduce two more exponents

$$\nu^{-1} = -\frac{\partial\beta_\theta}{\partial\theta}\bigg|_{FP}, \tag{5.3.49}$$

$$p = \frac{2}{2+\gamma_{zc}^\star}. \tag{5.3.50}$$

Next from the method of characteristics we obtain the scaling behavior as follows

$$\begin{aligned} G_{jk} &= F_{jk}(X,Y,Z), \\ U &= F_U(X,Y,Z), \\ M &= (zT)^{\mu_c}F_M(X,Y,Z). \end{aligned} \tag{5.3.51}$$

Here F_{jk}, F_U and F_M are unspecified regular functions, whereas

$$X = (zT)^{-\kappa}\Delta\theta, \qquad \kappa = \frac{p}{2\nu}, \tag{5.3.52}$$

$$Y = (zT)^{-\mu_c}\zeta, \qquad \mu_c = p-1 \tag{5.3.53}$$

denotes the *relevant* scaling variables and

$$Z = (zT)^{\mu_\sigma}\Delta\sigma, \qquad \mu_\sigma = -\frac{py_\sigma}{2}, \tag{5.3.54}$$

is the *irrelevant* one. It is worthwhile to mention that now temperature involves in the scaling relations (5.3.51) only in the combination zT.

In the case of Coulomb interaction when $c' = 1$ from the outset the only *relevant* scaling variable that remains is the X. It can be written also as

$$X = \left(\frac{L_\phi}{\xi}\right)^{1/\nu} \tag{5.3.55}$$

where $\xi \propto |\Delta\theta|^{-\nu}$ is the divergent correlation length and

$$L_\phi \propto (zT)^{-p/2} \tag{5.3.56}$$

is the dephasing length introduced by the Coulomb interaction. We mention that the results for the divergent correlation length and dephasing length are very similar (except the precise values of critical exponents) to what we had in the Fermi-liquid fixed point.

As one can see from Eqs.(5.3.51)-(5.3.54) at the non-Fermi liquid fixed point the "magnetization" scales with temperature as

$$M \propto (zT)^{\mu_c}. \tag{5.3.57}$$

Thence, the quasiparticle density of states

$$\rho_{qp}(\omega) \propto \omega^{\mu_c} \tag{5.3.58}$$

that yields the non-Fermi liquid behavior of the specific heat

$$c_v \propto T^p. \tag{5.3.59}$$

We qualify it as a non-Fermi liquid behavior because the exponent $p > 1$. The result (5.3.59) implies that the physical observable, associated with the "dephasing length" exponent p in the quantum Hall systems, is none other than the *specific heat* of the electron gas. A measurement of c_v should therefore provide the ultimate test on the consistency of the theory. This information is not present as of yet.

5.4 Conclusions

To understand the fundamental differences between Fermi liquid theory and the Coulomb interaction problem in the quantum Hall regime, a deeper understanding of the theory in $2+\epsilon$ spatial dimensions is absolutely essential. As we see in Chapter 3 the main reason is that the mobility edge problem in $2+\epsilon$ dimensions is the only place where the various different aspects of dynamical scaling of the electron gas can be established and evaluated explicitly. This includes not only the theory of quantum transport but also fundamental aspects such as the *specific heat*, the *multifractal* singularity spectrum of the electron gas [118] etc. that one usually do not probe in the experiments on the quantum Hall effect. In fact, the long standing problems associated with the theory of electron-electron interactions have in many ways turned out to be an outstanding laboratory for advanced methods in quantum field theory that cannot be studied in any different manner.

Finally, we can say that the results of this and previous Chapter explain why the scaling behavior of the non-interacting electron gas or electron gas with short-ranged interaction and the Coulomb interaction problem in strong magnetic fields looks so similar. In spite of the fact that the underlying theories are fundamentally different they have nevertheless important features in common such as asymptotic freedom, instantons, massless edge excitations etc. Since in both cases the topological concepts are the same it is natural to expect that the basic phenomena are the same, in particular the existence of *robust topological quantum numbers* that explain the observability and precision of the quantum Hall effect, as well as *quantum criticality* at $\theta = \pi$ that generally facilitates a *transition* to take place between different quantum Hall plateaus. Finally, by recognizing the fact that quantum Hall physics actually reveals itself as a generic, *super universal* feature of the instanton vacuum in asymptotically free field theory one has essentially laid the foundation for a more ambitious *unifying* theory that includes - besides integral quantum Hall physics - also the scaling behavior of the *abelian* quantum Hall states.

References

[1] For review see *The Quantum Hall effect*, edited by R. E. Prange and S. M. Girvin, (Springer-Verlag, Berlin, 1987).

[2] For review see T. Chakraborty and P. Pietiläinen, *The Fractional Quantum Hall Effect*, (Springer-Verlag, Berlin, 1988).

[3] L. D. Landau and E. M. Lifshitz, *Course in Theoretical Physics*, Vol. 8 (Pergamon Press, Oxford, 1984).

[4] A. A. Abrikosov, *Fundamentals of the theory of metals*, North-Holland, Amsterdam (1988).

[5] T. Ando, A. B. Fowler, and F. Stern, *Electronic properties of two-dimensional systems*, Rev. Mod. Phys. **54**, 437 (1982).

[6] K. von Klitzing, G. Dorda, and M. Pepper, *New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance*, Phys. Rev. Lett. **45**, 494 (1980).

[7] D. C. Tsui, H. L. Störmer, and A. C. Gossard, *Two-dimensional magnetotransport in the extreme quantum limit*, Phys. Rev. Lett. **48**, 1559 (1982).

[8] K. von Klitzing, *Nobel lecture: The quantized Hall effect*, Rev. Mod. Phys. **58**, 519 (1986).

[9] R. B. Laughlin, *Nobel lecture: Fractional quantization*, Rev. Mod. Phys. **71**, 863 (1999).

[10] H. L. Störmer, *Nobel lecture: The fractional quantum Hall effect*, Rev. Mod. Phys. **71**, 875 (1999).

[11] D. C. Tsui, *Nobel lecture: Interplay of disorder and interaction in two-dimensional electron gas in intense magentic field*, Rev. Mod. Phys. **71**, 891 (1999).

[12] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, *Scaling theory of localization: Absence of quantum diffusion in two dimensions*, Phys. Rev. Lett. **42**, 673 (1979).

[13] F. Wegner, *Disordered system with n orbitals per site: $n = \infty$ limit*, Phys. Rev. B **19**, 783 (1979); *The mobility edge problem: Continuous symmetry and conjecture*, Z. Phys. B **35**, 207 (1979).

[14] L. Schaefer and F. Wegner, *Disordered System with n Orbitals per Site: Lagrange Formulation, Hyperbolic Symmetry, and Goldstone Modes*, Z. Phys. B **38**, 113 (1980).

[15] K. B. Efetov, A. I. Larkin, D. E. Khemel'nitzkii, *Interaction between diffusion modes in localization theory*, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980)[Sov. Phys. JETP **52**, 568 (1980)].

[16] A. J. McKane and M. Stone, *Localization as an alternative to Goldstone's theorem*, Ann. Phys. (N.Y.) **131**, 36 (1981).

[17] A. M. M. Pruisken and L. Schaefer, *Field theory and the Anderson model for disordered electronic systems* , Phys. Rev. Lett. **46**, 490 (1981); *The Anderson model for electron localisation, non-linear σ model, asymptotic gauge invariance*, Nucl. Phys. B **200**, 20 (1982).

[18] H. Levine, S. Libby, and A. M. M. Pruisken, *Electron delocalization by a magnetic field in two dimensions*, Phys. Rev. Lett. **51**, 1915 (1983).

[19] A. M. M. Pruisken, *On localization in the theory of the quantized Hall effect: a two-dimensional realization of the θ-vacuum*, Nucl. Phys. B **235**, 277 (1984).

[20] A. M. M. Pruisken, *Quasiparticles in the theory of the integral quantum Hall effect (I)*, Nucl. Phys. B **285**, 719 (1987).

[21] A. M. M. Pruisken, *Quasiparticles in the theory of the integral quantum Hall efect (II). Renormalization of the Hall conductance or instanton angle theta*, Nucl. Phys. B **290**, 61 (1987).

[22] A. M. M. Pruisken, *Universal singularities in the integral quantum Hall effect*, Phys. Rev. Lett. **61**, 1297 (1988).

[23] H. P. Wei, D. C. Tsui, M. A. Palaanen, A. M. M. Pruisken, *Experiments on delocalization and universality in the integral quantum Hall effect*, Phys. Rev. Lett. **61**, 1294 (1988).

[24] B. Huckestein, *Scaling theory of the integer quantum Hall effect*, Rev. Mod. Phys. **67**, 357 (1995).

[25] A. D. Mirlin, *Statistics of energy levels and eigenfunctions in disordered systems*, Phys. Rep. 326 259 (2000).

[26] R. B. Laughlin, *Anomalous quantum Hall effect: An incompressible quantum fluid with fractionally charged excitations*, Phys. Rev. Lett. **50**, 1395 (1983).

[27] J. K. Jain, *Composite-fermion approach for the fractional quantum Hall effect*, Phys. Rev. Lett. **63**, 199 (1989).

[28] S. H. Simon, *The Chern-Simons Fermi-liquid description of fractional quantum Hall states*, arXiv: cond-mat/9812186.

[29] X. G. Wen, *Chiral Luttinger liquid and the edge excitations in the fractional quantum Hall states*, Phys. Rev. B **41**, 12838 (1990); *Theory of the edge states in fractional quantum Hall effects*, Int. J. Mod. Phys. B **6**, 1711 (1992).

[30] A.M. Finkelstein, *Electron liquid in disordered conductors*, vol. 14 of Soviet Scientific Reviews, ed. by I.M. Khalatnikov, Harwood Academic Publishers, London, (1990).

[31] A. M. M. Pruisken and M. A. Baranov, *Cracking Coulomb interactions in the quantum Hall regime*, Europhys. Lett. **31**, 543 (1995).

[32] A. M. M. Pruisken, M. A. Baranov, and B. Škorić, *Mishandling gauge invariance in the theory of the QHE (I)*, Phys. Rev. B **60**, 16807 (1999).

[33] M. A. Baranov, A. M. M. Pruisken, and B. Škorić, *Mishandling gauge invariance in the theory of the QHE (II)*, Phys. Rev. B **60**, 16821 (1999).

[34] A. M. M. Pruisken, B. Škorić, and M. A. Baranov, *Mishandling gauge invariance in the theory of the QHE (III)*, Phys. Rev. B **60**, 16838 (1999).

[35] B. Škorić and A. M. M. Pruisken, *The fractional quantum Hall effect: Chern-Simons mapping, duality, Luttinger liquids and the instanton vacuum*, Nucl. Phys. B **559**, 637 (1999).

[36] A. D'Adda, P. Di Vecchia, and M. Luescher, *A $1/n$ expandable series of non-linear σ models with instantons*, Nucl. Phys. B **146**, 63 (1978).

[37] E. Witten, *Instantons, the quark model, and the $1/N$ expansion*, Nucl. Phys. B **149**, 285 (1979).

[38] I. Affleck, *The role of instantons in scale-invariant gauge theories*, Nucl. Phys. B **162**, 461 (1980); *The role of instantons in scale-invariant gauge theories : (II). The short-distance limit*, Nucl. Phys. B **171**, 420 (1980).

[39] I. Affleck, *The quantum Hall effects, σ-models at $\theta = \pi$ and quantum spin chains*, Nucl. Phys. B **257**, 397 (1985).

[40] I. Affleck, *Exact critical exponents for quantum spin chains, non-linear σ-models at $\theta = \pi$ and the quantum Hall effect*, Nucl. Phys. B **265**, 409 (1986).

[41] I. Affleck, *Critical behaviour of $SU(n)$ quantum chains and topological non-linear σ-models*, Nucl. Phys. B **305**, 582 (1988).

[42] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, (Clarendon Press, Oxford, 1993).

[43] E. Brézin and J. Zinn-Justin, *Spontaneous breakdown of continuous symmetries near two dimensions*, Phys. Rev. B **14**, 3110 (1976).

[44] G. 't Hooft, *Computation of the quantum effects due to a four-dimensional pseudoparticle*, Phys. Rev. D **14**, 3432 (1976).

[45] I. Affleck, *On constrained instantons*, Nucl. Phys. B **191**, 429 (1981); M. Nielsen and N. K. Nielsen, *Explicit construction of constrained instantons*, Phys. Rev. D **61**, 105020 (2000) and references therein.

[46] A. M. M. Pruisken, M. A. Baranov, and M. Voropaev, *The large N theory exactly reveals the quantum Hall effect and θ renormalization*, arXiv: cond-mat/0101003 (unpublished).

[47] A. M. M. Pruisken, I. S. Burmistrov, and R. Shankar, *Massless excitations at $\theta = \pi$ in the CP^{N-1} model with large values of N*, arXiv: cond-mat/0602653.

[48] A. M. M. Pruisken, R. Shankar and N. Surendran, *General topological features and instanton vacuum in quantum Hall and spin liquids*, Phys. Rev. B **72**, 035329 (2005).

[49] V. Ambegaokar, U. Eckern, G. Schön, Phys. Rev. Lett. **48**, 1745 (1982).

[50] I. S. Burmistrov and A. M. M. Pruisken, in preparation.

[51] J. P. Eisenstein and H. L. Störmer, *The fractional quantum Hall effect*, Science **248**, 1510 (1990).

[52] B. Karmakar, M. R. Gokhale, A. P. Shah, B. M. Arora, D. T. N. de Lang, A. de Visser, L. A. Ponomarenko, A. M. M. Pruisken, *The effects of macroscopic inhomogeneities on the magnetotransport properties of the electron gas in two dimensions*, Physica E **24**, 187 (2004).

[53] A. M. M. Pruisken, D. T. N. de Lang, L. A. Ponomarenko, A. de Visser, *Quantum criticality, particle-hole symmetry, and duality of the plateau-insulator transition in the quantum Hall regime*, arXiv: cond-mat/0109043 (unpublished).

[54] W. Li, G. A. Csáthy, D. C. Tsui, L. N. Pfeiffer, and K. W. West, *Scaling and universality of integer quantum Hall plateau-to-plateau transitions*, Phys. Rev. Lett. **94**, 206807 (2005).

[55] S. S. Murzin, A. G. M. Jansen, and I. Claus, *Topological oscillations of the magnetoconductance in disordered GaAs layers*, Phys. Rev. Lett. **92**, 016802 (2004); S. S. Murzin and A. G. M. Jansen, *Murzin and Janssen reply*, Phys. Rev. Lett. **95**, 189702 (2005).

[56] F. Wegner, *Four-loop-order β function of nonlinear σ-models in symmetric spaces*, Nucl. Phys. B **316**, 663 (1989).

[57] S. Hikami, *Lecture notes of NATO ASI on interference phenomena in mesoscopic systems*, **254**, p.429, ed. B. Kramer, (Plenum, N.Y. 1990).

[58] C. Castellani, C. DiCastro, P. A. Lee, and M. Ma, *Interaction-driven metal-insulator transitions in disordered fermion systems*, Phys. Rev. B **30**, 527 (1984).

[59] C. Castellani, C. DiCastro, P. A. Lee, M. Ma, S. Sorella, and E. Tabet, *Enhancement of the spin susceptibility in disordered interacting electrons and metal-insulator transition*, Phys. Rev. B **33**, 6169 (1986).

[60] T. R. Kirkpatrick and D. Belitz, *The Anderson-Mott transition*, Rev. Mod. Phys. **66**, 261 (1994).

[61] A. Kamenev and A. Andreev, *Electron-electron interactions in disordered metals: Keldysh formalism*, Phys. Rev. B **60**, 2218 (1999).

[62] C. Chamon, A. W. W. Ludwig, and C. Nayak, *Schwinger-Keldysh approach to disordered and interacting electron systems: Derivation of Finkelstein's renormalization-group equations*, Phys. Rev. B **60**, 2239 (1999).

[63] R. T. F. van Schaijk, A. de Visser, S. M. Olsthoorn, H. P. Wei, and A. M. M. Pruisken, *Probing the plateau-insulator quantum phase transition in the quantum Hall regime*, Phys. Rev. Lett. **84**, 1567 (2000).

[64] D. T. N. de Lang, L. A. Ponomarenko, A. de Visser, C. Possanzini, S. M. Olsthoorn, and A. M. M. Pruisken, *Evidence for a quantum Hall insulator in an InGaAs/InP heterostructure*, Physica E **12**, 666 (2002).

[65] L. A. Ponomarenko, D. T. N. de Lang, A. de Visser, D. Maude, B. N. Zvonkov, R. A. Lunin and A. M. M. Pruisken, *New insights into the plateau-insulator transition in the quantum Hall regime*, Physica E **22**, 236 (2004).

[66] A. M. M. Pruisken, M. A. Baranov, and M. Voropaev, *The θ vacuum reveals itself as the fundamental theory (I)*, arXiv: cond-mat/0206011 (unpublished).

[67] S. Coleman, *Aspects of Symmetry*, University Press, Cambridge (1989).

[68] See e.g. O. Heinonen (Ed.), *Composite Fermions*, World Scientific (1998).

[69] D. J. Thouless in: R. Balian, R. Maynard, and G. Toulouse (Eds), *Ill-condensed Matter*, North-Holland/World Scientific, 1978, p.1.

[70] H. Weidenmüller, *Single electron in a random potential and a strong magnetic field*, Nucl. Phys. B **290**, 87 (1987); S. Xiong, N. Read, and A. Douglas Stone, *Mesoscopic conductance and its fluctuations at a nonzero Hall angle*, Phys. Rev. B **56**, 3982 (1997).

[71] See e.g. A. B. Zamolodchikov and Al. B. Zamolodchikov, *Massless factorized scattering and sigma models with topological terms*, Nucl. Phys. B **379**, 602 (1992).

[72] H. Levine, S. Libby, and A. M. M. Pruisken, *Theory of the quantized Hall effect (I)*, Nucl. Phys. B **240**, 30 (1984); *Theory of the quantized Hall effect (II)*, *ibid* **240**, 49 (1984); *Theory of the quantized Hall effect (II)*, **240**, 71 (1984) and references therein.

[73] C. Castellani and L. Peliti, *Multifractal wavefunction at the localisation threshold*, J. Phys. A **19**, L429 (1986); F. Wegner, *Anomalous dimensions for the nonlinear sigma-model, in $2+\epsilon$ dimensions (I)*, Nucl. Phys. B **280**, 193 (1987); *Anomalous dimensions for the nonlinear sigma-model, in $2+\epsilon$ dimensions (II)*, *ibid* **280**, 210 (1987).

[74] F. Wegner, *Inverse participation ratio in $2+\epsilon$ dimensions*, Z. Phys. B **36**, 209 (1979).

[75] A. M. M. Pruisken, *Participation ratio in the nonlinear σ model representation of localization*, Phys. Rev. B **31**, 416 (1985).

[76] J. T. Chalker and P. D. Coddington, *Percolation, quantum tunnelling and the integer Hall effect*, J. Phys. C **21**, 2665 (1988).

[77] B. Huckestein and B. Kramer, *One-parameter scaling in the lowest Landau band: Precise determination of the critical bahavior of the localization length*, Phys. Rev. Lett. **64**, 1437 (1990).

[78] B. Mieck, *Nonuniversal delocalization in a strong magnetic field*, Europhys. Lett. **13**, 453 (1990).

[79] Y. Huo and R. N. Bhatt, *Current carrying states in the lowest Landau level*, Phys. Rev. Lett. **68**, 1375 (1992).

[80] T. Ando, *Edge and bulk Landau states in quantum Hall regime*, J. Phys. Soc. Japan **61**, 415 (1992).

[81] D.-H. Lee, Z. Wang and S. Kivelson, *Quantum percolation and plateau transitions in the quantum Hall effect*, Phys. Rev. Lett. **70**, 4130 (1993).

[82] J. T. Chalker and C. J. Daniell, *Scaling, diffusion, and the integer quantized Hall effect*, Phys. Rev. Lett. **61**, 593 (1988).

[83] B. Huckestein and L. Schweitzer, *Relation between the correlation dimensions of multifractal wave functions and spectral measures in integer quantum Hall systems*, Phys. Rev. Lett. **72**, 713 (1994).

[84] S. Hikami, *Localization length and inverse participation ratio of two dimensional electron in the quantized Hall effect*, Prog. Theor. Phys. **76**, 1210 (1986).

[85] D.-H. Lee, *Transitions between Hall plateaux and the dimerization transition of a Hubbard chain*, Phil. Mag. Lett. **73**, 145 (1996).

[86] W. Pook and M. Janßen, *Multifractality and scaling in disordered mesoscopic systems* Z. Phys. B **82**, 295 (1991).

[87] B. Huckestein, B. Kramer, L. Schweitzer, *Characterization of the electronic states near the centres of the Landau bands under quantum Hall conditions*, Surf. Sci. **263**, 125 (1992).

[88] B. Huckestein, *Corrections to scaling in the integer quantum Hall effect*, Phys. Rev. Lett. **72**, 1080 (1994).

[89] F. Evers, A. Mildenberger, and A.D. Mirlin, *Multifractality of wave functions at the quantum Hall transition revisited*, Phys. Rev. B **64**, 241303 (2001).

[90] A. M. M. Pruisken, *The integral quantum Hall effect: Shortcomings of conventional localization theory*, Nucl. Phys. B **295**, 653 (1988).

[91] B. L. Altshuler, V. E. Kravtsov, and I. V. Lerner, in *Mesoscopic Phenomena in Solids*, ed. B. L. Altshuler, P. A. Lee, and R. A. Webb, North Holland, New York, p. 449 (1991).

[92] E. Brézin, S. Hikami and J. Zinn-Justin, *Generalized non-linear σ-models with gauge invariance*, Nucl. Phys. B **165**, 528 (1980).

[93] A. M. M. Pruisken and Z. Wang, *Some critical properties of G/H non-linear sigma models near two dimensions: (I). Universal scaling law for the coupling constant*, Nucl. Phys. B **322**, 721 (1989).

[94] W. Pauli and F. Villars, *On the invariant regularization in relativistic quantum theory*, Rev. Mod. Phys. **21**, 434 (1949).

[95] N. Mermin and H. Wagner, *Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models.*, Phys. Rev. Lett. **17**, 1133 (1966).

[96] D. Höf and F. Wegner, *Calculation of anomalous dimensions for the nonlinear sigma model*, Nucl. Phys. B **275**, 561 (1986).

[97] R. Shankar and N. Read, *The $\theta = \pi$ nonlinear sigma model is massless*, Nucl. Phys. B **336**, 457 (1990).

[98] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 4th ed., Academic Press (1980)

[99] H. P. Wei, D. C. Tsui, and A. M. M. Pruisken, *Localization and scaling in the quantum Hall regime*, Phys. Rev. B **33**, 1488 (1985).

[100] D. G. Polyakov and B. I. Shklovskii, *Activated Conductivity in the Quantum Hall Effect*, Phys. Rev. Lett. **73**, 1150 (1994).

[101] S. Koch, R. J. Haug, K. von Klitzing, and K. Ploog, *Direct measurement of critical exponents in the quantum Hall regime*, Surf. Sci. **263**, 108 (1992).

[102] M. Furlan, *Electronic transport and the localization length in the quantum Hall effect*, Phys. Rev. B **57**, 14818 (1998).

[103] F. Hohls, U. Zeitler, and R. J. Haug, *High frequency conductivity in the quantum Hall regime*, Phys. Rev. Lett. **86**, 5124 (2001).

[104] S.-R. E. Yang, A. H. MacDonald, and B. Huckestein, *Interactions, localization, and the integer quantum Hall effect*, Phys. Rev. Lett. **74**, 3229 (1995).

[105] B. Huckestein and M. Backhaus, *Integer quantum Hall effect of interacting electrons: dynamical scaling and critical conductivity*, Phys. Rev. Lett. **82**, 5100 (1999).

[106] A. M. M. Pruisken, M. A. Baranov, and I. S. Burmistrov, arXiv: cond-mat/0104387 (unpublished).

[107] E. Brézin, J. Zinn-Justin, and J. C. Le Gouillou, *Renormalization of the nonlinear σ model in $2+\epsilon$ dimensions*, Phys. Rev. D **14**, 2615 (1976).

[108] A. M. M. Pruisken and M. A. Baranov, unpublished.

[109] C. Castellani and C. Di Castro, *Effective Landau theory for disordered interacting electron systems: Specific-heat behavior*, Phys. Rev. B **34**, 5935 (1986).

[110] A. L. Efros and B. I. Shklovskii, *Coulomb gap and low temperature conductivity of disordered systems*, J. Phys. C **8**, L49 (1975).

[111] A. M. Polyakov, *Gauge fields and Strings*, (Harwood academic publishers, Shur, 1987).

[112] M. R. Zirnbauer, *Another critique of the replica trick*, arXiv: cond-mat/9903338 (unpublished).

[113] M. R. Zirnbauer, *Towards a theory of the integer quantum Hall transition: from the nonlinear sigma model to superspin chains*, Ann. d. Physik **3**, 513 (1994).

[114] J. J. M. Verbaarschot and M. R. Zirnbauer, *Critique of the replica trick*, J. Phys. A **18**, 1093 (1985).

[115] M. Luesher and B. Berg, *Computation of quantum fluctuations around multi-instanton fields from exact Green's functions: the CP^{n-1} case*, Commun. Math. Phys. **69**, 57 (1979).

[116] V .A. Fateev, I. V. Frolov, and A. S. Schwartz, *Quantum fluctuations of instantons in the nonlinear σ model*, Nucl. Phys. B **154**, 1 (1979).

[117] C. G. Callan, R. F. Dashen, and D. J. Gross, *A theory of hadronic structure*, Phys. Rev. D **19**, 1826 (1979); *Toward a theory of the strong interactions*, **17**, 2717 (1978); *A mechanism for quark confinement*, Phys. Lett. B **66**, 375 (1977).

[118] M. Voropaev, and A. M. M. Pruisken in preparation.

[119] M. H. Cohen and A. M. M. Pruisken, *Mesoscopic block models for macroscopic conductances*, Phys. Rev. B **49**, 4593 (1994).

[120] Z. Wang, M.P.A. Fisher, S.M. Girvin and J.T. Chalker, *Short-range interactions and scaling near integer quantum Hall transition*, Phys. Rev. B **61**, 8326 (2000).

Summary

The original ideas of an instanton parameter θ in the theory of the quantum Hall effect go back as far as twenty years ago. In spite of more than two decades of research, the subject matter is still far from being finished. New discoveries have been made, in particular the existence of massless chiral edge excitations, which clearly demonstrate that the quantum Hall effect is a generic feature of the θ vacuum in asymptotically free field theory in *general*. These discoveries have furthermore led to a complete revision of certain prevailing ideas in the field, in particular those regarding the meaning of instantons, the replica method etc. These ideas have primarily been motivated by incorrect historical results on the large N expansion of the CP^{N-1} model that for many years have been dramatically mistaken for "exact results".

In this thesis I revisit the instanton methodology in the theory of the quantum Hall effect and extend this methodology in many ways. The results of this thesis can generally be regarded as an integral part of a new concept which states the basic aspects of the quantum Hall effect are *super universal* features of the instanton angle θ in scale invariant theories.

The main physical objective of this thesis is quantum criticality of the disordered electron gas in strong magnetic fields, both in the presence and absence of electron-electron interactions. To establish the general meaning of θ renormalization by instantons I first consider the case of electrons without interactions and next the case of the interacting electron gas but without a strong magnetic field. Finally I consider the combined effects of electron-electron interactions and strong magnetic fields.

In Chapter 2 I study the Grassmanian $U(m+n)/U(m) \times U(n)$ non-linear sigma model in two dimensions. This theory in the limit $m = n = 0$ is known to describe the low energy dynamics of free electrons in the presence of a random potential and strong magnetic fields. This theory furthermore permits general discussions on the topological significance of the massless chiral edge excitations in the problem as well as the closely related topic of *physical observables* or *conductances* σ_{xx} and σ_{xy}. The remainder of this Chapter is devoted to the technical details of instanton calculus and the specific manner in which the non-perturbative contributions to the renormalization group β and γ functions of the theory can be extracted. The detailed predictions of the instanton methodology for the theory with $m = n = 0$ are next being compared with the numerical exponent values obtained from computer simulations on the free electron gas. A remarkable agreement is found. The results of this Chapter provides the theoretical platform for Chapter 4 where I address the problem of electron-electron interactions.

In Chapter 3 I turn to the theory with Coulomb interaction introduced in Refs [32, 33, 34]. The technical details are being presented of the mobility edge problem in a systematic expansion in $2+\epsilon$ dimensions to order ϵ^2. The complete non-Fermi liquid scaling behavior is obtained for the conductivity and specific heat of the electron gas with varying temperature, external frequency as well as electron density.

Finally, in Chapters 4 and 5 I make use of the advances reported in all the previous Chapters and embark on the main topic of this thesis which is the θ dependence of the general unifying theory incorporating the effects of electron-electron interactions. Technically speaking, the Chapter 4 represents one of the most interesting applications of the instanton methodology in general, and the procedure of *spatially varying masses* in particular. The reason being that the (singlet) interaction term, unlike ordinary mass terms, completely alters the ultraviolet singularity structure of the theory which is just one of the many peculiar features of the problem with electron-electron interactions. In spite of the herculean efforts that are needed to control the extraordinary amount of technical and analytical detail, what is truly remarkable is that an elegant simplicity is shining through at the end of all computations. The most important results of this Chapter are represented by a three dimensional renormalization group flow diagram with the conductance parameters σ_{xx} and σ_{xy} along two of the axes and only a single parameter c, representing the *range* of the electron-electron interaction, in the third direction. To understand the physical meaning of this flow diagram and, in particular, the non-Fermi liquid behavior of the Coulomb interaction problem, one is furthermore helped by the underlying symmetries such as $\mathcal{F}$ invariance. Moreover, by using the standard principles of the renormalization group one immediately extracts the general scaling forms for the conductances from this diagram with varying temperature, external frequency and magnetic field. These scaling forms are the principle objectives for experimental research on quantum criticality in the quantum Hall regime.

In summary, by employing the quantum Hall effect as a laboratory for investigating and exploring the instanton angle θ, important new insights have emerged that could not have been discovered otherwise. I mention in particular the various different aspects of *super universality* of the θ vacuum concept that were previously unrecognized. These new insights are in many ways an onslaught on the historical "arena of bloody controversies" and I can now say that much of these historical controversies were actually borne out of a complete lack of physical objectives. The results of this thesis are therefore likely to have interesting consequences for QCD where the issue of a θ parameter arose first and the algebra is the same, but experiments are impossible.

Samenvatting

Het oorspronkelijke idee van een instantonparameter θ in de theorie van het quantum-Hall-effect gaat terug tot twintig jaar geleden. Ondanks meer dan twee decennia onderzoek is het onderwerp nog verre van af. Nieuwe ontdekkingen, in het bijzonder het bestaan van massaloze chirale randexcitaties, tonen duidelijk aan dat het quantum-Hall-effect *in het algemeen* een generieke eigenschap is van het θ-vacuum in een asymptotisch vrije veldentheorie. Deze ontdekkingen hebben verder geleid tot een volledige herziening van bepaalde overheersende ideeën in het veld, in het bijzonder die met betrekking tot de betekenis van instantons, de replicamethode, enz. Deze ideeën zijn oorspronkelijk ontstaan op grond van onjuiste historische werken aan het CP^{N-1}-model met grote waarden van N. Deze werken zijn vele jaren lang verkeerd aangezien voor “exacte analyses”.

In dit proefschrift ga ik opnieuw in op de instanton-methodologie in de theorie van het quantum-Hall-effect en breid ik deze methodologie op veel manieren uit. De resultaten in dit proefschrift kunnen in het algemeen gezien worden als een integraal onderdeel van een nieuw concept, dat stelt dat de basisaspecten van het quantum-Halleffect *super-universele* eigenschappen zijn van de instantonparameter θ in schaal-invariante theorieën.

Het fysisch hoofddoel van dit proefschrift is het beschrijven van de quantum-critische eigenschappen van het ongeordende elektronengas in twee ruimtelijke dimensies en in sterke magnetische velden, zowel in aan- als in afwezigheid van elektron-elektroninteracties. Om de algemene betekenis van θ-renormalisatie door instantons vast te stellen beschouw ik eerst het geval van elektronen zonder interacties en vervolgens het geval van het wisselwerkende elektronengas, maar zonder sterk magnetisch veld. Ten slotte beschouw ik het gecombineerde effect van elektron-elektroninteracties en sterke magnetische velden.

In hoofdstuk 2 bestudeer ik het niet-lineaire sigmamodel gedefinieerd op het Grassmann manifold $U(m+n)/U(m) \times U(n)$ en in twee dimensies. Het is bekend dat deze theorie in de limiet $m = n = 0$ de lage-energiedynamica van vrije elektronen in aanwezigheid van ”random impurities” en sterke magnetische velden beschrijft. Voorts laat deze theorie algemene besprekingen toe van zowel de topologische betekenis van de massaloze chirale randexcitaties in het probleem als het naverwante onderwerp van *fysische observabelen*, met name de *electrische geleidingen* σ_{xx} en σ_{xy}. De rest van dit hoofdstuk is gewijd aan de technische details van instantonberekeningen en de specifieke wijze waarop de niet-perturbatieve bijdragen aan de β- en γ-functies van de renormalisatiegroep afgeleid kunnen worden. De gedetailleerde voorspellingen van de instanton-methodologie voor de theorie met $m = n = 0$ worden vervolgens vergeleken met numerieke waarden voor de critische exponenten verkregen uit computersimulaties van het vrije elektronengas, wat tot een opmerkelijke overeenkomst leidt. De resultaten in dit hoofdstuk leveren de theoretische grondslag voor hoofdstuk 4, waarin ik inga op het probleem van elektron-elektroninteracties.

In de hoofdstukken 4 en 5 maak ik gebruik van de resultaten die in de eerdere hoofdstukken zijn geboekt en snijd het hoofdonderwerp van dit proefschrift aan: de θ-afhankelijkheid van de algemene, onderliggende theorie met inbegrip van de effecten van elektron-elektroninteracties. Technisch gesproken vertegenwoordigt hoofdstuk 4 een van de interessantste toepassingen van de instantonmethodologie in het algemeen, en de procedure van *ruimtelijk variërende massa's* in het bijzonder. De reden is dat de (singlet-)interactieterm, anders dan gewone massatermen, de singuliere structuur van de theorie op kleine afstanden compleet verandert, hetgeen slechts een van de vele merkwaardige eigenschappen is van het probleem met elektron-elektroninteracties. Het is waarlijk opmerkelijk dat, ondanks de herculische inspanning die nodig is om de buitengewone hoeveelheid technisch en analytisch detail in bedwang te houden, aan het eind van de berekening een elegante eenvoud zichtbaar wordt. De belangrijkste resultaten van dit hoofdstuk worden samengevat door een renormalisatiegroepdiagram in drie dimensies die bestaan uit de geleidingsparameters σ_{xx} en σ_{xy} en een enkele parameter c, die de "range" van de elektron-elektroninteractie vertegenwoordigt. Om de fysische betekenis van dit stromingsdiagram en, in het bijzonder, het niet-Fermi-vloeistofgedrag van het Coulomb-interactieprobleem te begrijpen zijn de onderliggende symmetrieën zoals $\mathcal{F}$-invariantie van groot belang. Bovendien kan men, op basis van de renormalisatiegroep, onmiddellijk het algemene schaalgedrag van de macroscopische geleidingen uit dit diagram afleiden, als functie van de temperatuur, externe frequentie en magnetisch veld. Deze schaalfuncties zijn de belangrijkste doelstellingen van het experimenteel onderzoek aan de plateau-overgangen in het quantum-Hall-regime.

Samenvattend zijn, door het quantum-Hall-effect als fysische toepassing van de instantonparameter θ te benutten, belangrijke nieuwe inzichten bovengekomen die op andere wijze niet ontdekt zouden kunnen zijn. In het bijzonder noem ik de verschillende aspecten van *super-universaliteit* die voorheen niet gezien werden. Deze nieuwe inzichten zijn op velerlei wijze een aanval op de historische "arena van bloedige controverses" in quantumveldentheorie en ik kan nu stellen dat veel van deze historische controverses geboren zijn uit een volstrekt gebrek aan fysische doelstellingen. Daarom is het waarschijnlijk dat de resultaten van dit proefschrift interessante gevolgen hebben voor QCD, waar het concept van een θ-parameter het eerst opkwam en waar de algebra hetzelfde is, maar experimenten onmogelijk zijn.

Acknowledgement

First of all, I will be always grateful to my promotor Adrianus Pruisken for his guidance on my own way in Theoretical Physics. Aad, you have introduced me into the field of the quantum Hall effect more than five years ago. I appreciate the unique opportunity that you have provided for me. I was your PhD student but spent most of my time in Moscow. Everyone knows that a dissertation is not a product of a single person alone. Without your brilliant ideas and physical intuition the research presented in this book would never have been possible. I am intensively grateful to you for all your help during writing of this thesis. Forever your extended knowledge and deep understanding of Nature will indicate for me the level which one can reach in Physics. Aad, I hope you will remain to provide me your scientific advices in the future.

I am grateful to A. G. Abanov, A. A. Belavin, O. Dimitrova, A. M. Dyugaev, A. M. Finkelstein, Ya. V. Fominov, M. V. Feigelman, L. I. Glazman, P. D. Grigoriev, S. V. Iordansky, A. S. Ioselevich, A. Kamenev, A. B. Kashuba, M. Khodas, D. E. Khmelnitsky, I. V. Kolokolov, S. E. Korshunov, A. Lebedev, V. V. Lebedev, D. S. Lyubshin, Yu. Makhlin, S. S. Murzin, Th. M. Nieuwenhuizen, P. M. Ostrovsky, D. I. Podolsky, I. Polyubin, I. Protopopov, V. M. Pudalov, V. Rychkov, K. A. Saraikin, K. Schoutens, M. Voropaev, and P. B. Wiegmann for valuable discussions of different problems. I am indebted to my co-authors M. A. Baranov, A. P. Bogatov, N. M. Chtchelkatchev, R. Shankar, and M. A. Skvortsov for joint works. I wish to thank all my colleagues in the Institute for Theoretical Physics at University of Amsterdam, the Landau Institute for Theoretical Physics and, last but not least, the Amsterdam Condensed Matter Theory Group for arranging financial support from NWO. I appreciate the help of Y. Chong, L. Gilliéron, and P. Uijthoven during my visits in Institute of Theoretical Physics. Finally, I thank K. Ray for his warm hospitality during my stays in Amsterdam.

List of publications

1. A. P. Bogatov and I. S. Burmistrov, *Attenuation of an optical wave propagating in a waveguide, formed by layers of a semiconductor heterostructure, owing to scattering on inhomogeneities*, Quantum Electronics **29**, 500 (1999) [Kvantovaya Elektronika **27**, 223 (1999)].
2. I. S. Burmistrov, *Two-dimensional electron liquid with disorder in a weak magnetic field*, JETP **95**, 132 (2002) [ZhETF **122**, 150 (2002)].
3. M. A. Baranov, I. S. Burmistrov, and A. M. M. Pruisken, *Non-Fermi-liquid theory for disordered metals near two dimensions*, Phys. Rev. B **66**, 075317 (2002).
4. I. S. Burmistrov and M. A. Skvortsov, *On the effect of far impurities on the density of states of two-dimensional electron gas in a strong magnetic field*, JETP Letters **78**, 156 (2003) [Pis'ma v ZhETF **78**, 188 (2003)].
5. N. M. Chtchelkatchev and I. S. Burmistrov, *Andreev conductance of a domain wall*, Phys. Rev. B **68**, R01338 (2003).
6. I. S. Burmistrov and M. A. Baranov, *Mean-field phase diagram of two-dimensional electrons with disorder in a weak magnetic field*, Phys. Rev. B **68**, 155328 (2003).
7. I. S. Burmistrov, *The anisotropic conductivity of two-dimensional electrons on half-filled high Landau level*, JETP Lett. **79**, 177 (2004) [Pis'ma v ZhETF **79**, 212 (2004)]
8. I. S. Burmistrov, *Anisotropic conductivity tensor on a half-filled high Landau level*, Phys. Rev. B **71**, 035331 (2005).
9. A. M. M. Pruisken and I. S. Burmistrov, *The instanton vacuum of generalized CP^{N-1} models*, Ann. of Phys. (N.Y.) **316**, 285 (2005).
10. A. M. M. Pruisken, M. A. Baranov, and I. S. Burmistrov, *Non-Fermi liquid theory of the quantum Hall effects*, JETP Lett. **82**, 150 (2005) [Pis'ma v ZhETF **82**, 166 (2005)].
11. I. S. Burmistrov and N. M. Chtchelkatchev, *Domain wall effects in ferromagnet-superconductor structures*, Phys. Rev. B **72**, 144520, (2005).
12. A. M. M. Pruisken and I. S. Burmistrov, *Comment on topological oscillations of the magnetoconductance in disordered GaAs layers*, Phys. Rev. Lett. **95**, 189701 (2005).

13. A. M. M. Pruisken and I. S. Burmistrov, *θ renormalization, electron-electron interactions and super universality in the quantum Hall regime*, cond-mat/0502488, submitted to Phys. Rev. B.

14. A. M. M. Pruisken and I. S. Burmistrov, *Non-Fermi liquid criticality and super universality in the quantum Hall regime*, cond-mat/0507412, submitted to Phys. Rev. Lett.

15. I. S. Burmistrov, *On magnetic susceptibility of a spin-S impurity in nearly ferromagnetic Fermi liquid*, cond-mat/0602650, submitted to European Phys. Journal B.

16. A. M. M. Pruisken, I. S. Burmistrov and R. Shankar, *Massless excitations at $\theta = \pi$ in the CP^{N-1} model with large values of N*, cond-mat/0602653, submitted to Nucl. Phys. B.

17. A. M. M. Pruisken, M. A. Baranov, and I. S. Burmistrov, *Localization, Coulomb interaction, topological principles and the quantum Hall effect*, cond-mat/0104387 (unpublished).