# CONFINEMENT AND THE SUPERFLUID DENSITY IN THEORIES OF THE UNDERDOPED CUPRATES, AND STRONGLY COMMENSURATE DIRTY BOSONS IN THE LARGE-N LIMIT

by

Matthew J. Case

B.Sc. (Hons.), Memorial University of Newfoundland, 1999

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE DEPARTMENT

OF

PHYSICS



SIMON FRASER UNIVERSITY

Summer, 2005

# APPROVAL

**Name:** Matthew J. Case

**Degree:** Doctor of Philosophy

**Title of Thesis:** Confinement and the Superfluid Density in Theories of the Underdoped Cuprates, and Strongly Commensurate Dirty Bosons in the Large-N Limit

**Examining Committee:** Dr. Howard Trottier (Chair)

---

Dr. Igor Herbut, Senior Supervisor
Associate Professor, Department of Physics, SFU

---

Dr. Michael Plischke, Supervisor
Professor, Department of Physics, SFU

---

Dr. Michael Wortis, Supervisor
Professor Emeritus, Department of Physics, SFU

---

Dr. David Broun, Internal Examiner
Assistant Professor, Department of Physics, SFU

---

Dr. Marcel Franz, External Examiner
Associate Professor, Department of Physics and Astronomy, UBC

**Date Approved:** August 25, 2005

# SIMON FRASER UNIVERSITY

## Partial Copyright Licence

# Abstract

In this thesis, we study various issues arising from the QED theory of underdoped, high temperature superconductors in 2+1 dimensions. The theory breaks up roughly into two sectors: fermionic and bosonic. With regard to the fermionic sector, we consider confinement of the emergent gauge field which we take to be compact. In the absence of fermions, the interaction between monopoles is Coulombic and the well known result is that the pure gauge theory is permanently confining. With the addition of fermions, the interaction becomes logarithmic, and an analogy with the usual Kosterlitz-Thouless transition suggests a deconfinement transition for the fermions. We show, however, that, when screening is taken into account, the deconfined phase is destabilized and fermions remain permanently confined.

The bosonic sector models Cooper pair phase fluctuations, whose effect on the depletion of the superfluid density we examine in two separate studies. In the first of these, we study the quantum XY model, and show that the quasi-two dimensionality, low critical temperatures and large d-wave gap characteristic of underdoped cuprates severely constrain the form of the superfluid density. Under these assumptions, we find that phase fluctuations alone are insufficient to account for recent observations of deviations from Uemura scaling, and that the quasiparticle contribution is a necessity. We use our results to satisfactorily fit the recent data.

In the second study, we model the cuprates by a layered system of interacting bosons and examine the collective excitations in this system. Depending on the anisotropy and the interaction strength, we find find four different regimes of temperature dependence of the superfluid density. We argue that interactions in the underdoped cuprates are effectively short-ranged and weak.

Finally, we study the related issue of disordered, interacting bosons in the large-N limit and at strong commensuration. Perturbatively at weak disorder and numerically at strong, we show that the screening of the random potential due to interactions is insufficient to delocalize the single-particle states so that no superfluid transition occurs from the Mott insulator.

*For my parents*

# Acknowledgments

First and foremost, I would like to acknowledge my supervisor, mentor and friend, Igor Herbut. Our discussions always provided me with much greater insight, whether the topic was physics or basketball, and I invariably left these meetings feeling refreshed from his constant enthusiasm and support. I am also grateful to the past and present members of our larger group, namely Mohammad Amin, Kamran Kaveh, Dominic Lee and Babak Seradjeh, particularly Babak and Kamran, with whom I have shared many informative discussions over the years.

I would also like to thank my committee members Michael Plischke and Michael Wortis for their probing questions during our yearly get-togethers, as well as David Broun for his enlightening comments on the superfluid density in the cuprates. I have also benefited from my classes with Dr. Irwin, Dr. Trottier and Dr. Viswanathan. A sincere thanks is due to the Physics department staff: Dagni, Susan and, especially, Candida and Sada, for all their incredibly hard work.

A warm thanks also to my friends and colleagues at SFU: Geoff Archibald, Marco Bieri, Fergal Callaghan, Erin Chapple, Jordi Cohen, Mike Dugdale, Eldon Emberly, Nichol Furey, Josie Herman, Suckjoon Jun, Karn Kallio, Mahshid Karimi, Shabnam Kavousian, Peter Matlock, Matt Nobes, Philip Patty, Oliver Pitts, Mehrdad Rastan, Bram Sadlik, Vahid Shahrezaei, Paul Sikora, Michael Short, Borzou Toloui, Dan Vernon, Phillipe Westreich, Kit Wong, Yi Yang, and Yvonne Yuan. Extra special thanks are due to Arpal and Kamal for keeping me fed and healthy.

Finally, but most importantly, I would like to recognize my parents, grandparents, brother and sister, for their years of support and encouragement, and Cecilia and the Lei family, for accepting me into their hearts. I'm here because of all of you.

Thank you all!

# Contents

# List of Figures

# Chapter 1

# Introduction

*"We can blame Bednorz and Müller."*

This is the opening proclamation of one writer's[1] attempted synthesis of the fast-paced field of high temperature superconductivity. Indeed, the 1986 discovery of superconductivity in the ceramic compound $Ba_xLa_{5-x}Cu_5O_{5(3-x)}$ by the above mentioned gentlemen[2] has sparked a renaissance in the field which most people thought was a closed book after the extremely successful theory of Bardeen, Cooper and Schrieffer (BCS) in 1956[3]. The blame and, therefore, the credit lies squarely with them, and for the achievement they were duly awarded with the Nobel prize in 1987,[1] one of the quickest turnarounds in the Swedish Institute's history.

With now over 100 materials belonging to this class of cuprate superconductors, and with critical temperatures as high as 138 K, the importance of their discovery is undeniable. The incredible growth of the field is due in large part to their high transition temperatures, well above the temperature of liquid nitrogen and, thus, experimentally accessible for most researchers. The technological implications are also very attractive, from superconducting magnets for use in medical imaging devices and magnetic levitation trains to supercooled wires in electrical power grids. For theoretical physicists, these materials posed an immediate challenge, as it was quickly realized that the superconductivity in the cuprates differed drastically from the conventional BCS kind, and the struggle to find a coherent explanation began.

[1] ...for their important breakthrough in the discovery of superconductivity in ceramic materials.

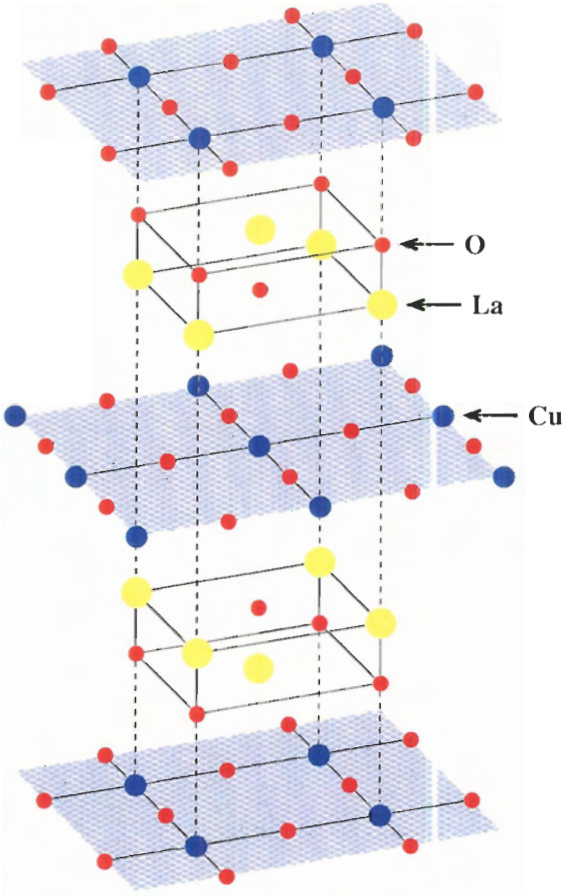


Figure 1.1: Schematic structural diagram of $La_2CuO_4$

## 1.1 Introducing the cuprates

The cuprates are characterized by layered planes of copper and oxygen atoms, with as few as one plane per unit cell in the lanthanum compounds and as many as seven in mercury based materials[4]. A single unit cell of undoped $La_2CuO_4$ is shown in Figure 1.1, highlighting the $CuO_2$ planes sandwiched between 'charge reservoirs' of La and O atoms. The crystallographic directions are in (a and b) or out of (c) the plane. Myriad evidence now points to these $CuO_2$ layers as crucial players in the physics of high temperature superconductivity, and there even seems to be a direct relationship between higher critical temperatures and an increased number of layers, but only up to three layers, after which the critical temperature decreases[4].

With such an anisotropic crystal structure, one might speculate that these materials also display other anisotropic properties, and they do. Measurements such as electrical transport[5], penetration depth[6] and magnetic torque[7] all indicate that cuprate superconductivity is quite two-dimensional, occurring primarily in the $CuO_2$ layers. For example, the ratio of $c$-axis to in-plane penetration depths can be as high as 100 in $YBa_2Cu_3O_{6-x}$[8, 9].

It must be noted, of course, that the pure, undoped materials, like the one shown in Figure 1.1, do not display superconductivity. In fact, all cuprates are believed to be antiferromagnetically ordered *insulators* at half-filling, with Néel temperatures of around 300 K. This is in contrast to conventional superconductors which are typically very good metals above $T_c$, conforming to Landau's fermi liquid paradigm[10], and for which magnetism is quite detrimental. Superconductivity is achieved in the cuprates through a process known as doping, where oxygen is added or out of plane atoms are partially substituted for others, such as strontium in place of lanthanum in the case of $La_{2-x}Sr_xCuO_4$. In this way, mobile holes are introduced into the $CuO_2$ planes, and these quickly destroy the antiferromagnetism.

With increased doping, resistivity measurements demonstrate that these materials become superconducting at some critical value $x_{\rm ud}$ and below the critical temperature $T_c$. This is known as the *underdoped* regime and is characterized by increasing critical temperatures and quasi two dimensionality. Eventually, a maximum $T_c$ is reached at what is called optimal doping, after which the transition temperature declines in the overdoped regime until the superconducting state is destroyed at $x_{\rm od}$. At the same time, the anisotropy between $ab$-plane and $c$-axis properties also decreases and the materials become more three dimensional[11]. Beyond $x_{\rm od}$, the materials are believed to behave like conventional metals. The situation is summarized in the generic phase diagram shown in Figure 1.2.

## 1.2 Understanding the phase diagram

In contrast to conventional superconductors, which are well described by the BCS theory at weak coupling, the cuprates are strongly correlated electronic systems, and their low-doping behaviour can be well understood in this context. The strong Coulomb repulsion near half-filling ($x = 0$) forces the conduction electrons to singly occupy the available sites, thus completely blocking the motion of charge and resulting in a Mott insulator. Particle exchange can only occur through virtual processes which require oppositely aligned spins, due to Pauli's exclusion principle, and antiferromagnetism

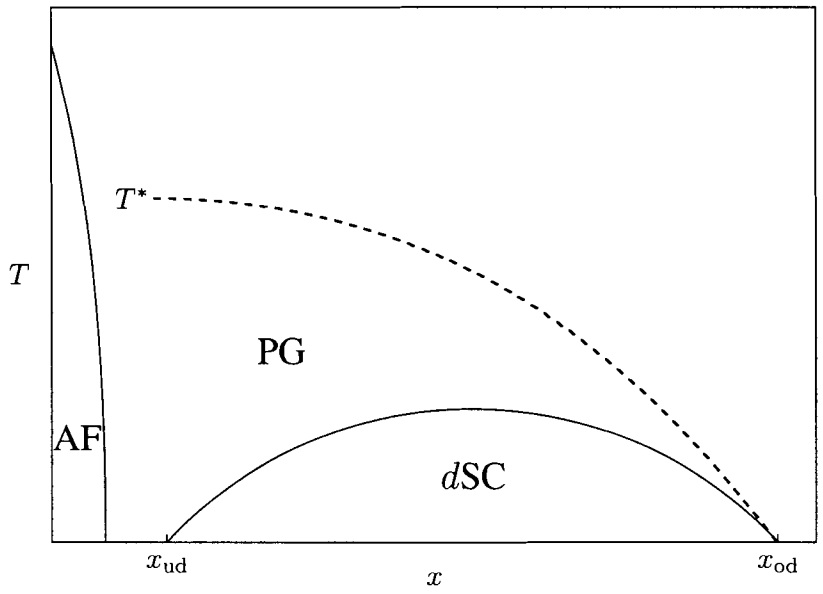


Figure 1.2: Generic temperature-doping phase diagram of cuprate superconductors, showing the antiferromagnetic (AF), superconducting ($d$SC) and pseudogap (PG) phases.

is therefore the natural outcome.

Understanding the phase diagram as holes are added is much more complicated. Within the framework of the BCS theory, the emergence of superconductivity from the antiferromagnetically ordered Mott insulating state seems paradoxical. In conventional materials, the superconducting state arises due to an instability of the weakly interacting Fermi liquid to phonon mediated pairing between electrons[3]; it is difficult to imagine how this pairing can occur in the face of the strong interactions present in the cuprates. To further complicate the situation, experiments early on found no sign of an isotope effect in high temperature superconductors[12], thus effectively ruling out the conventional pairing mechanism.

Another candidate for electron pairing has a magnetic origin and is, thus, consistent with the cuprate phase diagram. The essence of this mechanism can be seen in the $t$-$J$ model, which is believed to capture the important physics near half-filling. The $t$-$J$ model is derived from the Hubbard model for large interaction $U$:

$$H_{t\text{-}J} = -t \sum_{\langle i,j\rangle,\sigma} c^{\dagger}_{i,\sigma} c_{j,\sigma} + J \sum_{\langle i,j\rangle} \left( \mathbf{S}_i \cdot \mathbf{S}_j - \frac{1}{4} n_i n_j \right), \tag{1.1}$$

where $J = t^2/U$. The $J$ term describes virtual spin exchange processes, and results in antiferromagnetism at half-filling. Fourier transforming, this term becomes[13]

$$H_J = -\frac{J}{N}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \left(\varphi_{s^*}(\mathbf{k})\varphi_{s^*}(\mathbf{k}') + \varphi_d(\mathbf{k})\varphi_d(\mathbf{k}')\right) c^\dagger_{\mathbf{k}+\mathbf{q}/2,\uparrow} c^\dagger_{-\mathbf{k}+\mathbf{q}/2,\downarrow} c_{-\mathbf{k}'+\mathbf{q}/2,\downarrow} c_{\mathbf{k}'+\mathbf{q}/2,\uparrow}, \quad (1.2)$$

which displays *attraction* in the '$s^*$' and '$d$' channels, described by the order parameters $\varphi_{s^*}(\mathbf{k}) = (\cos k_x + \cos k_y)$ and $\varphi_d(\mathbf{k}) = (\cos k_x - \cos k_y)$, respectively. The mean-field theory of the $t$-$J$ model demonstrates that the the ground state is indeed superconducting[13]. Beyond mean-field, other states are possible[14, 15], and there is not currently a consensus on whether superconductivity survives in the phase diagram.

However, the above argument does show that superconductivity can, in principle, arise from mechanisms other than by phonons, and that this superconductivity is of quite an unusual type. Conventional superconductivity leads to an order parameter with $s$-wave symmetry, which is completely isotropic in $k$-space. In the above derivation, the order parameter can have very particular $k$-dependence, characterized by the form of the function $\varphi_{s^*}$ or $\varphi_d$; these two types of order are extended $s$-wave ($s^*$) and $d$-wave ($d$). When the superconducting gap opens, these two orders presumably compete for precedence, but it is the $d$-wave state which is of lower energy. This is because, near half-filling, the $s^*$-wave gap is almost zero over the entire Fermi surface, while the $d$-wave gap vanishes only along the directions $k_x = \pm k_y$. Thus, the energy gained by condensing in the $s^*$-wave state is much lower than for the $d$-wave state[13].

## 1.3 Gap symmetry

The question of the order parameter symmetry has been extensively studied by experimentalists and it is now well established that cuprates are $d$-wave superconductors. The earliest evidence for this was provided by measurements of the in-plane penetration depth by microwave cavity perturbation techniques[6], where a linear low-temperature behaviour was observed, characteristic of an order parameter with nodes. The reason for this temperature dependence is that the nodes in the gap at the Fermi surface allow quasiparticles to become excited at arbitrarily low energies, and these quasiparticles can, therefore, efficiently destroy superconductivity. This should be compared to the $s$-wave case which has a full isotropic gap everywhere on the Fermi surface and, thus, quasiparticles are excited only when the threshold energy set by the gap is overcome. This leads to an exponentially activated form for the temperature dependence of the penetration depth[16]. More conclusive evidence

for $d$-wave pairing comes from angle-resolved photoemission spectroscopy (ARPES)[17], which is able to probe the gap along different directions in $k$-space. There, the gap is seen to vanish along the $(\pm\pi, \pm\pi)$ directions, and maxima are observed along $(0, \pm\pi)$ and $(\pm\pi, 0)$. The most direct evidence, however, comes from phase sensitive measurements of Josephson corner junctions[18], which indicate a sign change of the order parameter between the $a$ and $b$ directions[19].

## 1.4 The pseudogap

Some of these experimental techniques have also been used to shed light on the non-superconducting state above $T_c$, and have yielded very surprising results. As these materials are heated above their critical temperatures, a spin gap persists, slowly shrinking and finally vanishing at some much higher temperature $T^*$, which can be almost as high as room temperature. This effect is most prominent in underdoped samples and disappears in overdoped, though exactly where is unclear. This gapped, non-superconducting state has been dubbed the pseudogap and is indicated in the phase diagram 1.2 by a dashed line, since it is not believed to demark an actual phase transition.

Many different explanations of the pseudogap exist, including competing orders or the resonating valence bond (RVB) state; the scenario to be promoted here involves preformed pairs above $T_c$. This model interprets the line $T^*$ as the mean-field temperature at which Cooper pairs are formed, but before phase coherence is established. The pseudogap is then a phase fluctuating $d$-wave superconductor[20, 21], with a fully formed gap. Evidence for this comes from heat transport[22], which measures the gap magnitude at various dopings. The values extracted from these measurements seem to correlate extremely well with the observed pseudogap, indicating that the pseudogap and superconducting gap are one and the same thing.

## 1.5 Fluctuating superconductivity

To assess the importance of phase fluctuations, it is useful to know the phase stiffness, as measured by the superfluid density at zero temperature. This quantity is accessible experimentally via penetration depth[6] and muon spin relaxation ($\mu$SR)[23–25] measurements, for example. For conventional materials, the superfluid density is proportional to the total number of electrons, and it should, therefore, be expected to decrease as holes are doped into the planes of cuprate superconductors. However, the aforementioned experiments clearly demonstrate that the superfluid density *increases*

with doping and is, in fact, proportional to $x$. In particular, early $\mu$SR measurements[23, 24] saw that the zero-temperature superfluid density is proportional to the critical temperature, with a constant of proportionality which is universal for a wide array of cuprates. These observations imply a greatly increased role for phase fluctuations as compared to conventional superconductors, especially in the underdoped regime, and physical quantities should reflect these effects. This argument for phase fluctuations was first clearly put forth in an influential article by Emery and Kivelson[26].

In underdoped samples, fluctuation effects should be apparent above $T_c$ and below $T^*$. Indeed, recent experiments on the Nernst effect in underdoped lanthanum compounds[27, 28] indicate that vortex fluctuations are present near and above $T_c$. These experiments measure the voltage transverse to a thermal gradient in the presence of an applied magnetic field. Vortices diffuse along the direction of the thermal gradient, generating a transverse electric field via the Josephson effect. The results show a positive Nernst signal well above the critical temperature, but dying significantly below the apparent pseudogap temperature. This is good evidence for fluctuating superconductivity existing in the pseudogap state but does not obviously corroborate the preformed pairs viewpoint, since $T^*$ and the onset temperature for the Nernst signal are not identical[29].

Of course, phase fluctuations should also affect the physical properties in the superconducting state, such as the superfluid density. Unfortunately, the growing of clean, underdoped samples has been fraught with complications, and only very recently have reliable measurements been done[8]. The latest measurements of the lower critical field in ultra-pure $YBa_2Cu_3O_{6-x}$ indicate that the superfluid density decreases linearly from its zero temperature value, just as is seen from penetration depth measurements at optimal doping[6]. Recalling that the results at optimal doping were taken as evidence of $d$-wave pairing and, thus, due to quasiparticle excitations, the new low-doping data seems at odds with the phase fluctuation picture.

As the doping is increased and the temperature at which phase coherence sets in also increases, the pseudogap temperature steadily decreases. At some point, these two energy scales cross and the relevant temperature becomes the one at which pairs form, concomitantly with phase ordering. The effects of fluctuations would then only be expected in a narrow region near the critical temperature. In conventional superconductors, which have long coherence lengths, this critical fluctuation region is vanishingly small[30]. On the other hand, because the cuprates are strongly type-II superconductors with short coherence lengths, this region is much larger[30, 31] and has, in fact, been seen with a width of about 10 K in in-plane penetration depth measurements on optimally doped $YBa_2Cu_3O_{6.95}$[32]. In the new lower critical field data for underdoped samples[8], however, it

seems that this critical region has completely disappeared, again seemingly in contradiction with the phase fluctuation picture.

## 1.6 A unified view of the superconducting state

To tie together the distinct contributions from phase fluctuations and quasiparticles and to highlight their respective regions of influence, there exists a simple and elegant relation, first derived in the context of the RVB state[33] and known as the Ioffe-Larkin rule. In the case of, for example, the in-plane superfluid density, this relation is

$$\rho_{ab}^{-1}(T) = \rho_{\mathrm{pf}}^{-1}(T) + \rho_{\mathrm{qp}}^{-1}(T), \tag{1.3}$$

where $\rho_{\mathrm{pf}}$ and $\rho_{\mathrm{qp}}$ are the phase fluctuation and quasiparticle superfluid densities, respectively; this is essentially a mathematical expression of the Emery-Kivelson argument. From this, it is clear that the total superfluid density is determined primarily by the smaller of the two contributions, and that superconductivity is lost as soon as one of these becomes zero. In the case of optimally doped $YBa_2Cu_3O_{6.95}$, this relation describes what is observed experimentally extremely well. At low temperatures, the quasiparticle part is responsible for the linear temperature dependence, while near the critical temperature, it is the phase fluctuations that ultimately destroy superconductivity, with a wide critical region as their signature. To apply the Ioffe-Larkin rule to underdoped samples, it is crucial to have a clear understanding of the superfluid density resulting due to phase fluctuations. Only then can it be accurately determined to what degree this picture holds true for the cuprates and, in particular, whether the latest data are consistent with this viewpoint.

## 1.7 Theoretical approaches

There are, of course, many different ways of approaching the question of superconductivity in the cuprates. Traditionally, superconductivity has been viewed as a deriving from some 'normal' state, and this language is widely used today when discussing high temperature superconductivity. In the BCS theory, the normal state was, indeed, normal. These conventional materials were all good metals at room temperature, well described by Landau's Fermi liquid theory, and the superconducting state was finally understood as an instability towards pairing caused by lattice distortions. The success of this 'bottom-up' approach has motivated many researchers to adopt a similar strategy for the

cuprates.

Most current attempts begin with the $t$-$J$ model introduced earlier. It is the simplest model which is believed to capture all of the key features of a doped Mott insulator[14]. At half-filling, it describes the Mott insulating antiferromagnetic state, and doping is incorporated through the effective hopping term. Away from half-filling, the strong interactions must be taken into account by prohibiting the occupation of sites by two electrons. This is a complicated task, and most attempts seem invariably to lead to strongly interacting gauge theories[14]. Nonetheless, significant progress has been made, and a whole host of possible states have been identified which emerge upon doping. Among these are, of course, $d$-wave superconductivity[13], as well as the RVB state of Anderson[34] and various flux phases[35]. Clearly, however, the nature of the non-superconducting 'normal' state remains an unresolved issue.

On the other hand, in contrast to the strongly correlated Mott insulating state, the superconductor which derives from it is quite well understood. It has $d$-wave symmetry and well-defined nodal quasiparticle excitations[22]. This suggests a break from tradition, and thus a 'top-down' approach, starting with the superconducting state, and determining what other states derive from it. This approach has been taken by several authors[36, 37], most recently within the $QED_3$ formulation[20, 21, 38, 39]. These approaches also lead to gauge theories as the effective low energy descriptions, and have provided results in qualitative and quantitative agreement with experiment.

## 1.8 Scope of this thesis

In the present thesis, we will adopt the 'top-down' framework presented in Herbut [20]. Our philosophy is in keeping with the preformed-pairs interpretation of the pseudogap state, and, in Chapter 2, we will review the theoretical implications of phase disordering a $d$-wave superconductor with a well formed gap. We will see that the low energy theory can be described by an effective action with two emergent gauge fields which derive from the fluctuating vortex loops inherent to the two-dimensional system at zero temperature. The action for the phase fluctuations will take the form of a theory for two types of bosons coupled to the fermionic action via the gauge fields, while the fermionic theory will take the form of three-dimensional quantum electrodynamics ($QED_3$) in the non-superconducting state. Due to a special symmetry, hidden in the original theory, the destruction of superconductivity is accompanied by the dynamical generation of spin density wave (SDW) order and, so, the zero temperature non-superconducting ground state is connected to the parent

antiferromagnet at half-filling.

After this introduction to the theoretical framework, we will proceed in Chapter 3 to a discussion of an issue which arises in all gauge theories – confinement. In the context of the cuprate superconductors, confinement is related to the idea of spin-charge separation, and we will see that, just as in the pure gauge theory studied previously[40], the fermions in compact $QED_3$ with matter fields are permanently confined. This work was first published in Case et al. [41].

The next two Chapters (4 and 5) will be dedicated to a discussion of the superfluid density in underdoped superconductors. Chapter 4 covers research previously published in Herbut and Case [42], where we probed the importance of phase fluctuations in determining the temperature dependence of the superfluid density in anisotropic systems. In trying to explain classic $\mu$SR measurements, we will see that phase fluctuations in this model are insufficient to describe the experimental situation, and that quasiparticles must be included via the Ioffe-Larkin rule to improve the agreement. We then predict the low-doping form of the superfluid density in this scenario and, also, argue that we can explain new observations on the relation between the zero temperature superfluid density and the critical temperature.

The contents of Chapter 5 can also be found in Case and Herbut [43]. There, we recognize the caveats of the previous Chapter's arguments, and model the strong anisotropicity inherent to the cuprates by a layered system of bosons with Coulomb interactions. We will see that the effect of the layers is to screen the Coulomb interaction into a short-ranged one. The results from dilute bosonic systems then apply, and we find four regimes of temperature dependence of the superfluid density. Examining experimental data, we find that the underdoped cuprates are in the two-dimensional, short-ranged regime with weak to moderate interaction strength. We then argue that this explains the most recent lower critical field data on very underdoped $YBa_2Cu_3O_{6+x}$.

The final content Chapter (Chapter 6) recalls early work done on the combined effect of disorder and interactions on two-dimensional bosons and was previously published in Case and Herbut [44]. It is shown there, in the limit of a large number of order parameter components, that, while interactions partially screen the disorder, the ground state remains localized and no superfluid state obtains. While not directly related to the other studies, we will argue that the study of dirty bosons is of significance to the underdoped cuprates and that, in fact, the discovery of the elusive Bose-glass phase could be a potential candidate for the pseudogap state.

While each Chapter will contain its own conclusions, in Chapter 7 we will provide an outlook to further studies related to the present thesis.

# Chapter 2

# Theoretical framework

In this Chapter, we introduce the theoretical framework which sets the context for the work done in subsequent chapters. We will see that the theory breaks up roughly into two parts: a fermionic (spinonic) and a bosonic. These two sectors are coupled by two U(1) gauge fields which mediate the strong-correlation effects. The derivation presented here is based primarily on the articles of Herbut[20, 21], although other derivations exist[38, 39].

## 2.1 Preamble

The cuprate superconductors are known to be of $d$-wave type, conforming to the usual BCS phenomenology. Importantly, it has been experimentally determined that quasiparticles in this state are well defined, with very long lifetimes[45, 46]. These observations suggest that we begin with a theory in terms of $d$-wave quasiparticles and determine what states arise when the superconductivity is destroyed.

In principle, this destruction can happen in several ways. In conventional superconductors, the primary means is by driving the gap amplitude to zero through thermal fluctuations. This situation is well described by the Landau-Ginzburg mean field theory, which holds except in a very narrow temperature range around $T_c$[30]. In the cuprates, on the other hand, the saddle-point value of the gap qualitatively follows the observed pseudogap[47], which is divorced from the actual critical temperature on the underdoped side of the phase diagram. As already discussed, our interpretation is to identify the pseudogap with the formation of the superconducting gap, so, at temperatures well below $T^*$, amplitude fluctuations can be safely ignored. This argument has been made more precise

in the context of the mean-field theory of the $t$-$J$ model[13].

In the $t$-$J$ model, we saw that magnetic exchange interactions allow ordering in the extended $s$-wave channel as well as in the $d$-wave channel. Fluctuations towards this type of ordering, or any other, could then also lead to the loss of $d$-wave superconductivity in the cuprates. At least in the case of $s^*$-wave order, the $d$-wave saddle point has been argued to be stable to these fluctuations, as well[13].

The final mechanism for the destruction of superconductivity is by fluctuations in the phase of the order parameter. These can proceed classically (thermally) or quantum mechanically (due to interactions). We have already argued that the observed low superfluid density indicates an increased role for phase fluctuations, and we will assume that this is the relevant mechanism for the underdoped cuprates. Further justification can be found in Paramekanti et al. [13].

## 2.2 BCS superconductivity and the effective low-energy action

We begin by considering the usual finite-temperature BCS action:

$$S = T \sum_{\omega_n, \mathbf{k}, \sigma} \left[ (i\omega_n - \xi(\mathbf{k})) c^\dagger_\sigma(k) c_\sigma(k) - \frac{\sigma}{2} \Delta(\mathbf{k}) c^\dagger_\sigma(k) c^\dagger_{-\sigma}(-k) + \text{h.c.} \right], \tag{2.1}$$

where $c$ and $c^\dagger$ are (electron) quasiparticle annihilation and creation operators, respectively. The projection of spin along the quantization axis is denoted by $\sigma = \pm 1$ and we are using the shorthand $k = (\omega_n, \mathbf{k})$. Quasiparticle interactions are being neglected in the present discussion, and, anyway, they are irrelevant if weak; their effects have been considered elsewhere. Since we are interested in describing the cuprate superconductors, we will assume the $d_{x^2-y^2}$ form for the gap function:

$$\Delta(\mathbf{k}) = \frac{\Delta_0}{2} (\cos(k_x) - \cos(k_y)). \tag{2.2}$$

Evidently, this form displays nodes along the diagonal directions $k_x = \pm k_y$ and it is around these points that we will focus.

Referring to Figure 2.1, the nodes $(i)$-$(iv)$ are located at wave vectors $\pm \mathbf{K}_{\text{I,II}}$. At these points, the fermi-surface is ungapped so we expect the low-energy theory to be dominated by contributions coming from wave vectors in the vicinity of the nodes. Proceeding with this in mind, we can construct a four-component representation of the original action (2.1) by defining the fields

$$\Psi^\dagger_i(\mathbf{q}', \omega_n) \equiv \left( c^\dagger_\uparrow(\mathbf{K}_i + \mathbf{q}', \omega_n), c_\downarrow(-\mathbf{K}_i - \mathbf{q}', -\omega_n), c^\dagger_\uparrow(-\mathbf{K}_i + \mathbf{q}', \omega_n), c_\downarrow(\mathbf{K}_i - \mathbf{q}', -\omega_n) \right), \tag{2.3}$$

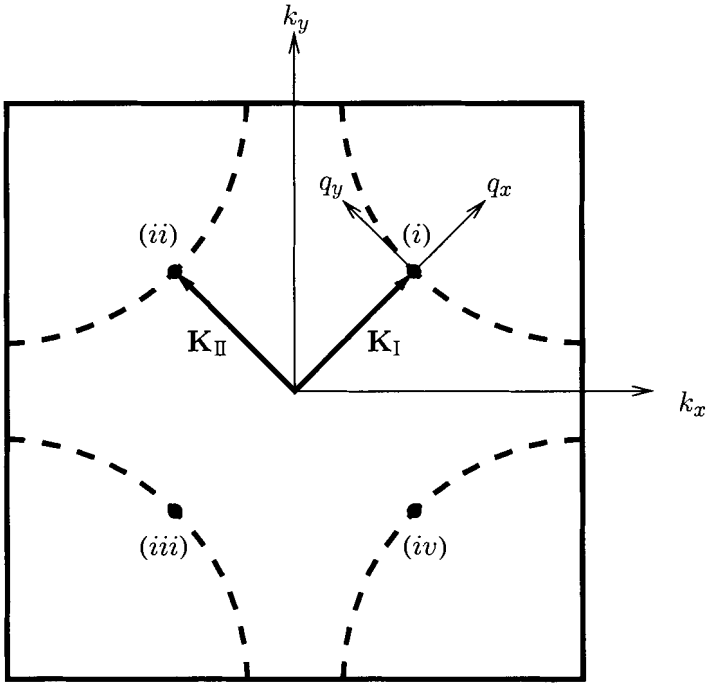


Figure 2.1: Schematic diagram of the fermi surface (dashed line) indicating the four nodes $(i - iv)$ in the $d$-wave order parameter, located at $\pm\mathbf{K}_{\mathrm{I,II}}$.

where $|\mathbf{q}'| \ll |\mathbf{K}_i|$, reflecting the fact that we are interested in the low-energy effective theory corresponding to (2.1). The construction implemented here clearly is not unique. In our definition (2.3), each field is essentially composed of two Nambu-Gor'kov spinors, one for each of the nodes in a diagonally opposed pair; in another popular prescription, it is spin-reversed states which are combined in a similar manner. The choice one makes in this regard is dictated by the particular symmetry of the problem which is being brought to light. For the second choice above, we would recover the full SU(2) spin-rotation symmetry of the problem, while our choice (2.3) yields a (hidden) *chiral* SU(2) symmetry. This second point will be elaborated as we proceed.

The functions $\xi(\mathbf{k})$ and $\Delta(\mathbf{k})$ can now be expanded in the vicinity of the nodes. For nearest-neighbour hopping, the kinetic term can be written:

$$\xi(\mathbf{k}) = -2t(\cos(k_x) + \cos(k_y)) - \mu, \tag{2.4}$$

where $t$ is the hopping parameter and $\mu$ is the chemical potential. Noting that $\xi(\mathbf{K}_i) = 0$, we find,

for example,

$$\xi(\mathbf{K}_{\rm I}+\mathbf{q}') \approx 2t\sin(k_f)(q'_x+q'_y) \tag{2.5}$$

since $|K_{i,x}| = |K_{i,y}| \equiv k_f$. Further, rotating the $q'$-frame by $\pi/4$ leads to

$$\xi(\mathbf{K}_{\rm I}+\mathbf{q}') \approx 2\sqrt{2}t\sin(k_f)q_x \equiv v_f q_x. \tag{2.6}$$

Similarly, from the definition (2.2), we find

$$\Delta(\mathbf{K}_{\rm I}+\mathbf{q}') \approx \sqrt{2}\Delta_0\sin(k_f)q_y \equiv v_\Delta q_y. \tag{2.7}$$

Also, for $\mathbf{k} \approx \mathbf{K}_i$ it can seen that

$$\xi(\mathbf{k}) = -\xi(\mathbf{k}-2\mathbf{K}_i), \tag{2.8}$$

$$\Delta(\mathbf{k}) = -\Delta(\mathbf{k}-2\mathbf{K}_i). \tag{2.9}$$

Finally, we can rewrite the original action (2.1) in terms of the nodal fields $\Psi_i$ to arrive at the low-energy effective theory:

$$S_f[\Psi_i] = T\sum_{\omega_n}\int\frac{\mathrm{d}^2\mathbf{q}}{(2\pi)^2}\left\{\Psi^\dagger_{\rm I}(q)\left[i\omega_n + iM_1 v_f q_x + iM_2 v_\Delta q_y\right]\Psi_{\rm I}(q) + ({\rm I}\to{\rm II}, x\leftrightarrow y)\right\} \tag{2.10}$$

or in real space:

$$S_f[\Psi_i] = \int_0^\beta \mathrm{d}\tau\int\mathrm{d}^2\mathbf{r}\left\{\Psi^\dagger_{\rm I}(r)\left[\partial_\tau + M_1 v_f\partial_x + M_2 v_\Delta\partial_y\right]\Psi_{\rm I}(r) + ({\rm I}\to{\rm II}, x\leftrightarrow y)\right\}, \tag{2.11}$$

where $\beta = 1/T$ and the Fourier transformed fields are

$$\Psi_i(r) = T\sum_{\omega_n}\int\frac{\mathrm{d}^2\mathbf{q}}{(2\pi)^2}e^{i\omega_n\tau + i\mathbf{q}\cdot\mathbf{r}}\Psi_i(q). \tag{2.12}$$

An upper (UV) cutoff $\Lambda$ is implicit in all momentum integrals, the scale of $\Lambda$ being set by the pseudogap energy, $T^*$. The $4\times 4$ matrices in the above equations (2.10,2.11) are given by $M_1 \equiv -i\sigma_3\otimes\sigma_3$ and $M_2 \equiv i\sigma_3\otimes\sigma_1$, where the $\sigma_i$ are the usual Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.13}$$

## 2.3 Order parameter fluctuations

The form of the effective theory (2.11) has been derived neglecting fluctuations in the order parameter field, $\Delta(\mathbf{r})$, and it is the inclusion of these to which we turn now. Returning to the original action (2.1), we can write the pairing term in real space as

$$S_\Delta = -\frac{1}{2}\int_0^\beta \mathrm{d}\tau \sum_{\mathbf{r},\mu} \left\{ \Delta_\mu(r) \left[ c_\uparrow^\dagger(\mathbf{r},\tau) c_\downarrow^\dagger(\mathbf{r}+\hat{e}_\mu,\tau) - \uparrow\leftrightarrow\downarrow \right] + \text{h.c.} \right\}, \tag{2.14}$$

where the sum is over nearest-neighbour sites. In real space, the pairing field is given by $\Delta_x(r) = -\Delta_y(r) \equiv \Delta_d$. Fluctuations are incorporated by expressing this as

$$\Delta_\mu(r) = \Delta_d(1+\eta_\mu(r))e^{i\Phi_\mu(r)}, \tag{2.15}$$

where $\eta_\mu(r)$ and $\Phi_\mu(r)$ represent amplitude and phase fluctuations, respectively. $\Phi_\mu(r)$ can be further decomposed into $\Phi_x(r) = \phi(r)$ and $\Phi_y(r) = \pi + \theta(r) + \phi(r)$ so that $\theta(r)$ accounts for deviations away from the $d$-wave saddle-point, while $\phi(r)$ are uniform phase fluctuations. Since we expect that the fields $\eta$ and $\theta$ are massive (i.e., gapped), we neglect their fluctuations and concentrate only on the field $\phi$.

At this point, it is useful to rederive the effective low-energy theory corresponding to the pairing term (2.14), but in a different way from that of the preceding section, so as to clarify how fluctuations are included in the final form of the action. Fourier transforming the fields (2.3) to real space using the definition (2.12), we find

$$\Psi_{i,\alpha}(r) = \begin{pmatrix} e^{-i\mathbf{K}_i\cdot\mathbf{r}} c_\uparrow(r) \\ e^{-i\mathbf{K}_i\cdot\mathbf{r}} c_\downarrow^\dagger(r) \\ e^{i\mathbf{K}_i\cdot\mathbf{r}} c_\uparrow(r) \\ e^{i\mathbf{K}_i\cdot\mathbf{r}} c_\downarrow^\dagger(r) \end{pmatrix}. \tag{2.16}$$

From this, we can write

$$c_\uparrow^\dagger = e^{-i\mathbf{K}_\mathrm{I}\cdot\mathbf{r}}\Psi_{\mathrm{I},1}^\dagger = e^{i\mathbf{K}_\mathrm{I}\cdot\mathbf{r}}\Psi_{\mathrm{I},3}^\dagger = e^{-i\mathbf{K}_\mathrm{II}\cdot\mathbf{r}}\Psi_{\mathrm{II},1}^\dagger = e^{i\mathbf{K}_\mathrm{II}\cdot\mathbf{r}}\Psi_{\mathrm{II},3}^\dagger \tag{2.17}$$

$$c_\downarrow^\dagger = e^{i\mathbf{K}_\mathrm{I}\cdot\mathbf{r}}\Psi_{\mathrm{I},2} = e^{-i\mathbf{K}_\mathrm{I}\cdot\mathbf{r}}\Psi_{\mathrm{I},4} = e^{i\mathbf{K}_\mathrm{II}\cdot\mathbf{r}}\Psi_{\mathrm{II},2} = e^{-i\mathbf{K}_\mathrm{II}\cdot\mathbf{r}}\Psi_{\mathrm{II},4}, \tag{2.18}$$

so that, when we compute products of electron operators, we can choose to write the $\Psi$-pairings in the most convenient way. To be precise, we will consider only terms which connect elements at the

same node. As an example, consider the following contribution to the action from node $(i)$:

$$S_\Delta^{(i)} \rightarrow -\frac{1}{2}\int_0^\beta d\tau \sum_{\mathbf{r}} \Big\{ \Delta_x(r) \left( e^{ik_f} \Psi_{\mathrm{I},1}^\dagger(r) \Psi_{\mathrm{I},2}(r+\hat{e}_x) - e^{-ik_f} \Psi_{\mathrm{I},2}(r) \Psi_{\mathrm{I},1}^\dagger(r+\hat{e}_x) \right) + \Delta_y(r) \left( e^{ik_f} \Psi_{\mathrm{I},1}^\dagger(r) \Psi_{\mathrm{I},2}(r+\hat{e}_y) - e^{-ik_f} \Psi_{\mathrm{I},2}(r) \Psi_{\mathrm{I},1}^\dagger(r+\hat{e}_y) \right) + \mathrm{h.c.} \Big\}. \quad (2.19)$$

Near half-filling ($k_f = \pi/2$) and expanding in gradients of the $\Psi$ fields, we find

$$S_\Delta^{(i)} = \frac{1}{2}\int_0^\beta d\tau \int d^2\mathbf{r} \left\{ v_\Delta e^{i\phi(r)} \left( \Psi_{\mathrm{I},1}^\dagger(r) i\partial_y \Psi_{\mathrm{I},2}(r) - \left( i\partial_y \Psi_{\mathrm{I},1}^\dagger(r) \right) \Psi_{\mathrm{I},2}(r) \right) + \mathrm{h.c.} \right\} \quad (2.20)$$

after rotating the coordinate axes by $\pi/4$ such that $\partial_y - \partial_x \rightarrow \sqrt{2}\partial_y$. We have also explicitly included the phase fluctuations by writing $\Delta_i = \pm v_\Delta e^{i\phi}$ with $v_\Delta$ defined in (2.7). Integrating by parts the second term and using the relation

$$\frac{1}{2}\left( e^{i\phi} i\partial_y + i\partial_y e^{i\phi} \right) = e^{i\phi/2} i\partial_y e^{i\phi/2}, \quad (2.21)$$

we can write, now incorporating contributions from all nodes,

$$S_\Delta = \int_0^\beta d\tau \int d^2\mathbf{r} \left\{ v_\Delta \Psi_{\mathrm{I}}^\dagger \mathcal{B}_y^\phi \Psi_{\mathrm{I}} + (\mathrm{I} \rightarrow \mathrm{II}, y \rightarrow x) \right\}; \quad (2.22)$$

the matrix $\mathcal{B}_i^\phi$ is defined as

$$\mathcal{B}_i^\phi \equiv \begin{pmatrix} 0 & e^{i\phi/2} i\partial_i e^{i\phi/2} & 0 & 0 \\ e^{-i\phi/2} i\partial_i e^{-i\phi/2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -e^{i\phi/2} i\partial_i e^{i\phi/2} \\ 0 & 0 & -e^{-i\phi/2} i\partial_i e^{-i\phi/2} & 0 \end{pmatrix}. \quad (2.23)$$

In the absence of fluctuations, this is identical to the form derived in the previous section. However, we now clearly see how the action becomes modified in the presence of fluctuations.

It is usual at this stage to absorb the order parameter phase by a redefinition of the fields. There are countless ways of performing this change of variables, but most lead to multi-valued fields in the presence of vortex configurations. For example, consider the change of variables,

$$c_\uparrow \rightarrow e^{i\phi/2} c_\uparrow \quad (2.24)$$

$$c_\downarrow \rightarrow e^{i\phi/2} c_\downarrow. \quad (2.25)$$

As we traverse a path encircling a vortex, the phase changes by $2\pi$ and, as a result, our fields acquire an overall minus sign. Due then to the non-singular nature of the transformation, branch cuts must

be introduced which unnecessarily complicate the problem[39]. To avoid such complications, we proceed by writing the phase as a sum of a regular ("spin-wave") part and a singular part coming from vortices[39, 48]:

$$\phi(r) = \phi_{\rm reg}(r) + \phi_{\rm vort}(r). \tag{2.26}$$

Further dividing the vortex configurations into two completely arbitrary groups, $A$ and $B$, we can now choose to absorb the *entire* phase of group $A(B)$ vortices into spin up(down). In this way, as we encircle a vortex from group A (for example) with a spin up or down field, acquiring a phase of $2\pi$ or 0, respectively, our field doesn't change sign; hence, our transformation is single-valued.

To implement this procedure, introduce a new field

$$\widetilde{\Psi}_i(r) \equiv U(r)\Psi_i(r) \tag{2.27}$$

where $U(r)$ is the diagonal $4\times 4$ unitary matrix with elements $\{e^{-i\phi_A(r)}, e^{i\phi_B(r)}, e^{-i\phi_A(r)}, e^{i\phi_B(r)}\}$ and $\phi_\alpha = \phi_{\rm reg}/2 + \phi_{{\rm vort},\alpha}$, $\alpha = (A,B)$. As we can see, spin up and spin down are transformed differently, the former associated with group $A$, the latter with group $B$. Substituting this new field into our action (2.11) and including phase fluctuations via (2.22), we have

$$\begin{aligned} S_f[\widetilde{\Psi}_i] &= \int_0^\beta {\rm d}\tau \int {\rm d}^2\mathbf{r}\Big\{ \widetilde{\Psi}_{\rm I}^\dagger(r)U(r)\left[\partial_\tau + v_f M_1 \partial_x + v_\Delta \mathcal{B}_y^\phi\right] U^{-1}(r)\widetilde{\Psi}_{\rm I}(r) \\ &\quad +({\rm I}\to{\rm II}, x\leftrightarrow y)\Big\}. \end{aligned} \tag{2.28}$$

Examining this term-by-term, we will be led to consider derivatives $\partial_\mu\phi_{A,B}(r)$, which we will find convenient to rewrite as

$$\partial_\mu\phi_A(r) = \frac{1}{2}\partial_\mu(\phi_A+\phi_B) + \frac{1}{2}\partial_\mu(\phi_A-\phi_B) \equiv v_\mu + a_\mu \tag{2.29}$$

$$\partial_\mu\phi_B(r) = \frac{1}{2}\partial_\mu(\phi_A+\phi_B) - \frac{1}{2}\partial_\mu(\phi_A-\phi_B) \equiv v_\mu - a_\mu. \tag{2.30}$$

With these definitions, the action (2.28) becomes

$$\begin{aligned} S_f[\widetilde{\Psi}_i] &= \int_0^\beta {\rm d}\tau \int {\rm d}^2\mathbf{r}\left\{ \widetilde{\Psi}_{\rm I}^\dagger(r)\left[(\partial_\tau + ia_\tau) + v_f M_1(\partial_x + ia_x) + v_\Delta M_2(\partial_y + ia_y)\right]\widetilde{\Psi}_{\rm I}(r)\right. \\ &\quad \left. +({\rm I}\to{\rm II}, x\leftrightarrow y) + iv_\mu J_\mu\right\}. \end{aligned} \tag{2.31}$$

The current $J_\mu$ is defined as

$$J_\mu = \left(\sum_i \widetilde{\Psi}_i^\dagger(r)(\mathbb{1}\otimes\sigma_3)\widetilde{\Psi}_i(r), -iv_f\widetilde{\Psi}_{\rm I}^\dagger(r)(\sigma_3\otimes\mathbb{1})\widetilde{\Psi}_{\rm I}(r), -iv_f\widetilde{\Psi}_{\rm II}^\dagger(r)(\sigma_3\otimes\mathbb{1})\widetilde{\Psi}_{\rm II}(r)\right), \tag{2.32}$$

which, in terms of the original electronic operators, can be written

$$J_0 \propto \sum_{\sigma,n} c_\sigma^{n\dagger} c_\sigma^n \tag{2.33}$$

(where the sum is over spin $\sigma$ and the four nodes $n$) and

$$J_1 \propto v_f(c_\uparrow^{(i)\dagger} c_\uparrow^{(i)} - c_\downarrow^{(i)\dagger} c_\downarrow^{(i)}) + ((i) \to (iii)) \tag{2.34}$$

$$J_2 \propto v_f(c_\uparrow^{(ii)\dagger} c_\uparrow^{(ii)} - c_\downarrow^{(ii)\dagger} c_\downarrow^{(ii)}) + ((ii) \to (iv)). \tag{2.35}$$

We interpret these as the total electron density at the nodes (2.33) and the current carried by the electrons at the nodes (2.34,2.35). Taken together $J_\mu$ is the physical 3-current and we, thus, deduce that the true electromagnetic gauge potential $A_\mu$ can be introduced through the replacement $v_\mu \to v_\mu + A_\mu$. This should be contrasted with the minimally coupled way by which the field $a_\mu$ enters the theory. This also suggests how to include the effect of moving away from half-filling: tuning the electronic chemical potential $h$ is equivalent to shifting $A_0 \to A_0 + ih$. We will explicitly include this later, in Chapter 5.

As was mentioned earlier, our goal in transforming the action (2.1) as we have is to elucidate the symmetry properties of the theory. Apart from the SU(2) symmetry alluded to at the outset, we now recognize two emergent *gauge* symmetries related to the fields $a$ and $v$ in (2.31). We pause now to consider these, leaving the discussion of the global properties until a later section.

We can isolate the (as yet uncertain) charge of the $\Psi$ fields by making the usual gauge transformation $A_\mu \to A_\mu + \partial_\mu \chi$. However, we can make use of the gauge freedom afforded by $v_\mu$ to absorb this shift, leaving $a$ and, in particular, $\Psi$ unchanged. Hence, the $\Psi$ fields are electrically neutral fermions, and we will refer to them as spinons[34]. The action (2.31) is also invariant under the transformation $a_\mu \to a_\mu + \partial_\mu \chi$, $\Psi \to \Psi e^{-i\chi}$, with $v_\mu$ unchanged. Recalling the definition of the gauge field $a$, we see that this invariance reflects the freedom we had in dividing the phase field into $A$ and $B$. That is, shifting $a$ as described is tantamount to redefining the regular part of $\phi_A$ as $\phi_{\rm reg}/2 + \chi$ and that of $\phi_B$ as $\phi_{\rm reg}/2 - \chi$. As usual, we must eventually deal with this freedom by gauge fixing. We also recognize that there is an additional arbitrariness in dividing the *vortex* part of the phase into $A$ and $B$: all groupings are completely equivalent. This can be dealt with by averaging over all possible divisions, as we shall see next when considering the dynamics for the gauge fields.

## 2.4 Duality and gauge field dynamics

The gauge fields $a$ and $v$ are dependent on the fluctuating positions of the vortex defects and will acquire dynamics from them. To determine this, we assume that the fluctuations are described by the 2+1 dimensional XY-model given on the lattice by

$$Z_{XY} = \int_0^{2\pi} \left( \prod_i \mathrm{d}\phi_i \right) \exp\left\{ \widetilde{K} \sum_{i,\mu} \cos(\phi_{i+\hat{e}_\mu} - \phi_i) \right\}, \tag{2.36}$$

where $i$ labels lattice sites and $\hat{e}_\mu$ is a unit vector in the $\mu = \tau, x, y$ direction; $\widetilde{K}$ is the unrenormalized phase stiffness. Approximating the cosine using the Villain transformation[49], we have

$$Z_{XY} \approx \int_0^{2\pi} \left( \prod_i \mathrm{d}\phi_i \right) \sum_{n=-\infty}^{\infty} \exp\left\{ -\frac{\widetilde{K}}{2} \sum_{i,\mu} \left( \phi_{i+\hat{e}_\mu} - \phi_i - 2\pi n_{i,\mu} \right)^2 \right\}, \tag{2.37}$$

which can be decoupled by the Hubbard-Stratonovich transformation[10]

$$\begin{aligned} Z_{XY} \;\rightarrow\; & \int_0^{2\pi} \left( \prod_i \mathrm{d}\phi_i \right) \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}L_{i,\mu} \right) \qquad (2.38) \\ & \times \exp\left\{ -\sum_{i,\mu} \left[ \frac{1}{2\widetilde{K}} L_{i,\mu}^2 - iL_{i,\mu} \left( \phi_{i+\hat{e}_\mu} - \phi_i - 2\pi n_{i,\mu} \right) \right] \right\} \\ = \; & \int_0^{2\pi} \left( \prod_i \mathrm{d}\phi_i \right) \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}L_{i,\mu} \right) \qquad (2.39) \\ & \times \exp\left\{ -\frac{1}{2\widetilde{K}} \sum_{i,\mu} L_{i,\mu}^2 - 2\pi i \sum_i L_i \cdot n_i - i \sum_{i,\mu} \phi_i (L_{i,\mu} - L_{i-\hat{e}_\mu,\mu}) \right\}. \end{aligned}$$

Now, we can integrate over the phase field which simply acts as a Lagrange multiplier, imposing the 'no-divergence' constraint

$$\Delta \cdot L_i = \sum_\mu (L_{i,\mu} - L_{i-\hat{e}_\mu,\mu}) = 0. \tag{2.40}$$

This can be satisfied by rewriting $L_i$ as a lattice curl,

$$L_{i,\mu} = (\Delta \times S_i)_\mu = \sum_{\nu,\sigma} \epsilon_{\mu\nu\sigma} (S_{i-\hat{e}_\sigma,\sigma} - S_{i-\hat{e}_\sigma-\hat{e}_\nu,\sigma}). \tag{2.41}$$

Substituting into (2.39) and defining $m_i = \Delta \times n_i$,

$$Z_{XY} = \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}S_{i,\mu} \right) \exp\left\{ -\frac{1}{2\widetilde{K}} \sum_i (\Delta \times S_i)^2 + 2\pi i \sum_i S_i \cdot m_i \right\}. \tag{2.42}$$

The fields $S_i$ mediate long-range (Coulomb) interactions between the integers $m_i$ which we can, thus, interpret as the density of vortices of the original angle-valued fields, $\phi_i$. Due to the condition $\Delta \cdot m_i = 0$ arising from their definition, the vortices close upon themselves and we refer to (2.42) as the vortex loop representation of the XY-model.

Having isolated the singular contribution of the fluctuating phase field, we can introduce the groups $A$ and $B$ by dividing $m_i = m_{A,i} + m_{B,i}$ such that $\Delta \cdot m_{A,i} = \Delta \cdot m_{B,i} = 0$. The gauge fields $a$ and $v$ can then be incorporated by recognizing that

$$\Delta \times a = \frac{2\pi(m_A - m_B)}{2}; \; \Delta \times v = \frac{2\pi(m_A + m_B)}{2}, \tag{2.43}$$

which follows from the defining relation of the vortex density $\nabla \times \nabla\phi = 2\pi m$. The partition function is now

$$\begin{aligned} Z_{XY} \;=\; & \sum_{m_{A,B}=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}[S, a, v, r, t]_{i,\mu} \right) \\ & \times \exp\Bigg\{ -\frac{1}{2\widetilde{K}} \sum_i (\Delta \times S_i)^2 + 2\pi i \sum_i S_i \cdot (m_{A,i} + m_{B,i}) \\ & -i \sum_i \Big[ r_i \cdot (\Delta \times v_i + \Delta \times a_i - 2\pi m_{A,i}) + t_i \cdot (\Delta \times v_i - \Delta \times a_i - 2\pi m_{B,i}) \Big] \Bigg\}, \end{aligned} \tag{2.44}$$

where $r$ and $t$ are Lagrange multiplier fields imposing (2.43). The usefulness of this representation is now clear: we are able to completely sum over all possible divisions of the vortices into two groups, thus removing the arbitrariness of this assignment and the associated gauge freedom. This is done by way of the Poisson summation formula (see, for example, Jose et al. [50]), which then forces the combinations $l_A = S + r$ and $l_B = S + t$ to be integer. Then,

$$\begin{aligned} Z_{XY} \;=\; & \sum_{l_{A,B}=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}[S, a, v]_{i,\mu} \right) \exp\Bigg\{ -\frac{1}{2\widetilde{K}} \sum_i (\Delta \times S_i)^2 \\ & -i \sum_i \Big[ (\Delta \times v_i) \cdot (l_{A,i} + l_{B,i} - 2S_i) + (\Delta \times a_i) \cdot (l_{A,i} - l_{B,i}) \Big] \Bigg\}. \end{aligned} \tag{2.45}$$

Integrating over $S$ and rearranging slightly, we find

$$\begin{aligned} Z_{XY} \;=\; & \sum_{l_{A,B}=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}[a, v]_{i,\mu} \right) \exp\Bigg\{ -2\widetilde{K} \sum_i v_i^2 \\ & +i \sum_i \Big[ (v_i + a_i) \cdot (\Delta \times l_{A,i}) + (v_i - a_i) \cdot (\Delta \times l_{B,i}) \Big] \Bigg\}. \end{aligned} \tag{2.46}$$

We can now replace the $\Delta \times l_{A,B}$ terms with new (integer, divergenceless) fields $h_{A,B}$. Then loosening the integer-$h$ restriction using the Poisson summation formula, we have

$$\begin{aligned} Z_{XY} &= \lim_{x\to 0} \sum_{k_{A,B}=-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \prod_i \mathrm{d}[a, v, h, \psi_{A,B}]_{i,\mu} \right) \exp\Bigg\{ -2\widetilde{K} \sum_i v_i^2 \\ &\quad - \sum_i \left( x h_{A,i}^2 - i\Big[ (v_i + a_i) \cdot h_{A,i} + 2\pi h_{A,i} \cdot k_{A,i} + \Delta\psi_{A,i} \cdot h_{A,i} \Big] \right) \\ &\quad - \sum_i \left( x h_{B,i}^2 - i\Big[ (v_i - a_i) \cdot h_{B,i} + 2\pi h_{B,i} \cdot k_{B,i} + \Delta\psi_{B,i} \cdot h_{B,i} \Big] \right) \Bigg\}, \end{aligned} \tag{2.47}$$

where the Lagrange multiplier fields, as usual, impose the required constraints. We have also added a small chemical potential which shouldn't affect the results in the limit $x \to 0$. We can now recognize in the second and third lines of (2.47) expressions similar to the partition function of (2.39). Indeed, reversing the steps which led to (2.39) from (2.36), we arrive at

$$\begin{aligned} Z_{XY} &= \lim_{x\to 0} \int_{\infty}^{\infty} \left( \prod_i \mathrm{d}[a, v, \psi_{A,B}]_i \right) \exp\Bigg\{ -2\widetilde{K} \sum_i v_i^2 \\ &\quad + \frac{1}{2x} \sum_i \cos(\Delta\psi_{A,i} + v_i + a_i) + \frac{1}{2x} \sum_i \cos(\Delta\psi_{B,i} + v_i - a_i) \Bigg\}. \end{aligned} \tag{2.48}$$

From symmetry considerations[51, Part II, Chap. 5], it is straightforward to deduce the continuum version of the partition function: $Z_{XY} = \int \mathcal{D}[a, v, b_n] \exp(-S_b)$ with

$$\begin{aligned} S_b[b_n] &= \int_0^{\beta} \mathrm{d}\tau \int \mathrm{d}^2\mathbf{r} \Bigg\{ 2\widetilde{K} v_\mu^2 + \frac{1}{2} \sum_{n=1}^{2} \left| \left( \partial_\mu - i(v_\mu + (-1)^n a_\mu) \right) b_n \right|^2 \\ &\quad + \alpha \sum_{n=1}^{2} |b_n|^2 + \frac{\beta_1}{2} \left( \sum_{n=1}^{2} |b_n|^2 \right)^2 + \frac{\beta_2}{2} \sum_{n=1}^{2} |b_n|^4 \Bigg\}. \end{aligned} \tag{2.49}$$

A remarkable consequence of the above duality transformation is that the electronic charge that was 'lost' in going from the original BCS action (2.1) to the spinon action (2.31) has been recovered. The transformation (2.27) effectively fractionalized the electrons, the charge being carried away by the fluctuating phase and becoming attached to the bosonic fields $b_n$ as a result of the duality. To see this, introduce the electromagnetic gauge field $A$ into the original form of the XY-model (2.36) with the substitution $\Delta\phi \to \Delta\phi + Q_T A$, where $Q_T$ is the (unknown) total charge. Repeating the steps leading to the continuum dual action (2.49) produces the same final form, but with $2\widetilde{K} v_\mu^2 \to 2\widetilde{K}(v_\mu + (Q_T/2)A_\mu)^2$. Now, the usual gauge transformation $A_\mu \to A_\mu + \partial_\mu \chi$ requires that $v_\mu \to$

$v_\mu - (Q_T/2)\partial_\mu\chi$, and so the bosonic fields must transform as $b_n \rightarrow b_n \exp(-i(Q_T/2)\chi)$, proving that they each carry charge $Q_T/2$. To make this gauge transformation consistent with the invariance of the fermionic sector of the theory (2.31) as discussed in the last section, we further require that $Q_T = 2$, which is simply the charge of the original Cooper pairs. Thus, each $b$ field carries charge *one*.

Since we are interested in the dynamics of the gauge fields, we should at this point integrate out the bosonic fields so that we have an action entirely in terms of $a$, $v$ and $A$. Consider first the condensed phase of bosons, $\alpha < 0$. Minimizing the action with respect to the $b_i$, we find that

$$|\langle b_1\rangle|^2 = |\langle b_2\rangle|^2 = \frac{|\alpha|}{2\beta_1 + \beta_2} \tag{2.50}$$

and the action becomes

$$S_b \rightarrow \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^2\mathbf{r}\left\{2\widetilde{K}(v_\mu + A_\mu)^2 + \frac{|\alpha|}{2\beta_1 + \beta_2}(v_\mu^2 + a_\mu^2)\right\}. \tag{2.51}$$

We see that all of the gauge fields are massive and, thus, we identify this as the superconducting phase at low energies. Note also that $a$ and $v$ are decoupled so that we can decompose the full action (spinons + fluctuations) into spin and charge sectors. Writing

$$S[\Psi_i, b_n, a, v] = S_f[\Psi_i, a, v] + S_b[b_n, a, v] = S_{\text{spin}} + S_{\text{charge}}, \tag{2.52}$$

where $S_f$ and $S_b$ are given by equations (2.31) and (2.49), respectively, we define

$$S_{\text{spin}} = S_f - i(v_\mu + A_\mu)J_\mu + \frac{|\alpha|}{2\beta_1 + \beta_2}a_\mu^2 \tag{2.53}$$

and

$$S_{\text{charge}} = i(v_\mu + A_\mu)J_\mu + 2\widetilde{K}(v_\mu + A_\mu)^2 + \frac{|\alpha|}{2\beta_1 + \beta_2}v_\mu^2. \tag{2.54}$$

In the symmetric phase, $\langle b_i\rangle = 0$ and integrating over bosons to one loop yields

$$S_b \rightarrow \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^2\mathbf{r}\left\{2\widetilde{K}(v_\mu + A_\mu)^2 + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu v_\sigma)^2}{24\pi\alpha} + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu a_\sigma)^2}{24\pi\alpha}\right\}. \tag{2.55}$$

A further integration over $v$ generates a Maxwell term for $A$, as well, and so this clearly is not a superconducting phase. In fact, the Maxwell term for $A$ implies a charge gap and, therefore, the system is insulating. We also note that, again, $a$ and $v$ are decoupled so we can decompose the theory into spin and charge sectors given by

$$S_{\text{spin}} = S_f - i(v_\mu + A_\mu)J_\mu + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu a_\sigma)^2}{24\pi\alpha} \tag{2.56}$$

and

$$S_{\text{charge}} = i(v_\mu + A_\mu)J_\mu + 2\widetilde{K}(v_\mu + A_\mu)^2 + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu v_\sigma)^2}{24\pi\alpha}. \tag{2.57}$$

Having dealt with the gauge fields and their dynamics, we can now return to consideration of the global symmetry of the theory. In particular, we will focus on the breaking of this symmetry as we leave the superconducting state, and on the resulting order.

## 2.5 The Dirac theory and its global symmetry properties

To identify the global symmetries of the theory, it is useful to bring the action (2.31) to the convenient Dirac form. This can be obtained by constructing three matrices $\gamma_0, \gamma_1$ and $\gamma_2$ which satisfy $\gamma_0^2 = \mathbb{1}$, $\gamma_0\gamma_1 = M_1$ and $\gamma_0\gamma_2 = M_2$. It is easy to find that there are four sets (along with all continuous rotations between),

$$1: \quad \gamma_0 = \sigma_1 \otimes \mathbb{1},\ \gamma_1 = -\sigma_2 \otimes \sigma_3,\ \gamma_2 = \sigma_2 \otimes \sigma_1 \tag{2.58}$$

$$2: \quad \gamma_0 = \sigma_2 \otimes \mathbb{1},\ \gamma_1 = \sigma_1 \otimes \sigma_3,\ \gamma_2 = -\sigma_1 \otimes \sigma_1 \tag{2.59}$$

$$3: \quad \gamma_0 = \sigma_3 \otimes \sigma_2,\ \gamma_1 = \mathbb{1} \otimes \sigma_1,\ \gamma_2 = \mathbb{1} \otimes \sigma_3 \tag{2.60}$$

$$4: \quad \gamma_0 = iM_1M_2,\ \gamma_1 = iM_2,\ \gamma_2 = -iM_1, \tag{2.61}$$

and that these satisfy the Clifford algebra, $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu,\nu}$[52]. In terms of these matrices, the action (2.31) becomes

$$\begin{aligned} S[\Psi_i] \;=\; & \int_0^\beta \mathrm{d}\tau \int \mathrm{d}^2\mathbf{r} \left\{ \bar{\Psi}_{\mathrm{I}}(r) \left[\gamma_0(\partial_\tau + ia_\tau) + v_f\gamma_1(\partial_x + ia_x) + v_\Delta\gamma_2(\partial_y + ia_y)\right] \Psi_{\mathrm{I}}(r) \right. \\ & \left. +(\mathrm{I} \to \mathrm{II}, x \leftrightarrow y) + i(v_\mu + A_\mu)J_\mu \right\}, \end{aligned} \tag{2.62}$$

where $\bar{\Psi}_i \equiv \Psi_i^\dagger\gamma_0$ and we are dropping tildes on the $\Psi$ fields from here on in.

The convenience of the Dirac form stems from its symmetry properties made manifest by the properties of the $\gamma$-matrices. For example, it can be shown that the matrices

$$L_{\alpha\beta} \equiv [\gamma_\alpha, \gamma_\beta];\ \alpha, \beta = 0, 1, 2 \tag{2.63}$$

satisfy the Lorentz algebra and, thus, generate the Lorentz group provided the $\gamma$ matrices satisfy the Clifford algebra. Then, the $L_{0i}$ generate Lorentz boosts and the $L_{ij}$ generate spatial rotations $(i, j = 1, 2)$.

One could correctly note that a two-component (Dirac) representation would suffice to represent this algebra in 3 (= 1+2) dimensions, with the Pauli matrices playing the role of the $\gamma_\mu$. However, there is no fourth matrix which anticommutes with the Pauli matrices and so the Lorentz symmetry is the maximal global symmetry in this representation.

By contrast, in the four-component representation, it is in most cases possible to find such a fourth (and fifth) anticommuting matrix. Then, an additional U(2) symmetry (per Dirac field), referred to as *chiral* in the field theory literature and hereafter denoted $U_c(2)$, can be identified, its generators given by $\{\mathbb{1}, \gamma_3, \gamma_5, \gamma_{35}\}$ where $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$ and $\gamma_{35} = i\gamma_3\gamma_5$. Referring to the numbering of (2.58-2.61):

$$1: \quad \gamma_3 = \sigma_2 \otimes \sigma_2,\ \gamma_5 = \sigma_3 \otimes \mathbb{1},\ \gamma_{35} = -\sigma_1 \otimes \sigma_2 \tag{2.64}$$

$$2: \quad \gamma_3 = \sigma_3 \otimes \mathbb{1},\ \gamma_5 = \sigma_1 \otimes \sigma_2,\ \gamma_{35} = -\sigma_2 \otimes \sigma_2 \tag{2.65}$$

$$3: \quad \gamma_3 = \sigma_1 \otimes \sigma_2,\ \gamma_5 = \sigma_2 \otimes \sigma_2,\ \gamma_{35} = -\sigma_3 \otimes \mathbb{1} \tag{2.66}$$

$$4: \quad \text{Can't be done.} \tag{2.67}$$

The full theory, therefore, has a global $U_c(2)\times U_c(2)$ symmetry, one factor for each $\Psi$ field; in the isotropic limit, $v_f = v_\Delta$, this expands to $U_c(4)$.

The significance of the chiral symmetry is evident when we consider the massive theory. The term $m\Psi_i^\dagger\gamma_0\Psi_i$ breaks the chiral symmetry, reducing the $SU_c(2)$ subgroup to $U_c(1)$ generated by $\gamma_{35}$. This should be compared with the two-component representation of the theory in which the symmetry is the same whether the fields are massless or not.

Such a term can arise through a process known as dynamical mass generation, which is an example of spontaneous symmetry breaking whereby some order parameter not sharing the same invariance properties as the action acquires a non-zero vacuum expectation value. In the present case, the order parameter is a composite operator $(\bar{\Psi}_i\Psi_i)$ and, by convention, we refer to its condensation as *dynamical* symmetry breaking. In the low-energy action (2.62) we are considering, the relevant dynamics are associated with the fluctuating gauge field $a$ and were determined in the last section (2.51 and 2.55).

Of course, the exact form of the action for $a$ will determine whether the mass term is generated or not. In the next section, we'll focus on the non-superconducting state and consider the role of dynamical symmetry breaking there.

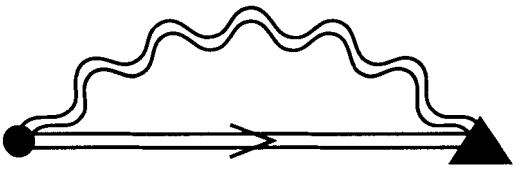

Figure 2.2: The fermion self energy , $\Sigma(q)$. The straight, double line is the full fermion propagator, the wiggly, double line is the full gauge field propagator, $D_{\mu\nu}(p-q)$; the circle is the bare vertex $\gamma_\mu$ and the triangle is the full vertex function.

## 2.6 $QED_3$ and chiral symmetry breaking

In the insulating state, the spin sector of the theory (2.56) can be written,

$$S_{\rm spin} = \int_0^\beta d\tau \int d^2\mathbf{r} \Big\{ \bar{\Psi}_{\rm I} \left[ \gamma_0(\partial_\tau + ia_\tau) + v_f\gamma_1(\partial_x + ia_x) + v_\Delta\gamma_2(\partial_y + ia_y) \right] \Psi_{\rm I} + ({\rm I} \to {\rm II}, x \leftrightarrow y) + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu a_\sigma)^2}{24\pi\alpha} \Big\}. \tag{2.68}$$

We can neglect the charge sector (2.57) by observing that integration over $v$ will yield only short-range interactions between the $\Psi$ fields, which are, thus, irrelevant at low-energies[20]. Making the further simplification that $v_f = v_\Delta = 1$, the low-energy effective theory for the insulating state is then described, at zero temperature, by quantum electrodynamics in 3(=1+2) dimensions:

$$S_{\rm QED_3} = \int d^3r \left\{ \bar{\Psi}_n \gamma_\mu D_\mu \Psi_n + \frac{(\epsilon_{\mu\nu\sigma}\partial_\nu a_\sigma)^2}{24\pi\alpha} \right\}. \tag{2.69}$$

We use the notation $D_\mu = \partial_\mu + ia_\mu$ and assume a sum over $N_f$ fermion flavours; for us, $N_f$=2.

$QED_3$ is a theory which is well known to exhibit dynamical symmetry breaking by the generation of a fermion mass term[53, 54]. To see this, we write down the fermion self energy (see Fig. 2.2) as

$$\Sigma(q) = \int \frac{d^3p}{(2\pi)^3} \frac{\gamma_\mu D_{\mu\nu}(p-q)\Sigma(p)\gamma_\nu}{p^2 + \Sigma^2(p)}, \tag{2.70}$$

where the gauge field propagator is, in the Landau gauge,

$$D_{\mu\nu}(p) = 12\pi\alpha \left( \frac{\delta_{\mu\nu} - \hat{p}_\mu\hat{p}_\nu}{p^2 + \Pi(p)} \right). \tag{2.71}$$

In writing (2.70), we have neglected wave function renormalization and vertex corrections; these effects should be included in more sophisticated treatments, but apparently most results are in agreement with this simple treatment[55, 56]. Assuming a finite mass for the fermions, $\Sigma(0) = m$, the

polarization to lowest order in $1/N_f$ is

$$\Pi(p) = 3mN_f\alpha\left[2+\frac{p^2-4m^2}{mp}\arcsin\left(\frac{p}{\sqrt{p^2+4m^2}}\right)\right] \quad (2.72)$$

$$\equiv 3mN_f\alpha F(p/m). \quad (2.73)$$

The self energy (2.70) can then be written as

$$1=\frac{4}{\pi N_f}\int_0^{\Lambda/m}\mathrm{d}p\frac{p^2\Sigma(p)}{\Big(p^2+\Sigma^2(p)\Big)\Big(mp^2/(3N_f\alpha)+F(p)\Big)}, \quad (2.74)$$

where we have rescaled $p/m \to p$ and $\Sigma(p)/m \to \Sigma(p)$. It is interesting to compare this expression with the (perhaps) more familiar gap equation in the BCS theory of superconductivity. Indeed, the symmetry breaking which occurs there is also dynamical, since the order parameter is again a composite operator. Now, we notice that the righthand side of equation (2.74) is a decreasing function of the mass, so, it is clear that this equation can always be satisfied by taking $N_f$ to be small enough. We, thus, conclude that dynamical mass generation occurs in $QED_3$ for $N_f$ less than some critical number, $N_f^c$. The value of $N_f^c$ is determined from equation (2.74) at $m=0$:

$$N_f^c=\frac{4}{\pi}\int_0^{\infty}\mathrm{d}p\frac{p^2\Sigma(p)}{\Big(p^2+\Sigma^2(p)\Big)F(p)}. \quad (2.75)$$

Evaluation of the above integral obviously depends on $\Sigma(p)$ which must be self-consistently determined from (2.70); the full solution yields $N_f^c = 32/\pi^2$. The mass for fixed $N_f$ can also be determined from this procedure and one finds that $m \approx (192\pi\alpha)\exp\left[-2\pi/\sqrt{N_f^c/N_f-1}\right]$.

Importantly, the critical value $N_f^c = 3.24$ is greater than the $N_f = 2$ in our case. More complete treatments also generally find $N_f^c \approx 3$, although there is still some controversy with regard to this. In any case, whether $N_f^c$ is greater than or less than 2, we expect the chiral symmetry to be broken at some point, if not immediately upon exiting the superconducting state, then deeper in the insulating state due to irrelevant terms such as repulsive interactions[20].

The physical significance of chiral symmetry breaking can be seen by rewriting the dynamically generated mass term using the original electronic operators. Of course, the form that obtains depends on the choice made for the matrix $\gamma_0$. Referring to the notation of (2.58-2.61), we denote the respective matrix as ${}^s\gamma_0$, $s=1\ldots4$. For $s=1$,

$$m\Psi_i^\dagger({}^1\gamma_0)\Psi_i = m\left[c_\uparrow^\dagger(\mathbf{K}_i+\mathbf{q})c_\uparrow(-\mathbf{K}_i+\mathbf{q})-c_\downarrow^\dagger(\mathbf{K}_i-\mathbf{q})c_\downarrow(-\mathbf{K}_i-\mathbf{q})+\text{h.c.}\right]$$

$$\to 2m\sum_{\sigma=\pm1}\sigma\cos(2\mathbf{K}_i\cdot\mathbf{r})c_\sigma^\dagger(r)c_\sigma(r), \quad (2.76)$$

which we can recognize as staggered spin density wave (SDW) order[20, 57], the periodicity of which is set by the node-spanning wave vector, $2\mathbf{K}_i$. This is particularly intriguing since the state at half-filling is known to be an anti-ferromagnetic insulator, thus, the SDW is a natural candidate for the correct doping-evolved state. Repeating the above exercise for $s = 2$ yields a sine-SDW rather than the cosine-SDW of $s = 1$.

For $s = 3$, the situation is slightly different; in this case, the mass acts like a particle-particle potential, encouraging the opening of a pairing gap in addition to the already established $d$-wave gap. The sign of the potential is reversed for opposing nodes, and so we interpret this as $p$-wave pairing. The resultant state, then, is a $d + ip$ insulator.

The mass term associated with the fourth set of gamma matrices is quite different from the other three. As we previously noted, the fourth set does not allow a representation of the chiral $\mathrm{SU_c}(2)$, and so the mass term does not break this symmetry. The mass term does, however, break parity (e.g., $x \rightarrow -x$) and time-reversal symmetry ($t \rightarrow -t$), both of which are expected to be preserved in $\mathrm{QED_3}$. For this reason, the term $m\Psi_i^\dagger({}^4\gamma_0)\Psi_i$ is not expected to become dynamically generated[20].

It is useful to notice that the sets of generators of $\mathrm{SU_c}(2)$ (2.64-2.66), corresponding to the different $\gamma_0$ matrices, are simply permutations of each other. Consequently, under the action of a broken generator, the order parameter is rotated in the direction of the $\gamma_0$ for which the broken generator is its $\gamma_{35}$. That is, the transformation of the fields $\Psi \rightarrow \exp[i\vartheta\widetilde{\gamma}_{35}]\Psi$ rotates $\gamma_0$ as

$$e^{-i\vartheta\widetilde{\gamma}_{35}}\gamma_0 e^{i\vartheta\widetilde{\gamma}_{35}} = \cos(2\vartheta)\gamma_0 + \sin(2\vartheta)\widetilde{\gamma}_0, \tag{2.77}$$

where $\widetilde{\gamma}_{35}$ is either $\gamma_3$ or $\gamma_5$. This is nicely summarized in Figure 2.3. Thus, the different orders are degenerate and related by chiral rotations. In the isotropic $\mathrm{U_c}(4)$ case, the manifold of possible (degenerate) states is increased and includes, for example, one-dimensional charge stripe ordering and a $d + is$ insulator[20]. However, including perturbations such as a weak repulsion or anisotropy lifts the degeneracy, and it has been found that the lowest energy state is the SDW[58].

## 2.7 Outlook

The $\mathrm{QED_3}$ theory has been used recently to study many aspects of cuprate superconductors and the pseudogap state. The fine structure of chiral symmetry breaking has been studied in Seradjeh and Herbut [58], and the effect of velocity anisotropy on the antiferromagnetic instability has been

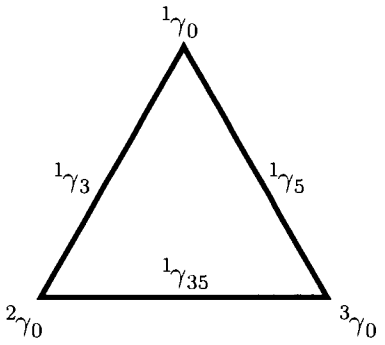


Figure 2.3: $SU_c(2)$ triangle representing the relationship between different orders and broken generators [20].

studied in Lee and Herbut [59]; the lattice version was examined in Lee and Herbut [60] and spin response was considered in Herbut and Lee [61]. This theory was also applied to the question of coexistence between antiferromagnetism and superconductivity in the pseudogap[62].

In the remainder of this thesis, we will apply this framework to the questions of spinon confinement (Chapter 3) and the superfluid density (Chapters 4 and 5); in Chapter 6, we will consider the closely related dirty boson problem.

# Chapter 3

# Confinement

In discussing the emergent gauge fields $a$ and $v$ in the preceding Chapter, we ignored a subtlety arising in the dual formalism of the XY model describing the phase fluctuations. Re-examining the partition function (2.46), we notice that the $v$ field appears in a quadratic term, a feature not shared by $a$. The consequence of this term is to render the action non-invariant under shifts of $v$, while $a$ may be shifted by integer multiples of $2\pi$ with no effect on the theory. The gauge field $a$ then appears to take on the significance of an angular variable, defined on the compact interval $[0, 2\pi)$, while $v$ takes on values from the entire real line. Thus, we are tempted to distinguish between $a$ and $v$ as being *compact* and *non-compact* gauge fields, respectively.

However, there is a further subtlety, since $a$ and $v$ both derive from the same set of vortex defects which form closed loops. A necessary characteristic of compact gauge fields is that they give rise to vortex lines, terminating in sources of magnetic flux (monopoles). Clearly, $a$ does not satisfy this condition and we must conclude that it is not compact.

Compact gauge fields do arise, though, in other theories of strongly correlated systems[63], and their study constitutes an important and open problem. We will consider the issue of confinement in such theories, mentioning later what is expected for the non-compact versions.

## 3.1 Introduction

Gauge theories of compact U(1) fields in three dimensions have long been of interest to researchers in high energy physics. For particle physicists, they serve as relatively simple models exhibiting non-perturbative phenomena such as chiral symmetry breaking[53, 54] and confinement[40], be-

lieved to be crucial to our understanding of more realistic theories like quantum chromodynamics. Additionally, assuming compactness for the gauge fields of nature provides, by consequence, a natural explanation for the quantization of charge[64], an issue which remains unresolved.

Now, as we have mentioned, such 'toy models' arise naturally as real, effective descriptions of strongly correlated systems in two spatial dimensions and at zero temperature, elevating them from the realm of training grounds for more fundamental theories, to the level of actual physically relevant models. As such, the intriguing questions studied in the high energy community are now important for condensed matter physicists. For example, we have already seen how the antiferromagnetic state of the cuprate superconductors arises naturally from the breaking of chiral symmetry. In the present Chapter, we will address the issue of confinement in the context of our effective compact $QED_3$ ($cQED_3$) description of the pseudogap state. This work has previously been published in Case et al. [41].

Confinement is a crucial issue in all compact U(1) theories and is believed to result from the unbinding of magnetic monopoles, which are invariably present due to the compact nature of the gauge field[40]. In a pioneering work[40], Polyakov showed that, in pure compact quantum electrodynamics without matter and in $d = 3$, confinement is permanent for all values of the gauge coupling. The situation where the gauge field is coupled to matter is more subtle and has been the subject of considerable debate. It has been argued that coupling to relativistic massless fermions transforms the usual Coulombic interaction between monopoles into the much longer-ranged logarithmic interaction at large distances[33, 65–69]. When applied to a single monopole-antimonopole pair, this would suggest that monopoles may bind into dipoles, in analogy with the celebrated Berezinskii[70], Kosterlitz and Thouless[71] (BKT) transition in two dimensions. However, while the effects of a finite density of monopoles on the BKT transition in $d = 2$ are well understood[71, 72], the situation in $d = 3$ appears less clear[73]. The difficulty lies in the fact that, while the screening in the dipole phase in $d = 2$ just amounts to a renormalization of the dielectric constant, in $d = 3$ it changes the form of the interaction[74–77]. This point of view was expounded in Herbut and Seradjeh [74], using a renormalization group treatment of the problem, and extended to the case of non-relativistic fermions in Herbut et al. [75], the results being that screening transforms the logarithmic potential back into the Coulombic form, and the deconfined phase is always unstable to monopole unbinding.

In the remainder of the Chapter, we introduce the anomalous sine-Gordon action, which is dual to the original $cQED_3$. Using a variational method, we first recall Polyakov's original result within this framework and, then, present a generalized method which allows us to account for screening.

We will see that, in the absence of screening, the theory indeed undergoes a deconfinement transition, but that this solution is unstable when screening is taken into account. We begin by clarifying the difference between the non-compact and compact versions of QED.

## 3.2 Compact vs. non-compact QED

The distinction between the non-compact and compact versions of QED arises when we consider the theory on a lattice. The pure gauge sector can be written in two ways, both of which must reproduce the continuum limit[78],

$$S_{\rm G} = \frac{1}{4q^2}\int \mathrm{d}^3\mathbf{r}(\nabla\times\mathbf{a})^2, \tag{3.1}$$

as the lattice constant $l \to 0$. Writing the curl in the above equation in its simplest discrete form,

$$F_{\mu\nu} = \frac{1}{l}\left[a_\nu(\mathbf{r}+\hat{e}_\mu) - a_\nu(\mathbf{r}) + a_\mu(\mathbf{r}+\hat{e}_\nu) - a_\mu(\mathbf{r})\right], \tag{3.2}$$

the *non-compact* lattice version of (3.1) is simply

$$S_{\rm nc} = \frac{1}{4\hat{q}^2}\sum_{\mathbf{r},\mu,\nu}(l^2F_{\mu\nu})^2, \tag{3.3}$$

where the combination $l^2F_{\mu\nu}$ is dimensionless; $\hat{q}$ is the dimensionless, bare charge, about which we will have more to say later. In another formulation, use is made of the elements $U_\mu(\mathbf{r})$ of the *compact* group $U(1)$:

$$U_\mu(\mathbf{r}) = e^{ila_\mu(\mathbf{r})}. \tag{3.4}$$

The quantity $la_\mu$ takes on values from the compact interval $(0, 2\pi]$; hence, the designation 'compact QED'. As $l \to 0$, the domain of $a$ must span the entire real axis, and, so, the continuum limit is clearly non-compact. The gauge invariant action for $a$ is the product of these $U(1)$ factors around an elementary plaquette, and has the form,

$$S_{\rm c} = \frac{1}{2\hat{q}^2}\sum_{\mathbf{r},\mu,\nu}\left[1-\cos(l^2F_{\mu\nu})\right]. \tag{3.5}$$

This form has the proper continuum limit (3.1) and is clearly different from the non-compact case (3.3).

It should be noted that, in taking the continuum limit, we must identify $q^2 = \hat{q}^2/l$. Thus, $\hat{q}$ cannot be held fixed as $l \to 0$ but, instead, must also vanish such that $q$ is constant in this limit.

The continuum limit then corresponds to either $l \longrightarrow 0$ or $\hat{q} \longrightarrow 0$. The relevance of this will become apparent when we consider the compact, pure gauge theory (3.5).

Pure gauge compact $QED_3$ (3.5) was first studied by Polyakov, who showed that non-dynamical fermions are permanently confined due to the monopoles arising from the compact gauge field. We will consider this simpler example before moving on to the theory with *dynamical*, massless fermions coupled to the compact gauge field. First, however, we will introduce the duality transformations on the action (3.5) to arrive at an action for the monopoles only.

## 3.3 Duality transformations

Compactness of the gauge field leads to topologically non-trivial configurations which are known as monopoles, in exact analogy with the vortices of the XY model. As we saw in the last chapter, it is possible to perform transformations on the original action to arrive at a theory in terms of these defect degrees of freedom. The duality transformations on the action (3.5) are quite similar to those of Chapter 2 and we will try to adopt the same notation here. We begin as usual by writing the cosine term in the Villain approximation. However, to facilitate later results, let us generalize this as

$$S_{\rm c} \approx \frac{1}{2} \sum_{\mathbf{r},\mu,\nu} \Big( F_{\mu\nu}(\mathbf{r}) - 2\pi n_{\mu\nu}(\mathbf{r}) \Big) u(\mathbf{r}) \Big( F_{\mu\nu}(\mathbf{r}) - 2\pi n_{\mu\nu}(\mathbf{r}) \Big). \tag{3.6}$$

For the present case of the pure gauge theory, $u(\mathbf{r}) = 1/(2\hat{q}^2)$. In writing the above action, we have dropped the constant term and set $l = 1$ for clarity. A Hubbard-Stratonovich transformation then yields

$$S_{\rm c} \longrightarrow \frac{1}{2} \sum_{\mathbf{r},\mu,\nu} M_{\mu\nu}(\mathbf{r}) u^{-1}(\mathbf{r}) M_{\mu\nu}(\mathbf{r}) + i \sum_{\mathbf{r},\mu,\nu} M_{\mu\nu}(\mathbf{r}) \Big( F_{\mu\nu}(\mathbf{r}) - 2\pi n_{\mu\nu}(\mathbf{r}) \Big). \tag{3.7}$$

Representing $F_{\mu\nu} = \epsilon_{\mu\nu\sigma}\epsilon_{\sigma\upsilon\rho}\Delta_\upsilon a_\rho$, and defining $L_\sigma = \epsilon_{\sigma\mu\nu}M_{\mu\nu}$ and $m_\sigma = (1/2)\epsilon_{\sigma\mu\nu}n_{\mu\nu}$, (3.7) becomes

$$S_{\rm c} = \frac{1}{4} \sum_{\mathbf{r},\sigma} L_\sigma(\mathbf{r}) u^{-1}(\mathbf{r}) L_\sigma(\mathbf{r}) + i \sum_{\mathbf{r}} L(\mathbf{r}) \cdot \Big( \Delta \times a(\mathbf{r}) - 2\pi m(\mathbf{r}) \Big). \tag{3.8}$$

Integrating over $a$ will now force $L$ to be curl-free, and so we can write $L_\sigma = \Delta_\sigma \varphi$. A further integration over $\varphi$ leads to

$$S_{\rm c} = \frac{1}{2} \sum_{\mathbf{r},\mathbf{r}'} \rho(\mathbf{r}) V(\mathbf{r},\mathbf{r}') \rho(\mathbf{r}'), \tag{3.9}$$

where the monopole density is $\rho = \Delta_\sigma m_\sigma$ and $V^{-1} = -1/(8\pi^2)\Delta_{\mathbf{r},\sigma} u^{-1} \Delta_{\mathbf{r}',\sigma}$. The monopole density may be expressed as $\rho(\mathbf{r}) = \sum q_a \delta_{\mathbf{r},\mathbf{r}_a}$, so this action (3.9) represents a gas of charges

(monopoles) interacting via a potential $V(|\mathbf{r}_a - \mathbf{r}_b|)$. For the case $u = 1/(2\hat{q}^2)$, the potential has the Coulomb form,

$$\begin{aligned} V(|\mathbf{r}_a - \mathbf{r}_b|) &= \frac{4\pi^2}{\hat{q}^2} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{e^{i\mathbf{k}\cdot(\mathbf{r}_a - \mathbf{r}_b)}}{k^2} \\ &= \frac{\pi}{\hat{q}^2 |\mathbf{r}_a - \mathbf{r}_b|}. \end{aligned} \tag{3.10}$$

The action (3.9) with the Coulomb potential is well known to be equivalent to the sine-Gordon equation[50]. We present a derivation here, but with a general expression for $u(\mathbf{r})$.

Return to the action (3.8) after having integrated out $a$, and add a fugacity for the monopoles so that

$$S_{\mathrm{c}} \to \frac{1}{4} \sum_{\mathbf{r},\sigma} \Delta_\sigma \varphi(\mathbf{r}) u^{-1}(\mathbf{r}) \Delta_\sigma \varphi(\mathbf{r}) + 2\pi i \sum_{\mathbf{r}} \varphi(\mathbf{r}) \rho(\mathbf{r}) + \frac{1}{4y} \sum_{\mathbf{r}} \rho^2(\mathbf{r}); \tag{3.11}$$

the limit $y \to \infty$ is assumed. Removing the integer constraint on the density $\rho$ by way of the Poisson formula[50], and integrating over the resulting real field, we find

$$S_{\mathrm{c}} \to -\frac{1}{4} \sum_{\mathbf{r},\sigma} \varphi(\mathbf{r}) \Delta_\sigma u^{-1}(\mathbf{r}) \Delta_\sigma \varphi(\mathbf{r}) + y \sum_{\mathbf{r}} (2\pi\varphi(\mathbf{r}) - 2\pi n(\mathbf{r}))^2. \tag{3.12}$$

Finally, rescaling $2\pi\varphi \to \varphi$, the latter action becomes, up to a Villain approximation,

$$S_{\mathrm{c}} \to \frac{1}{2} \sum_{\mathbf{r}} \varphi(\mathbf{r}) V^{-1}(\mathbf{r}) \varphi(\mathbf{r}) - 2y \sum_{\mathbf{r}} \cos\varphi(\mathbf{r}), \tag{3.13}$$

which, for $u(\mathbf{r}) = 1/(2\hat{q}^2)$, is the usual sine-Gordon equation with effective temperature $\hat{q}^2/(4\pi^2)$.

We will focus on the dual action (3.13) to establish confinement in the original theory (3.6). In the next section, we reproduce Polyakov's result that no transition occurs in pure gauge cQED$_3$. The remainder of the chapter will then be dedicated to the theory coupled to massless fermions.

## 3.4 Permanent confinement in pure gauge cQED$_3$

Polyakov focused on the sine-Gordon theory given by

$$S_{\mathrm{SG}} = \int \mathrm{d}^3\mathbf{r} \left[ \frac{\hat{q}^2}{4\pi^2} (\nabla\varphi)^2 - 2y \cos\varphi \right] \tag{3.14}$$

and showed that no transition exists for any value of the gauge coupling $\hat{q}$ or, equivalently, effective temperature $\widetilde{T} \equiv \hat{q}^2/(4\pi^2)$. We will demonstrate this result using a variational approach based on the Gibbs-Bogoliubov-Feynman (GBF) inequality[79].

The GBF inequality establishes a strict upper bound on the free energy of a system through the relation

$$F \leq F_{\rm var} \equiv F_0 + \langle S - S_0 \rangle_0 \,, \tag{3.15}$$

where $S_0$ is some trial action chosen to approximate $S$; $F_0$ is the free energy associated associated with $S_0$ and $\langle ... \rangle_0$ represents an average within this ensemble. We are free to choose any form for the trial action, the inequality (3.15) being completely general, but for simplicity in evaluating averages, we choose the Gaussian form[80]

$$S_0 = \frac{1}{V} \sum_{\mathbf{q}} \frac{1}{2} \varphi(\mathbf{q}) G_0^{-1}(\mathbf{q}) \varphi(\mathbf{q}). \tag{3.16}$$

The evaluation of $F_{\rm var}$ is now straightforward. With $F = F_{\rm SG}$, this yields (up to a constant)

$$\begin{aligned} \frac{F_{\rm var}}{V} &= -\frac{1}{2} \int \frac{{\rm d}^3\mathbf{q}}{(2\pi)^3} \ln \left( G_0(\mathbf{q}) \right) + \frac{\widetilde{T}}{2} \int \frac{{\rm d}^3\mathbf{q}}{(2\pi)^3} q^2 G_0(\mathbf{q}) \\ &\quad -2y \, \exp \left\{ -\frac{1}{2} \int \frac{{\rm d}^3\mathbf{q}}{(2\pi)^3} G_0(\mathbf{q}) \right\}. \end{aligned} \tag{3.17}$$

In deriving the above expression, we have used the result for the propagator

$$\left\langle \varphi(\mathbf{q}) \varphi(\mathbf{q}') \right\rangle_0 = V G_0(\mathbf{q}) \delta^3(\mathbf{q} + \mathbf{q}') \tag{3.18}$$

and the identity

$$\langle \cos \varphi(\mathbf{r}) \rangle_0 = e^{-\frac{1}{2} \langle \varphi^2(\mathbf{r}) \rangle_0}. \tag{3.19}$$

To take full advantage of the inequality (3.15), we now minimize $F_{\rm var}$ with respect to $G_0$ to produce the optimal Gaussian theory approximating $F_{\rm SG}$:

$$\frac{\delta F_{\rm var}}{\delta G_0} = 0 \implies G_0^{-1}(\mathbf{q}) = \widetilde{T} q^2 + m, \tag{3.20}$$

where the 'mass' $m$ is determined implicitly through

$$\begin{aligned} m &= 2y \exp \left\{ -\frac{1}{2} \int \frac{{\rm d}^3\mathbf{q}}{(2\pi)^3} \frac{1}{\widetilde{T} q^2 + m} \right\} \\ &= 2y \exp \left\{ \frac{1}{4\pi^2 \widetilde{T}} \left( \Lambda - \sqrt{\frac{m}{\widetilde{T}}} \arctan \sqrt{\frac{\widetilde{T} \Lambda^2}{m}} \right) \right\}; \end{aligned} \tag{3.21}$$

$\Lambda$ is the ultra-violet (UV) cutoff. The right-hand side (RHS) has the following limits:

$$\text{RHS} \to \begin{cases} 2y e^{\frac{\Lambda}{4\pi^2 \widetilde{T}}}, & m \to 0 \\ 2y, & m \to \infty \end{cases}. \tag{3.22}$$

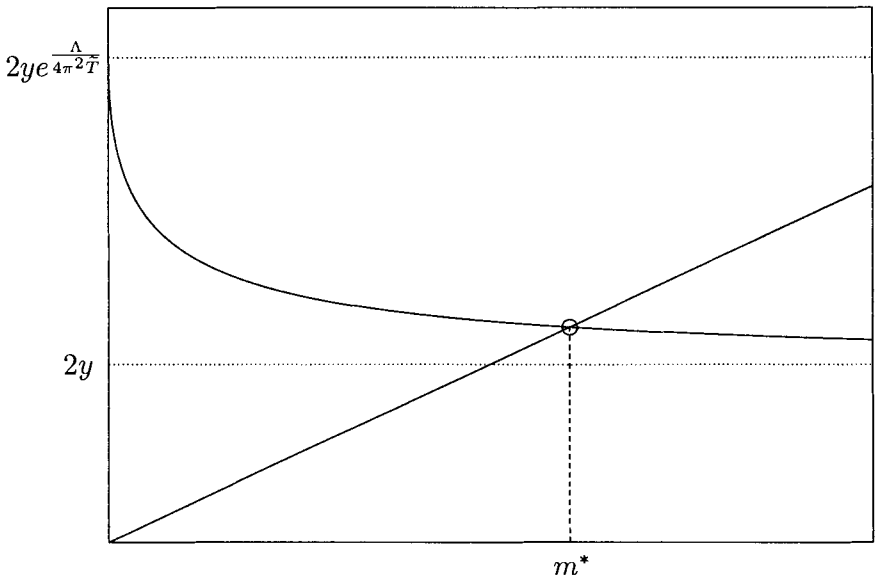


Figure 3.1: The solid curves are the left- and right-hand sides of (3.21); the dotted lines are the limits given in (3.22). The solution $m^*$ is determined by the intersection of the solid lines, as indicated.

Thus, the only solution of (3.21) for any value of $\widetilde{T}$ has a finite mass, $m^*$. This is depicted graphically in Figure 3.1 where we plot both the left- and right-hand side, and the solution is the intersection of these curves.

To understand what the finite mass solution means, look at the monopole density. This is determined from

$$\begin{aligned} \rho_{\rm M} &= -\frac{1}{V}\frac{\partial F_{\rm var}}{\partial \mu} \\ &= -\frac{y}{V}\frac{\partial F_{\rm var}}{\partial y} \\ &= m, \end{aligned} \tag{3.23}$$

where we have used the definition of the fugacity $y = \exp\{\mu\}$. It is clear then that a finite $m$ solution corresponds to a phase of (free) monopoles. Since only a finite $m$ solution exists for (3.21), we conclude that the monopoles are always free in the theory (3.14). To be completely clear, we mention that a solution with $m^* = 0$ corresponds to $\rho_M = 0$, from which we would conclude that monopoles form into bound dipoles. In two dimensions, the theory (3.14) indeed possesses such a solution, and the accompanying transition is of the celebrated Berezinskii-Kosterlitz-Thouless (BKT) type.

To elucidate the relation between the phase of monopoles and confinement, we require a (gauge invariant) operator which distinguishes between the confined and deconfined phase. The Wilson loop is believed to be one such[64], and is defined as

$$W(C) = \exp\left\{ i \oint_C a_\mu \mathrm{d}x_\mu \right\}. \tag{3.24}$$

Polyakov showed[64] that the expectation value of (3.24) satisfies the 'area law':

$$\langle W(C) \rangle_{\mathrm{c}} = e^{-\gamma S_C}, \tag{3.25}$$

where $S_C$ is the minimal area bounded by the contour $C$. In words, the Wilson loop measures the energy required to separate two 'test charges' (non-dynamical fermions) by a distance $R$, propagate the pair for a time $T$, and bring them back together; this defines the contour of integration in (3.24). The exponent can be written $\gamma S_C = E(R)T$ where $E(R) = \gamma R$ is the energy required to separate the fermions. Thus, the area law indicates a potential between the test charges which increases linearly with separation, the hallmark of confinement.

At this point, it is useful to construct a phase diagram for the fermions in $\mathrm{QED}_3$, based on the results we have discussed so far. We have just seen that the theory with $N_{\mathrm{f}} = 0$ has no transition for any value of the coupling constant, $\hat{q}$. On the other hand, we saw in the last chapter that, in the continuum limit $\hat{q} = 0$, the theory has a chiral symmetry breaking transition for a critical number of fermions $N_{\mathrm{f}} = N_{\mathrm{f}}^{\mathrm{c}}$. These two limits correspond to the axes of the diagram shown in Figure (3.2). Coupling the theory to fermions ($N_{\mathrm{f}} \neq 0$) leaves open the issue whether the transition in the continuum limit extends to finite $\hat{q}$. Indeed, there are compelling reasons to believe that perhaps it does, as we discussed earlier, and such a possibility is indicated in the Figure by the line labeled $N_{\mathrm{f}}^{?}$. This is just the question we will address in the remainder of this chapter, using the dual formalism introduced above. A final subtlety of the phase diagram as drawn has to do with the relationship between confinement and chiral symmetry breaking. We are making the assumption that the theory has a single transition, if at all, so that the two phenomena go hand-in-hand[81]. Furthermore, based on numerical results[81], we expect that confinement will coincide with chiral symmetry *breaking*, thus explaining our choice of plotting $N_{\mathrm{f}}$ rather than $\widetilde{T} \sim 1/N_{\mathrm{f}}$.

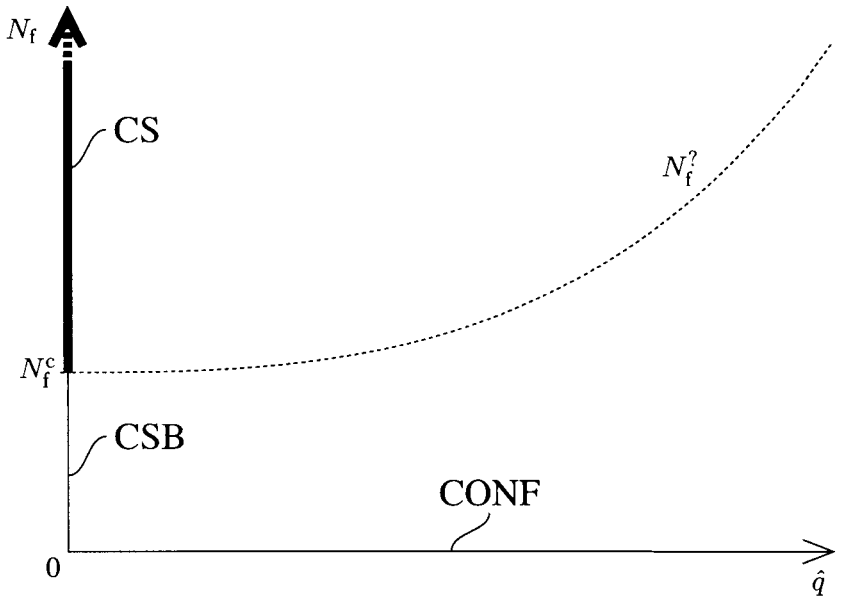


Figure 3.2: The known and possible phases of cQED$_3$. The transition in the continuum ($\hat{q} = 0$) is between chirally symmetric (CS) and chiral symmetry broken (CSB) phases; pure gauge cQED$_3$ ($N_f = 0$) is permanently confining for any value of the coupling. The possibility of a boundary between confined and deconfined phases is indicated by the dashed line $N_f^?$.

## 3.5 Introducing fermions: the anomalous sine-Gordon theory

We are interested in the phases of cQED$_3$, with the gauge field coupled to massless relativistic fermions on a lattice:

$$S[\chi, a] = S_F[\chi, a] - \frac{1}{2\hat{q}^2} \sum_{\mathbf{r},\mu,\nu} \cos\left(F_{\mu\nu}(\mathbf{r})\right). \tag{3.26}$$

The sites of the three dimensional quadratic lattice are labeled by $r_\mu = \{\tau, x, y\}$. $S_F$ is the lattice action of massless fermions coupled to the gauge field which reduces in the continuum limit to QED$_3$ with $N_f$ flavours of four-component Dirac spinors. Using staggered fermions[78], this takes the form

$$S_F[\chi, a] = \frac{1}{2} \sum_{\mathbf{r},\mu} \sum_{n=1}^{N_f/2} \eta_\mu(\mathbf{r}) \left[\bar{\chi}_n(\mathbf{r}) e^{ia_\mu(\mathbf{r})} \chi_n(\mathbf{r} + \hat{e}_\mu) - \bar{\chi}_n(\mathbf{r} + \hat{e}_\mu) e^{-ia_\mu(\mathbf{r})} \chi_n(\mathbf{r})\right], \tag{3.27}$$

where $\eta_1 = 1$, $\eta_2 = (-1)^x$ and $\eta_3 = (-1)^{x+y}$.

In the case of continuum $QED_3$, the fermion polarization to one-loop order is [53].

$$\Pi_{\mu\nu}(p) = \frac{N_f}{16} p \left\{ \delta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2} \right\}. \tag{3.28}$$

Incorporating compactness of $a_\mu$ in the spirit of the Villain approximation, this suggests that we consider a theory closely related to $cQED_3$

$$S[a] = \frac{1}{2} \sum_{\mathbf{r},\mu,\nu} \left\{ (F_{\mu\nu}(\mathbf{r}) - 2\pi n_{\mu\nu}(\mathbf{r})) \left( \frac{1}{2\hat{q}^2} + \frac{N_f}{16\,|\Delta|} \right) (F_{\mu\nu}(\mathbf{r}) - 2\pi n_{\mu\nu}(\mathbf{r})) \right\}, \tag{3.29}$$

where the $n_{\mu\nu}$ are integers. The action (3.29) has the same continuum limit as $cQED_3$ to the leading order in large $N_f$ and may be understood as a compact quadratic approximation to it. We assume that the original $cQED_3$ and the theory (3.29) are in the same universality class.

In the presence of fermions, when $N_f \neq 0$, the original Maxwell term proportional to $1/\hat{q}^2$ becomes irrelevant at large distances, and can be neglected with respect to the second term in (3.29). Now, in our previous notation, $u = N_f/(16|\Delta|)$ and the potential between monopoles is

$$\begin{aligned} V(|\mathbf{r}_a - \mathbf{r}_b|) &= -\frac{\pi^2 N_f}{2} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{e^{i\mathbf{k}\cdot(\mathbf{r}_a - \mathbf{r}_b)}}{k^3} \\ &\sim \ln(|\mathbf{r}_a - \mathbf{r}_b|) \end{aligned} \tag{3.30}$$

at large distances. Inserting this into the general expression (3.13) yields the *anomalous* sine-Gordon (ASG) action

$$S_{\mathrm{ASG}}[\varphi] = \int \mathrm{d}^3\mathbf{r} \left\{ -\frac{\widetilde{T}}{2} \varphi\, |\nabla|^3\, \varphi - 2y \cos\varphi \right\}, \tag{3.31}$$

where $\widetilde{T} = 2/(N_f \pi^2)$. The non-analytic gradient term proportional to $|q|^3$ is a consequence of the coupling to massless, relativistic fermions.

The most remarkable feature of (3.31) is that the coupling to fermions modifies the potential between monopoles, from Coulombic to the much longer-ranged logarithm in three dimensions. Such a logarithmic interaction is more famously encountered in the 2DXY model[70, 71] where the monopoles (vortices), because of the long-ranged nature of the potential, undergo a BKT transition to a bound dipole phase. In the next section, we review how the competition between the energy and entropy can lead to such a transition. In the present context, this would imply a deconfinement transition for fermions, in contrast to the situation in the pure gauge theory. We contend, however, that screening due to the medium must be taken into account and present an electrostatic argument showing that the bound phase is destabilized in the presence of dipoles.

## 3.6 Stability of the dipole phase

The change in free energy due to a single monopole in a sample of linear size $L$ can be written[74]

$$\Delta F = \frac{1}{4\pi^2} \ln L - \widetilde{T} \ln L^3. \tag{3.32}$$

For $\widetilde{T} < \widetilde{T}_c \equiv 1/(12\pi^2)$, the free energy increases in the presence of free monopoles, and we, thus, expect that the system will favour the formation of bound dipoles. This corresponds to a deconfinement transition at the critical number of fermions, $N_{\mathrm{f}}^{\mathrm{c}} = 24$. This simple argument works surprisingly well in the case of the BKT transition where screening of the medium merely renormalizes the effective temperature. However, the situation in three dimensions is quite different, as we will now demonstrate.

The potential at a distant point $\mathbf{r}$ due to a collection of logarithmically interacting monopoles can be expanded in the spirit of a multipole expansion as

$$V(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \rho(\mathbf{r}') \left\{ (-\ln r) + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^2} + \ldots \right\}, \tag{3.33}$$

from which the various moment densities can be established. In a sea of monopoles and dipoles, the potential of an isolated charge $Q$ will be modified from the bare logarithm due to polarization of the medium. To account for this, the potential can be expressed as

$$V(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \left\{ (-\ln |\mathbf{r} - \mathbf{r}'|) \rho(\mathbf{r}') + \frac{\langle \mathbf{P} \rangle \cdot \mathbf{r}}{|\mathbf{r} - \mathbf{r}'|^2} \right\}, \tag{3.34}$$

where $\langle \mathbf{P} \rangle = n_{\mathrm{d}} \langle \rho(\mathbf{r}') \mathbf{r}' \rangle$ is the thermally averaged dipole moment density ($n_{\mathrm{d}}$ is the dipole number density). For a small dipole of unit charge ($\rho = 1$) in the electric field of $Q$, the energy is $\mathcal{E} = -\mathbf{r}' \cdot \mathbf{E}$, and the thermal average can be evaluated:

$$\begin{aligned} \langle \mathbf{P} \rangle &= \frac{\int \mathrm{d}^3\mathbf{r}' \, (n_{\mathrm{d}} r' \cos\theta) \exp\left\{ \frac{Er' \cos\theta}{T} \right\}}{\int \mathrm{d}^3\mathbf{r}' \, \exp\left\{ \frac{Er' \cos\theta}{T} \right\}} \\ &\approx \frac{n_{\mathrm{d}} \langle r'^2 \rangle}{3T} E, \end{aligned} \tag{3.35}$$

which defines the electric susceptibility $\chi_{\mathrm{e}} \equiv (n_{\mathrm{d}}/3T) \langle r'^2 \rangle$. Writing the electric field as $E = -\nabla V$, the potential, after an integration by parts, is now

$$V(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' (-\ln |\mathbf{r} - \mathbf{r}'|) \left\{ \rho(\mathbf{r}') + \chi_{\mathrm{e}} \nabla^2 V(\mathbf{r}') \right\}. \tag{3.36}$$

In Fourier space, assuming an isolated charge $\rho(\mathbf{r}') = Q\delta(\mathbf{r}')$, we can solve for the screened potential

$$V(\mathbf{q}) = \frac{Q}{|\mathbf{q}|^3 + \chi q^2}. \tag{3.37}$$

Transforming back to real space, the potential felt by a distant external charge no longer has the logarithmic form, but is instead the Coulombic , $V \sim 1/x$. Thus, the presence of a finite density of dipoles destabilizes the bound phase, leading to free monopoles only. In the language of the previous energy-entropy argument, screening modifies the energy of an isolated monopole, which now scales as $1/L$ and, therefore, entropy always dominates at large distances.

To address this issue in a more systematic way, we next study the anomalous sine-Gordon action (3.31) using the GBF variational approximation employed in the pure gauge theory. To the lowest order, we will see that this approach agrees with the naïve energy-entropy argument, owing to the neglect of screening. We will then introduce a systematic generalization of this method meant to capture such non-trivial effects, and show that a finite mass solution does indeed exists at all temperatures, in agreement with our electrostatic argument.

## 3.7 Variational approach

Again, we choose the trial action to have the Gaussian form

$$S_0[\varphi] = \frac{1}{V}\sum_{\mathbf{q}} \frac{1}{2}\varphi(\mathbf{q})G_0^{-1}(\mathbf{q})\varphi(-\mathbf{q}), \tag{3.38}$$

and calculate $F_{\rm var}$:

$$\begin{aligned} \frac{F_{\rm var}}{V} &= -\frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\ln\left(G_0(\mathbf{q})\right) + \frac{\widetilde{T}}{2}\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}|q|^3 G_0(\mathbf{q}) \\ &\quad -2y\,\exp\left\{-\frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}G_0(\mathbf{q})\right\}. \end{aligned} \tag{3.39}$$

Minimizing $F_{\rm var}$ with respect to $G_0(q)$ yields the optimal Gaussian theory that approximates $F_{\rm ASG}$:

$$\frac{\delta F_{\rm var}}{\delta G_0} = 0 \Longrightarrow G_0^{-1}(\mathbf{q}) = \widetilde{T}|q|^3 + m, \tag{3.40}$$

with the 'mass' $m$ determined self-consistently through

$$\begin{aligned} m &= 2y\,\exp\left\{-\frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\frac{1}{\widetilde{T}|q|^3 + m}\right\} \\ &= 2y\left(1 + \frac{\widetilde{T}\Lambda^3}{m}\right)^{-\frac{\widetilde{T}_c}{\widetilde{T}}}. \end{aligned} \tag{3.41}$$

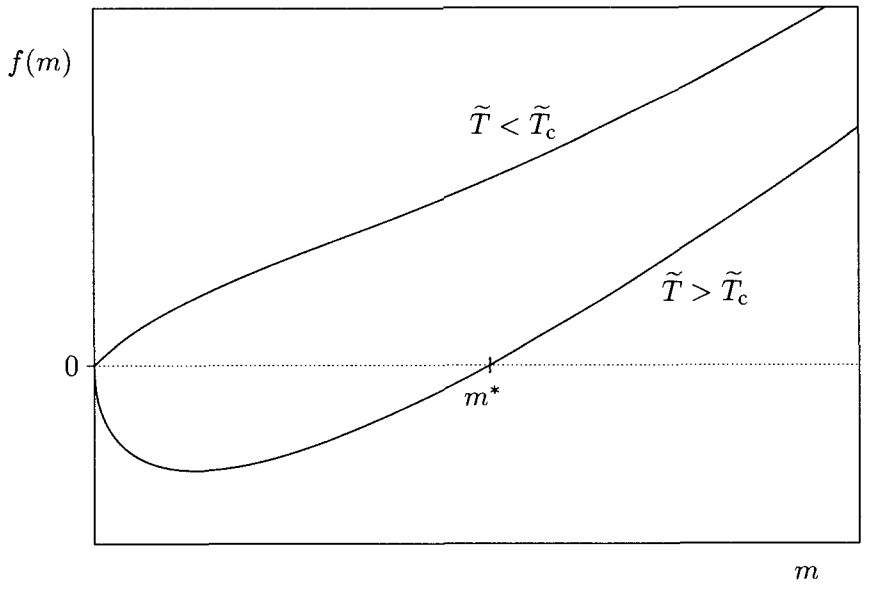


Figure 3.3: The solid curves are the function $f(m)$ for $\widetilde{T} > \widetilde{T}_c$ and $\widetilde{T} < \widetilde{T}_c$, as indicated. The solution $m^*$ is determined by the intersection of the solid lines with the $f(m) = 0$ axis. A finite solution exists only for $\widetilde{T} > \widetilde{T}_c$.

$\Lambda$ is the ultraviolet cutoff and $\widetilde{T}_c \equiv 1/(12\pi^2)$. Determining the solutions of (3.41) amounts to identifying the roots of the function

$$f(m) = m - 2y\left(1 + \frac{\widetilde{T}\Lambda^3}{m}\right)^{-\frac{\widetilde{T}_c}{\widetilde{T}}}. \tag{3.42}$$

It is evident that $m = 0$ is one such root for all values of $\widetilde{T}$. We next demonstrate that a solution with finite $m^*$ exists for $\widetilde{T} > \widetilde{T}_c$. In the limit of small $m$, $f(m)$ has the form

$$f(m \ll \Lambda^3) = \begin{cases} m, & \widetilde{T} < \widetilde{T}_c \\ -m^{\widetilde{T}_c/\widetilde{T}}, & \widetilde{T} > \widetilde{T}_c \end{cases}, \tag{3.43}$$

while for large $m$

$$f(m \gg \Lambda^3) = m, \quad \forall\, \widetilde{T}. \tag{3.44}$$

For $\widetilde{T} > \widetilde{T}_c$, $f(m)$ changes sign and, thus, has a root with $m^* > 0$, while only the $m = 0$ solution exists for $\widetilde{T} < \widetilde{T}_c$ [82]. This situation is depicted graphically in Figure 3.3.

The stability of the $m = 0$ solution for $\widetilde{T} > \widetilde{T}_c$ can be determined from the variational free energy (3.39) with the solution (3.40) for $G_0^{-1}$. Evaluating the free energy we find

$$\frac{F_{\rm var}(m)}{V} = \widetilde{T}_c \Lambda^3 \ln\left(m + \widetilde{T}\Lambda^3\right) - 2y\left(1 + \frac{\widetilde{T}\Lambda^3}{m}\right)^{-\widetilde{T}_c/\widetilde{T}}. \tag{3.45}$$

Then

$$\frac{1}{V}\left(F_{\rm var}(m^*) - F_{\rm var}(0)\right) = \frac{m^*(\widetilde{T}_c - \widetilde{T})}{\widetilde{T}} + O(m^2), \tag{3.46}$$

so that for $\widetilde{T} > \widetilde{T}_c$ any solution with $m^* > 0$ is of lower free energy than with $m = 0$. That is, the stable solution at $\widetilde{T} > \widetilde{T}_c$ has finite $m^*$.

Again, we identify the finite $m^*$ solution with the phase of free monopoles, while $m = 0$ indicates the dipole phase. The simple variational calculation would, therefore, suggest that monopoles undergo a binding-unbinding transition at $\widetilde{T} = \widetilde{T}_c$ (i. e. at $N = N_f^c = 24$) in exact analogy with the equivalent calculation one can perform for the standard BKT transition. The value of $\widetilde{T}_c$ also agrees with the simple energy-entropy argument which we presented in the last section.

An obvious objection to this simple calculation is that minimization of the variational free energy (3.39) by construction cannot yield any momentum dependence of the self-energy, but can only determine its constant part, the 'mass' $m$. The renormalization group [74] treatment of the ASG theory suffers from the same problem to the lowest order in fugacity, and would likewise naïvely suggest the BKT transition. The same holds for the direct perturbative evaluation of the self-energy in the ASG. However, it is easy to check that the self-energy does become momentum dependent to the *second* order in fugacity, with the leading analytic term $\sim q^2$ at low momenta. This is just what one would expect based on the simple electrostatic analysis of the problem where this term translates into the Coulombic interaction in real space. The presence of such a term would, however, drastically alter our present considerations. Indeed, if we add by hand the term $Qq^2$ with $Q \neq 0$ in the denominator of the integrand in the self-consistent equation (3.41), we find

$$f(m) = \begin{cases} -2y\left(1 + \frac{\Lambda\widetilde{T}}{\chi}\right)^{-3\widetilde{T}_c/\widetilde{T}}, & \sigma \ll \Lambda^3 \\ m, \quad m \gg \Lambda^3 \end{cases} \tag{3.47}$$

for all $\widetilde{T}$. Hence, the non-trivial solution would exist for all temperatures, exactly as in Polyakov's original treatment of the pure gauge theory. This is natural since $Q \neq 0$ means that the original logarithmic interaction between monopoles is, even without free monopoles and only with a finite

density of dipoles, screened into the Coulomb interaction for which the standard argument for the confined phase readily applies.

In the next section we propose a modified self-consistent calculation which provides a systematic perturbative approximation to the free energy and which reduces to the GBF method to the lowest order. As we will see in Section 3.9, such an approach has the advantage of including the screening effects in a self-consistent way, therefore overcoming the limitations of the purely variational theory discussed in this section.

## 3.8 Self-consistent perturbative approach

There are many ways in which one may generalize the variational method of the previous section. For instance, one may add a second-order term $-\frac{1}{2}\langle(S_{\rm ASG}-S_0)^2\rangle_0+\frac{1}{2}\langle S_{\rm ASG}-S_0\rangle_0^2$ to $F_{\rm var}$ and extremize the new energy functional. Such a second-order extension, however, has little variational justification. For a more systematic generalization, we go back to the GBF inequality (3.15). Exchanging the roles of $S_{\rm ASG}$ and $S_0$, we find

$$F_< \equiv F_0+\langle S_{\rm ASG}-S_0\rangle \leq F_{\rm ASG}. \tag{3.48}$$

Extremizing $F_<$ with respect to a quadratic action $S_0$ yields

$$\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0=\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle, \tag{3.49}$$

which is simply the equation for the *exact* propagator in the ASG theory. The right hand side (RHS) of the equation may be rewritten as

$$\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle\equiv\frac{\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})e^{-\Delta S}\rangle_0}{\langle e^{-\Delta S}\rangle_0}, \tag{3.50}$$

with $\Delta S\equiv S_{\rm ASG}-S_0$. In this form, (3.49) may be understood as a self-consistent equation for the action $S_0$, which we may attempt to solve by expanding the RHS in powers of $\Delta S$, for example. To the first order in $\Delta S$ this becomes

$$\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0-\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\langle\Delta S\rangle_0=0, \tag{3.51}$$

which is precisely the relation one would obtain from extremizing $F_{\rm var}$ with respect to $S_0$ in (3.15). That is, the first order approximation to (3.49) reproduces the GBF result from the previous section.

The equation (3.49) forms the basis of our modified variational approximation to $F_{\rm ASG}$. To the first order in $\Delta S$ it reduces to the GBF equation of the previous section, and when solved self-consistently to all orders, gives the best variational lower bound to the free energy, provided by $F_<$ in (3.48).

In addition, consider the expansion of (3.49) to order $(\Delta S)^n$. One can show (see Appendix A.2) that the resulting expression is the same as the one that would arise from extremizing the function

$$F_{\rm var}^{(n)} \equiv \frac{F^{(1)} + F^{(2)} + \cdots + F^{(n)}}{n}. \tag{3.52}$$

Here $F^{(n)}$ stands for the expansion of the true free energy of the system, $F_{\rm ASG}$, in powers of $\Delta S$, truncated at $(\Delta S)^n$. As an illustration, consider the $n = 2$ case. From $F^{(1)} = F_0 + \Delta S$ and $F^{(2)} = F^{(1)} - \frac{1}{2}\left(\langle(\Delta S)^2\rangle_0 - \langle\Delta S\rangle_0^2\right)$, we can compute the functional derivatives

$$\frac{\delta F^{(1)}}{\delta G_0(\mathbf{q})} = \left(\frac{-1}{2(2\pi)^3 G_0(\mathbf{q})}\right)\Big\{\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\langle\Delta S\rangle_0 - \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0\Big\}, \tag{3.53}$$

$$\begin{aligned}\frac{\delta F^{(2)}}{\delta G_0(\mathbf{q})} &= \left(\frac{-1}{2(2\pi)^3 G_0(\mathbf{q})}\right)\Big\{\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\left[-\frac{1}{2}\langle\Delta S^2\rangle_0 + \langle\Delta S\rangle_0^2\right] \\ &+ \frac{1}{2}\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})(\Delta S)^2\rangle_0 - \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0\langle\Delta S\rangle_0\Big\}.\end{aligned} \tag{3.54}$$

In the above, we have used

$$\frac{\delta F_0}{\delta G_0(\mathbf{q})} = \left\langle\frac{\delta S_0}{\delta G_0(\mathbf{q})}\right\rangle_0 = \frac{-1}{2(2\pi)^3[G_0(\mathbf{q})]^2}\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0, \tag{3.55}$$

$$\frac{\delta\langle g\rangle_0}{\delta G_0(\mathbf{q})} = \frac{\delta F_0}{\delta G_0(\mathbf{q})}\langle g\rangle_0 + \left\langle\frac{\delta g}{\delta G_0(\mathbf{q})} - g\frac{\delta S_0}{\delta G_0(\mathbf{q})}\right\rangle_0. \tag{3.56}$$

Now, adding the derivatives together, and setting the result to zero, we find

$$\frac{\delta F_{\rm var}^{(2)}}{\delta G_0} = \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0^{\rm c} - \frac{1}{2}\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})(\Delta S)^2\rangle_0^{\rm c} = 0, \tag{3.57}$$

where we have defined the connected averages

$$\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0^{\rm c} \equiv \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0 - \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\langle\Delta S\rangle_0, \tag{3.58}$$

$$\begin{aligned}\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})(\Delta S)^2\rangle_0^{\rm c} &\equiv \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})(\Delta S)^2\rangle_0 - \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\langle(\Delta S)^2\rangle_0 \\ &- 2\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\Delta S\rangle_0\langle\Delta S\rangle_0 + 2\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0\langle\Delta S\rangle_0^2.\end{aligned} \tag{3.59}$$

This is exactly the result which follows from expanding (3.49) to second order.

The utility of identifying the appropriate variational free energy to be minimized is evident when it is cast into the alternate form (see Appendix A.1)

$$F_{\rm var}^{(n)} = F^{(n)} + \frac{F^{(n)} - F_{<}^{(n)}}{n}. \tag{3.60}$$

It is then clear that the sequence $\{F_{\rm var}^{(n)}\}$ converges to $F$ for any $S_0$. Therefore, the $S_0$ determined self-consistently from (3.49) yields the variational sequence that best approximates $F_{\rm ASG}$ within the family $\{F_{\rm var}^{(n)}[S_0]\}$.

Having established the validity of our systematic extension to the GBF inequality, we will apply this to the anomalous sine-Gordon model (3.31) using the second order result (3.57) derived above. In particular, we will show that the density of free monopoles is finite at all $T > 0$, and that charge should consequently be permanently confined in $\mathrm{cQED}_3$.

## 3.9 Confining Solution for $T > 0$

From the definitions of $S_{\rm ASG}$ and $S_0$ (3.31 and 3.38) it is straightforward to calculate the connected averages of (3.58) and (3.59). Our second order result (3.57) then yields the quadratic equation for $G_0^{-1}(\mathbf{q})$

$$\left[G_0^{-1}(\mathbf{q})\right]^2 - A[\mathbf{q}, G_0]G_0^{-1}(\mathbf{q}) + B[\mathbf{k}, G_0] = 0, \tag{3.61}$$

where

$$A[\mathbf{q}, G_0] = \frac{3}{2}\widetilde{T}|q|^3 + 3a + ab - 2a^2\left(c + \sum_{n=0}^{\infty}(-1)^n d_n q^{2n}\right) \tag{3.62}$$

$$B[\mathbf{q}, G_0] = \frac{1}{2}\widetilde{T}^2 q^6 + 2a\widetilde{T}|q|^3. \tag{3.63}$$

In the above expressions (3.62, 3.63), we have defined

$$a = y e^{-\frac{1}{2}D_0(0)}, \tag{3.64}$$

$$b = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\left(\frac{\widetilde{T}}{2}|k|^3 - \frac{1}{2}G_0^{-1}(\mathbf{k})\right)\left[G_0(\mathbf{k})\right]^2, \tag{3.65}$$

$$c = \int \mathrm{d}^3\mathbf{R}\left[1 - \cosh D_0(\mathbf{R})\right], \tag{3.66}$$

$$d_n = \int \mathrm{d}^3\mathbf{R}\frac{(R\cos\theta)^{2n}}{(2n)!}\sinh D_0(\mathbf{R}), \tag{3.67}$$

and the real-space propagator is $D_0(\mathbf{R}) = \int \mathrm{d}^3\mathbf{k}/(2\pi)^3 G_0(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{R}}$.

We can solve the quadratic of (3.61) and expand in powers of $|q|^3/A_0$ to yield the result for $G_0^{-1}(\mathbf{q})$:

$$G_0^{-1}(\mathbf{q}) = m + Q(m)q^2 + \mathcal{T}(m)|q|^3 + \cdots, \tag{3.68}$$

where the coefficients are defined as

$$m = \frac{1}{2}\{A_0 \pm |A_0|\}, \tag{3.69}$$

$$Q(m) = a^2 d_1 \left(1 \pm \frac{|A_0|}{A_0}\right), \tag{3.70}$$

$$\mathcal{T}(m) = \frac{3}{4}\widetilde{T} \pm \frac{|A_0|}{A_0}\left(\frac{3}{4}\widetilde{T} - \frac{2a\widetilde{T}}{A_0}\right); \tag{3.71}$$

and with $A_0 \equiv A[q = 0, G_0]$. For these equations, we should choose the solution corresponding to the upper sign in (3.69 – 3.71) to ensure that $m \geq 0$. In what follows, we neglect terms higher order in $q$ than $q^3$ as they should be irrelevant at low momenta.

As announced, the second order result includes additional renormalization of the bare terms as well as the generation of new momentum dependent terms. Most importantly, the leading term proportional to $q^2$ has now appeared.

In the analysis to the lowest order, we found that the bound phase of monopoles corresponded to low $\widetilde{T}$. In what follows we will restrict ourselves to low temperatures by assuming $\widetilde{T}\Lambda \ll Q$ and show that monopoles are unbound even for arbitrarily small temperatures. By continuity this would imply that they are free at all temperatures.

Let us start by examining $a$:

$$\begin{aligned} a &= y \exp\left\{-\frac{1}{2}D_0(0)\right\} \\ &\approx y \exp\left\{-\frac{1}{2}\int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\,\frac{1}{(Q(m)k^2 + m)}\right\} \\ &= y \exp\left\{-\frac{1}{4\pi^2 Q(m)}\left(\Lambda - \sqrt{\frac{m}{Q(m)}}\arctan\left(\Lambda\sqrt{\frac{Q(m)}{m}}.\right)\right)\right\}, \end{aligned} \tag{3.72}$$

When $m \to 0$, we will assume $m/Q(m) \to 0$, and justify this assumption *a posteriori*. The coefficient $a$ now takes the form

$$a = y \exp\left\{-\frac{\Lambda}{4\pi^2 Q(m)}\right\} + \mathcal{O}(\widetilde{T}), \quad m \ll \Lambda^3. \tag{3.73}$$

Next, we examine the equation for $b$

$$\begin{aligned} b &= \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\left(\frac{\widetilde{T}}{2}|k|^3 - \frac{1}{2}G_0^{-1}(\mathbf{k})\right)[G_0(\mathbf{k})]^2 \\ &= -\frac{1}{2}D_0(0) + \mathcal{O}(\widetilde{T}). \end{aligned} \tag{3.74}$$

From this we find

$$b = -\frac{\Lambda}{4\pi^2 Q(m)} + \mathcal{O}(\widetilde{T}), \quad m \ll \Lambda^3. \tag{3.75}$$

Next, as the terms $c$ and $d_0$ always appear together, we consider the combination

$$\begin{aligned} (c+d_0) &\approx \int \mathrm{d}^3\mathbf{R}\left(1 - \exp\left\{-\int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\,\frac{e^{i\mathbf{k}\cdot\mathbf{R}}}{(Q(m)k^2+m)}\right\}\right) + \mathcal{O}(\widetilde{T}) \\ &= \int_0^\infty \mathrm{d}R\,\frac{R\mathrm{e}^{-\sqrt{\frac{m}{Q(m)}}R}}{Q(m)} + \mathcal{O}(\widetilde{T}). \end{aligned} \tag{3.76}$$

Evaluating this yields

$$(c+d_0) = m^{-1} + \mathcal{O}(\widetilde{T}), \quad m \ll \Lambda^3. \tag{3.77}$$

Similar analysis applies to the coefficient $d_1$:

$$d_1 = \frac{1}{6}\int_0^\infty \mathrm{d}R\,\frac{R^3\mathrm{e}^{-\sqrt{\frac{m}{Q(m)}}R}}{Q(m)} + \mathcal{O}(\widetilde{T}), \tag{3.78}$$

which gives

$$d_1 = \frac{Q(m)}{m^2} + \mathcal{O}(\widetilde{T}), \quad m \ll \Lambda^3. \tag{3.79}$$

Evaluating (3.70) for $Q$ then we find

$$\begin{aligned} Q &= 2a^2 d_1 \\ &= 2y^2 \exp\left\{-\frac{\Lambda}{2\pi^2 Q}\right\}\frac{Q}{m^2} + \mathcal{O}(\widetilde{T}). \end{aligned} \tag{3.80}$$

Solving this for $Q \neq 0$ yields

$$Q = \frac{\Lambda}{4\pi^2}\left(\ln\frac{\sqrt{2}y}{m}\right)^{-1} + \mathcal{O}(\widetilde{T}), \tag{3.81}$$

and we see that $m/Q(m)$ indeed approaches zero as $m \to 0$, thus justifying our earlier assumption. Substituting this solution for $Q(m)$ into our mass equation (3.69) gives

$$\begin{aligned} m &= A_0 \\ &\approx \frac{m}{\sqrt{2}}\left[3 - \sqrt{2} - \ln\frac{\sqrt{2}y}{m}\right], \end{aligned} \tag{3.82}$$

which can finally be solved for $m \neq 0$ to give the finite mass solution

$$m^* = \sqrt{2}\mathrm{e}^{2\sqrt{2}-3}y. \tag{3.83}$$

The corresponding finite value of $Q$ is

$$Q^* = \frac{\Lambda}{2\pi^2(3-2\sqrt{2})}. \tag{3.84}$$

Note that $m^*$ is proportional to $y$ so that small fugacity translates to small $m^*$, in accord with our assumption that $m \ll \Lambda^3$.

To show that monopoles are free when $m \neq 0$, we calculate the monopole density as before (3.23). From (3.52) we see that the free energy associated with our second order equation (3.57) is

$$F_{\mathrm{var}}^{(2)} = F_0 + \langle \Delta S \rangle_0 - \frac{1}{4}\langle (\Delta S)^2 \rangle_0 + \frac{1}{4}\langle \Delta S \rangle_0^2. \tag{3.85}$$

From this, the monopole density can be calculated.

$$\begin{aligned} \rho_M^{(2)} &= -\frac{1}{V}\frac{\partial F_{\mathrm{var}}^{(2)}}{\partial \mu} \\ &= 2a + ab - 2a^2c. \end{aligned} \tag{3.86}$$

For $m = 0$, the monopole density vanishes, while for our finite $m$ solution

$$\begin{aligned} \rho_M^{(2)} &= \frac{m^*}{\sqrt{2}}\left(2\sqrt{2} - 1 + \frac{1}{16\pi}\sqrt{\frac{2m^*}{(Q^*)^3}}\right) \\ &> 0. \end{aligned} \tag{3.87}$$

From the free energy (3.85), it is also possible to show that the finite $m$ solution is the stable solution for all temperatures. In fact, the free energy diverges as $\log(1/m)$ as $m$ approaches zero, but has a finite value for finite $m$. It is then the free phase of monopoles which is favoured at all temperatures.

Thus we have demonstrated that for arbitrarily low $\widetilde{T}$ a finite mass solution always exists for the self-consistent equations (3.69 – 3.71). This implies that monopoles are always free at low temperatures, or, in terms of the original lattice model (3.26), that the electric charge is presumably confined for any number of fermion flavours.

## 3.10 Conclusions

We have studied cQED$_3$ where massless relativistic fermions coupled to the compact gauge field result in a logarithmic interaction between magnetic monopoles. One may suspect that this could

lead to a BKT-like transition where free monopoles bind into monopole-antimonopole pairs at low enough effective temperatures. Although the simplest mean-field approximation would predict such a transition, we argued that by design this treatment misses the screening effects, which are crucial in this problem. To address this issue we developed a combined variational-perturbative approach which allowed us to include screening self-consistently. The modified theory then leads to the plasma phase of free monopoles as being stable at all temperatures, in agreement with the renormalization group treatment of the problem [74]. Our conclusion is also in agreement with a numerical study of $cQED_3$[83].

This result has implications for the phase of fermions at finite coupling. We've already seen that a chiral symmetry breaking transition occurs for $\hat{q} = 0$, which corresponds to the continuum limit of the theory (3.26). For $\hat{q} > 0$, we effectively integrated out the fermions, and concentrated on the phase of the remaining monopole degrees of freedom. Now imagine having proceeded, instead, by integrating out the gauge field to determine the phase diagram of the theory in terms of fermions. It follows that we would find again a single phase, one of either preserved or broken chiral symmetry. In a numerical study comparing compact and non-compact $QED_3$, Fiebig and Woloshyn [81] found that free monopoles *enhance* chiral symmetry breaking; that is, the chiral condensate $\langle \bar{\Psi}\Psi \rangle$ is an increasing function of monopole density. This suggests that we must find fermions in the chiral symmetry broken phase for finite $\hat{q}$. This is summarized in the proposed phase diagram shown as Figure 3.4. The line $N_f^?$ suggested in Figure 3.2 does not mark a phase boundary, and so we omit it here, although the renormalization group treatment of the problem[74] does indicate a crossover behaviour occurring as $N_f$ is increased. In that work, the flow of the fugacity monotonically increases for $N_f$ below this line, which is the sign of free monopoles. However, above this line, the fugacity first decreases before ultimately running away, which suggests that fermions might actually appear deconfined at not too large distances.

The fermions in our picture, of course, represent the neutral low-energy spinon excitations of an underdoped cuprate superconductor. In the superconducting phase, the gauge fields are massive and so the theory, which is not QED, does not confine spinons. In this sense, there is an effective fractionalization of the electrons into its constituent spin and charge parts, as has been pointed out before. Upon exiting the superconducting state, these spinons become confined into (neutral) spin one objects and, thus, spin-charge separation does not persist into the spin density wave state, since there are no longer any uncharged, half-spin excitations in the particle spectrum.

A fractionalized, free spinon state corresponding to deconfinement is known as a spin liquid,

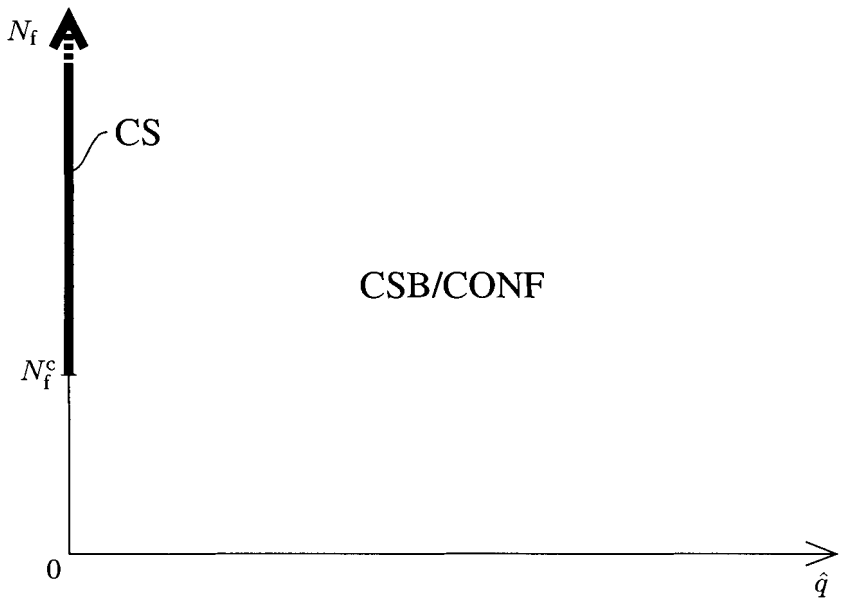


Figure 3.4: The actual phases of cQED$_3$. The transition in the continuum ($\hat{q} = 0$) is between chirally symmetric (CS) and chiral symmetry broken (CSB) phases; pure gauge cQED$_3$ ($N_f = 0$) is permanently confining for any value of the coupling. We have shown that permanent confinement persists in the presence of massless fermions and with $\hat{q} \neq 0$. In terms of the chiral symmetry, we expect that the single phase is the chiral symmetry broken phase.

and it has been shown[75] that such states are generically unstable in two dimensions for compact U(1) gauge theories. The present result is a specific example of this. It has recently been argued[84], though, that monopoles are, in fact, irrelevant and that deconfinement can occur, stablizing the spin liquid state. This argument suggests that our starting point (3.29) is flawed, and that we should ignore the compactness of $a$ arising from fermions. In the absence of monopoles, the theory is similar to the case in the continuum limit and a deconfinement transition may indeed be possible. However, the validity of this claim of the irrelevance of monopoles is unclear, and so the issue remains unresolved.

In contrast to the compact case, non-compact QED$_3$ is believed to undergo a transition[85] for finite coupling $\hat{q}$. Numerical results also seem to indicate that the critical number of fermion flavours increases as the coupling increases[85], as in Figure 3.2, though it is an open question whether the transition extends to infinite $\hat{q}$ or not. As before, the chirally symmetric phase corresponds to deconfinement, and so whether the critical number of flavours is above or below two determines

whether spinons are confined or not.

# Chapter 4

# Uemura scaling

In this Chapter, we will begin our discussion on the application of the theory (2.52) to the superfluid density in the underdoped cuprates. This research can be found in references [42] and [43], the former being the subject of the current Chapter, while the latter will be covered in the next.

## 4.1 Introduction

The superfluid density and its temperature dependence are fundamental properties of superfluids and superconductors. Both the zero temperature value and the low temperature behaviour are indicators of key properties involved in the condensation process. Perhaps the earliest experiments were performed by Andronikashvili[86] using a torsional pendulum suspended in $^4$He below the $\lambda$-transition. The period of oscillation of the pendulum is directly related to the *normal* fluid density, and these experiments provided the earliest justification of Landau's two fluid model of superfluidity[87].

In superconductors, the superfluid density is related to the depth of penetration of an applied magnetic field into the interior of the sample[6]. For conventional materials, or whenever the BCS theory applies, the temperature dependence is entirely due to the form of the superconducting gap function and, thus, may serve as an indicator of the pairing symmetry. For superconductors with isotropic $s$-wave symmetry, the full gap induces an exponentially activated form for the low-temperature behaviour[16], while the high-temperature superconductors display a linear in temperature behaviour[88], characteristic of an order parameter with nodes[89]. Indeed, this observation provided the earliest compelling evidence for $d$-wave symmetry in the cuprates[6].

The behaviour near the critical temperature is also notable, since it reflects the importance of

critical fluctuations. In optimally doped samples, a wide critical region has been observed where $\rho_s(T) \sim (T_c - T)^\nu$, with the critical exponent $\nu \approx 0.67$[32], indicative of 3DXY criticality[31].

These regimes of high and low temperature are intimately related by what has become known as Uemura scaling. Based on early muon spin relaxation relaxation measurements[23, 24], the zero temperature value of the superfluid density was seen to depend linearly on the critical temperature. This observation was subsequently seen in various other experiments and became adopted as a fact, and a necessary prediction for any candidate theory of high temperature superconductivity. Very recently, however, new experiments have begun to display marked deviations from this linear relation[11, 90], as sample preparation techniques have improved and experimental methods grown more reliable. The consensus which is beginning to grow is that $\rho_s(0)$ depends sublinearly on $T_c$, though with what power is still not well established[8].

To develop a theoretical understanding of the superfluid density and, thus, hopefully an understanding of the relation between $\rho_s(0)$ and $T_c$, we need to isolate the factors which determine the temperature dependence. In principle, of course, there are many things to consider; for example, we have already mentioned the superconducting gap, or quasiparticle contribution. Another important contribution comes from fluctuations around the 'saddle-point' value of the gap – amplitude and phase fluctuations. As we have already argued, though, the only relevant fluctuations at low dopings are in the phase, and so we will focus on those only. The energy required to induce such fluctuations is measured by the phase stiffness, which is related to the superfluid density and which, therefore, decreases with underdoping[88] (as also implied by the Uemura scaling). This simple observation led to the influential arguments of Emery and Kivelson[26], which assert that it is the phase fluctuations which are solely responsible for the depletion of the superfluid density at low dopings.

Besides the obvious logic of the above claim, it also readily explains the dome-like shape of the high temperature superconductors' phase diagram. In the BCS mean-field theory, the transition temperature can be determined from the gap equation, and steadily increases with underdoping[47], in contradiction with the observed critical temperature. On the other hand, the critical temperature in the absence of any disordering effect *except* phase fluctuations, is an increasing function of phase stiffness and, thus, of doping[26]. This last is an entirely empirical statement, and in particular does not explain why the phase stiffness is proportional to doping, as observed. In any case, the actual transition temperature is then determined by the lower of the two temperatures, implying the dome-like shape as observed.

This argument provides a definite framework in which to theoretically determine the behaviour of the superfluid density and, thus, to assess the validity of its assumptions. The goal of the current Chapter is to determine the superfluid density in the XY model of phase fluctuations, and to see what this predicts for the universal relation between $\rho_s(0)$ and the critical temperature. We will find that phase fluctuations alone are insufficient to account for the magnitude of the Uemura relation as seen in the old experiments, and that the contribution due to quasiparticles must be included to make the theory consistent with observations. Furthermore, because of a strong doping dependence of the quasiparticle contribution, a deviation from the linear relation is predicted, in agreement with the most recent experimental data.

## 4.2 Definitions of the superfluid density

The superfluid density is a quantity which provides a response to the imposition of a helical twist of the order parameter[91],

$$\psi_0 = \langle \psi(\mathbf{r}) \rangle \rightarrow \psi_0 e^{i\mathbf{k}_0 \cdot \mathbf{r}}, \tag{4.1}$$

with a wave number $k_0 = \theta/L$. It is clear that the effect of such a twist is to simply shift $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{k}_0$ in the kinetic energy term of the Hamiltonian, the result being

$$K = \frac{\hbar^2}{2m} \int \mathrm{d}^d\mathbf{r} |(\nabla + i\mathbf{k}_0)\psi(\mathbf{r})|^2. \tag{4.2}$$

We can see then, by expanding the square, that the cross-terms are the equivalent of a Galilean transformation to a reference frame moving with velocity $\mathbf{v} = \hbar\mathbf{k}_0/m$, i. e. $\epsilon(k) \rightarrow \epsilon(k) - \hbar\mathbf{k} \cdot \mathbf{v}$. The consequence of such a transformation is a net momentum flux, due to the flow of quasiparticles, given by

$$\left\langle \hat{P} \right\rangle = V \rho_n \mathbf{v}. \tag{4.3}$$

This last expression then defines the *normal* fluid (mass) density, that of the non-superfluid quasiparticle excitations. To find the *super*fluid density, we simple subtract this from the total density, $\rho_s = \rho - \rho_n$. This argument is completely equivalent to that of Landau, and in particular reproduces the famous result for interacting Bose systems, that $\rho_n \sim T^4$ in three dimensions[87].

Relating to the above definition, one could also consider the change in free energy due to a flowing superfluid. The velocity defined above could equally be considered a superfluid velocity (in the laboratory frame), and the associated increase in free energy from the non-flowing case can be

written[51]

$$\Delta\mathcal{F} = \frac{1}{2}\rho_s \mathbf{v}^2 V. \tag{4.4}$$

This definition, however, does not lend itself to ready calculation, as did the first. Another drawback, shared also by the previous definition, is the reliance on translational invariance. The total momentum is not well defined for lattice models, and so the arguments presented above are not valid in such circumstances.

A more general and rigourous definition, which avoids the complications arising from translational non-invariance, has been given in terms of the *helicity modulus* [91], which measures the free energy difference between systems with different boundary conditions. The idea is to impose a twist in the order parameter by fixing the phase at opposite ends of the system to differ by some angle $\theta$. The change in free energy is proportional to the square of the angle, the constant of proportionality being the helicity modulus $\Upsilon$:

$$\Delta\mathcal{F} \approx \frac{1}{2}\Upsilon(T)\theta^2 A/L; \tag{4.5}$$

$A$ is the cross-sectional area and $L$ is the length. The approximation becomes exact when the pitch of the twist goes to zero, which in this instance must be achieved by taking the thermodynamic limit, i. e.

$$\Upsilon(T) = \lim_{A,L\to\infty} \left(\frac{2L}{\theta^2 A}\right) \Delta\mathcal{F}. \tag{4.6}$$

Comparing (4.5) with our previous definition (4.4), with $k_0 = \theta/L$, we immediately recognize

$$\rho_s = \left(\frac{m}{\hbar}\right)^2 \Upsilon. \tag{4.7}$$

This now allows for the direct calculation of $\rho_s$ by evaluating the partition function under specified boundary conditions. Such a calculation, however, is not necessarily practical[92]. In particular, the free energy of the finite system must be determined, which is difficult. The final definition we will introduce makes use of the free energy in the thermodynamic limit, and should be completely equivalent to the preceding definitions.

In the definition (4.6), a fixed twist angle was applied along a finite length $L$, and the thermodynamic limit was imposed to take the pitch to zero. We could equally apply a twist with wave number $k_0$ to the infinite sample[93], and then the appropriate limit would be $k_0 \to 0$. The free energy would now have a term proportional to $k_0^2$, analogously to (4.5), suggesting a definition of the helicity modulus

$$\Upsilon(T) = \frac{1}{V}\left(\frac{\partial^2\mathcal{F}(k_0)}{\partial k_0^2}\right)_{k_0=0}. \tag{4.8}$$

That this is equivalent to the above definitions has been shown for several simple models[94], and we will assume it to be valid always. In particular, we will use it in the next section, when we establish the relationship between the superfluid density and the magnetic field penetration depth in superconductors.

In anisotropic materials like the high temperature superconductors, it is necessary to differentiate between the superfluid densities along the different crystallographic directions. The definition (4.8) can still be used, by substituting $k_0$ with $k_{0,i}$, where $i$ represents the direction along which $\rho_s$ is to be determined. We will primarily be concerned with the superfluid density in the CuO planes, $\rho_{ab}$.

## 4.3 Superfluid density and the penetration depth

The most conspicuous feature of superconducting materials is zero resistance to the motion of charge carriers[95, 96]. From Ohm's law[97], $\mathbf{j} = \sigma\mathbf{E}$ where $\mathbf{j}$ is the current density, $\sigma$ is the conductivity and $\mathbf{E}$ is the electric field, infinite conductivity, therefore, implies that the field is zero. When applied to the Maxwell equation $-c\nabla \times \mathbf{E} = \partial\mathbf{B}/\partial t$, this implies that the magnetic field in the interior of a superconductor is constant. However, what was observed experimentally[95] was the complete expulsion of flux from the superconductor, now known as the Meissner-Oschenfeld effect.

To explain this situation, the Londons[98] developed a phenomenological theory based upon Maxwell's electrodynamics. For a density $n_s$ of superconducting electrons, the current density is $\mathbf{j}_s = -en_s\mathbf{v}$, which can be used to write the Lorentz force law as

$$\frac{\partial \mathbf{j}_s}{\partial t} - \frac{1}{2en_s}\nabla j_s^2 + \frac{1}{en_s}\mathbf{j}_s \times (\nabla \times \mathbf{j}_s) = \frac{e^2 n_s}{m}\mathbf{E} - \frac{e}{mc}\mathbf{j}_s \times \mathbf{B}, \tag{4.9}$$

In the presence of a time-dependent magnetic field, Faraday's law of induction states that

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{4.10}$$

so that, by taking the curl of (4.9), we are led to the relation between the current density and magnetic field

$$\frac{\partial}{\partial t}\left(\nabla \times \mathbf{j}_s + \frac{e^2 n_s}{mc}\mathbf{B}\right) + \frac{1}{en_s}\nabla \times \left[\mathbf{j}_s \times \left(\nabla \times \mathbf{j}_s + \frac{e^2 n_s}{mc}\mathbf{B}\right)\right] = 0. \tag{4.11}$$

In light of the Meissner effect, the Londons proposed that the correct solution for superconductors is

$$\left(\nabla \times \mathbf{j}_s + \frac{e^2 n_s}{mc}\mathbf{B}\right) = 0, \tag{4.12}$$

so that the observed behaviour can be recovered. To see this, apply the Maxwell equation

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j}_{\mathrm{s}}, \tag{4.13}$$

so that the *London equation* (4.12) reduces to

$$\nabla^2 \mathbf{B} = \frac{4\pi e^2 n_{\mathrm{s}}}{mc^2}\mathbf{B}, \tag{4.14}$$

where we have also used the Maxwell equation $\nabla \cdot \mathbf{B} = 0$. In one dimension, we easily find that

$$\mathbf{B} = \mathbf{B}_0 e^{-x/\lambda} \tag{4.15}$$

satisfies the differential equation (4.14). This solution clearly shows that the magnetic field decays rapidly in the interior of the superconductor, over a characteristic length,

$$\lambda \equiv \sqrt{\frac{mc^2}{4\pi e^2 n_{\mathrm{s}}}}, \tag{4.16}$$

known as the penetration depth.

At this point, it is enlightening to comment upon the implications of the foregoing phenomenological approach. We saw that the assumption of zero resistance is not sufficient to fully describe superconductors, and as such, cannot be the defining feature of superconductivity. However, assuming the London equation (4.12) as our starting point, and substituting this back into (4.9), we find

$$\frac{\partial \mathbf{j}_{\mathrm{s}}}{\partial t} - \frac{1}{2en_{\mathrm{s}}}\nabla j_{\mathrm{s}}^2 - \frac{e^2 n_{\mathrm{s}}}{m}\mathbf{E} = 0. \tag{4.17}$$

In the interior of the sample, where $\mathbf{B} = 0$, (4.17) implies $E = 0$, from which we can infer zero resistivity. Thus, the Meissner-Oschenfeld effect implies superconductivity and is, therefore, the true defining feature of a superconductor, though not necessarily its most prominent.

Returning to the superfluid density, let us recall the definition (4.8)

$$\Upsilon(T) = \frac{1}{V}\left(\frac{\partial^2 \mathcal{F}(k_0)}{\partial k_0^2}\right)_{k_0=0}. \tag{4.18}$$

We saw that one way to introduce the helical twist was through the replacement $\nabla \rightarrow \nabla + i\mathbf{k}_0$ in the kinetic energy. Such a shift is very reminiscent of the minimally coupled way by which the electromagnetic gauge field enters, $\nabla \rightarrow \nabla - i(e^*/\hbar c)\mathbf{A}$, where $e^* = 2e$ is the charge of the Cooper pairs. In fact, identifying $\mathbf{k}_0 = -(e^*/\hbar c)\mathbf{A}$, we can re-express the helicity modulus in terms of $A$:

$$\Upsilon(T) = \frac{1}{V}\left(\frac{\hbar c}{e^*}\right)^2 \left(\frac{\partial^2 \mathcal{F}(A)}{\partial A^2}\right)_{A=0}. \tag{4.19}$$

This can be related to the penetration depth $\lambda$, by noting that $\partial(\mathcal{F}/V)/\partial A = -j_s/c$, and from the London equation (4.12), $j_s = -(c/4\pi\lambda^2)A$. Then,

$$\Upsilon(T) = -\frac{1}{c}\left(\frac{\hbar c}{e^*}\right)^2\left(\frac{\partial j_s}{\partial A}\right)_{A=0} = \left(\frac{\hbar c}{e^*}\right)^2 \frac{1}{4\pi\lambda^2(T)}, \tag{4.20}$$

which gives us experimental access to the superfluid density in superconductors, through measurement of $\lambda$[6].

In the high-temperature superconductors, which are quasi two dimensional materials with a layer thickness $d$, the helicity modulus in the CuO planes can be written in units of Kelvin as

$$\widetilde{\Upsilon}_{2D}(T) = \frac{\hbar^2 c^2 d}{16\pi k_B e^2 \lambda^2(T)}. \tag{4.21}$$

Inserting the requisite constants, one finds

$$\widetilde{\Upsilon}_{2D}(T) = 6.2\mathrm{K}\frac{d/10\text{Å}}{(\lambda(T)/\mu\mathrm{m})^2}, \tag{4.22}$$

which allows us to easily convert between the experimentally reported penetration depth, and the helicity modulus.

## 4.4 Scaling theory for the helicity modulus

The scaling theory of critical phenomena[99] predicts universal behaviour for critical exponents and amplitudes near continuous phase transitions. The basic idea is that there is a diverging length scale $\xi$ which is responsible for all singular behaviour as the critical point is approached. To see this, we can write, from dimensional analysis, the singular part of the free energy density

$$f_s = k_B T(\xi^{\pm})^{-D}\mathcal{A}^{\pm}, \tag{4.23}$$

where $D$ is the dimensionality. For the classical critical phenomena implied here, we will take the tuning parameter to be $t = (T - T_c)/T_c$, and so the $\pm$ superscript indicates the sign of $t$. The correlation length has the power-law divergence,

$$\xi^{\pm} = \xi_0^{\pm}|t|^{-\nu}, \tag{4.24}$$

determined by the critical exponent $\nu$, and we can now compute thermodynamic quantities such as, for example, the specific heat:

$$\begin{aligned} c_V &= -T\frac{\partial^2 f_s}{\partial T^2} \\ &\approx \mathcal{B}^{\pm}|t|^{-\alpha}. \end{aligned} \tag{4.25}$$

We have made an additional scaling *ansatz* for the singular part of $c_V$ to introduce the critical exponent $\alpha$. Inserting the free energy (4.23) into (4.25) leads to the usual hyper-scaling relation $2 - \alpha = \nu D$, which is the well known Josephson scaling law. Additionally, one can deduce the amplitude relation $\mathcal{B}^{\pm} = -(2-\alpha)(1-\alpha)k_{\mathrm{B}}(\xi_0^{\pm})^{-D}\mathcal{A}^{\pm}$.

At zero temperature, the situation changes slightly[100]. The integral over imaginary time $0 \leq \tau \leq (k_{\mathrm{B}}T)^{-1}$ extends over the entire positive real axis, so the temporal correlation length $\xi_\tau$ can diverge. That is to say, the correlation volume $\xi^D\xi_\tau$ grows unbounded in all directions as the critical point is approached. For such quantum critical phenomena, we assume a tuning parameter $\widetilde{\delta}$ to control the transition at $\widetilde{\delta} = 0$. The scaling form of the free energy can now be written

$$f_{\mathrm{s}} = (\xi_\tau)^{-1}(\xi^{\pm})^{-D}\mathcal{A}^{\pm}. \tag{4.26}$$

In analogy with the diverging spatial correlation length,

$$\xi^{\pm} = \xi_0^{\pm}|\widetilde{\delta}|^{-\nu}, \tag{4.27}$$

we define a temporal correlation length exponent via

$$\xi_\tau^{\pm} = \xi_{\tau,0}^{\pm}|\widetilde{\delta}|^{-\nu_\tau}. \tag{4.28}$$

Substituting these into (4.26), we find

$$f_{\mathrm{s}} = \frac{\mathcal{A}^{\pm}|\widetilde{\delta}|^{\nu(D+z)}}{\xi_{\tau,0}^{\pm}(\xi_0^{\pm})^D}, \tag{4.29}$$

where the dynamical critical exponent is defined to be $z = \nu_\tau/\nu$. The scaling form (4.4) simply demonstrates the fact the quantum critical phenomena are just like classical critical phenomena, but in $(D + z)$ dimensions. This suggests, for example, the modified Josephson scaling law $2 - \alpha = \nu(D + z)$.

As our interest is in the helicity modulus, we now turn to consideration of the free energy (4.4) in the presence of a twist in the order parameter phase. We make the simple modification

$$f_{\mathrm{s}} = \frac{\mathcal{A}^{-}|\widetilde{\delta}|^{\nu(D+z)}}{\xi_{\tau,0}^{-}(\xi_0^{-})^D}\Phi(\mathcal{C}k_0|\widetilde{\delta}|^{-\nu}), \tag{4.30}$$

as determined from dimensional analysis, and use the definition (4.8) to find

$$\Upsilon(T=0) = \frac{\mathcal{A}^{-}\mathcal{C}^2|\widetilde{\delta}|^{\nu(D+z-2)}}{\xi_{\tau,0}^{-}(\xi_0^{-})^D}. \tag{4.31}$$

Extending this to small but finite temperatures, we write

$$\Upsilon(T) = \frac{\mathcal{A}^{-}\mathcal{C}^2|\widetilde{\delta}|^{\nu(D+z-2)}}{\xi_{\tau,0}^{-}(\xi_0^{-})^D}\Psi(k_{\mathrm{B}}T\xi_\tau). \tag{4.32}$$

Now, assuming a line of critical points exists $T_{\mathrm{c}}(\widetilde{\delta})$, the critical temperature will scale as $T_{\mathrm{c}} \sim |\widetilde{\delta}|^{\nu z}$, and the helicity modulus, in the limit $\widetilde{\delta} \to 0$, behaves as

$$\Upsilon(T=0) \sim T_{\mathrm{c}}^{(D+z-2)/z}. \tag{4.33}$$

Remarkably, for 2D criticality, this relation implies $\Upsilon(0)/T_{\mathrm{c}} = Q$ ($Q$ constant) *independently* of the nature of the quantum critical point, i. e. $z$. This behaviour has been seen in experiments on the high temperature superconductors[23, 24], and is widely known as Uemura scaling.

At higher temperatures, the helicity modulus takes the form

$$\Upsilon(T) = k_{\mathrm{B}}T\left(\frac{\mathcal{A}^{-}\mathcal{C}^2}{(\xi_0^{-})^D}\right)|t|^{\nu(D-2)}. \tag{4.34}$$

In two dimensions, this predicts the universal jump of the helicity modulus at $T_{\mathrm{c}}$,

$$\frac{\Upsilon(T_{\mathrm{c}})}{k_{\mathrm{B}}T_{\mathrm{c}}} = \text{constant}, \tag{4.35}$$

which is known as the Nelson-Kosterlitz jump[101]. The renormalization group establishes that the constant in this relation is $2/\pi$[101].

In the next section, we will review the experimental results. In particular, we will see to what extent the Uemura scaling is satisfied, and extract the amplitude $Q$.

## 4.5 Uemura scaling

The apparent proportionality between the zero-temperature superfluid density and the critical temperature was first observed in muon spin relaxation measurements[23, 24] and for a wide array of cuprate superconductors. From their data, we can estimate the slope to be 3.0 K $\cdot$ $\mu$m$^2$, with a margin of error of $\pm 0.2$ for the YBaCuO data and $\pm 0.5$ for LaSrCuO[102]. Converting the slope to a dimensionless amplitude $Q$ via the formula (4.22) requires knowledge of $d$, the layer thickness. Assuming this to be the minimal thickness along the $c$-axis so that the material superconducts, $d$ can be determined from measurements on superconductor-insulator superlattices[103] or from crossing-point phenomena in magnetization measurements[102]. Otherwise, the width of a unit cell in the

$c$ direction can be used as an upper bound. For YBaCuO, we find 10 Å $< d <$ 12 Å producing $1.82 < Q < 2.48$, while for LaSrCuO, $d \approx 7.6$ Å[102] and $1.37 < Q < 1.88$. Evidently, $Q$ is not truly universal, due to its dependence on $d$, though we may say it is around 2 for most materials, and certainly greater than 1.37.

## 4.6 Depleting the superfluid density: the Ioffe-Larkin rule

In principle, the depletion of the superfluid density is due to several effects. However, we have argued that only $d$-wave nodal quasiparticles and order parameter phase fluctuations are relevant at low energies. The contributions from these two effects can be summarized in the so-called Ioffe-Larkin rule[33]:

$$\frac{1}{\rho_{\rm s}(T)} = \frac{1}{\rho_{\rm pf}(T)} + \frac{1}{\rho_{\rm qp}(T)}, \tag{4.36}$$

where $\rho_{\rm pf}$ results from phase fluctuations and $\rho_{\rm qp}$, quasiparticles. This simple relation can be easily derived from our effective theory, presented in Chapter 2, as we now show.

In the superconducting state, the full action (2.52) becomes

$$S[\Psi_i, b_n, a, v] = S_f[\Psi_i, a, v] + \int_0^\beta {\rm d}\tau \int {\rm d}^2\mathbf{r} \left\{ \frac{\widetilde{K}}{2}(2v_\mu + (e^*/\hbar c)A_\mu)^2 + 2J(T)(v_\mu^2 + a_\mu^2) \right\}. \tag{4.37}$$

In expressing this, we have restored the dimensionful constants, and written the condensate prefactor of the last term as $J(T)$. As far as the charge part of the action is concerned, integrating out spinons in the above action just causes the bare stiffness $\widetilde{K}$ to acquire a temperature dependence, so we can include the effect of $S_f$ by simply making the replacement $\widetilde{K} \to \widetilde{K}(T)$:

$$S_{\rm charge} = \int_0^\beta {\rm d}\tau \int {\rm d}^2\mathbf{r} \left\{ \frac{\widetilde{K}(T)}{2}(2v_\mu + (e^*/\hbar c)A_\mu)^2 + 2J(T)v_\mu^2 \right\}. \tag{4.38}$$

Integrating out $v$ is a trivial matter, and leads to

$$S_{\rm charge} \to \int_0^\beta {\rm d}\tau \int {\rm d}^2\mathbf{r} \left\{ \frac{1}{2} \left( \frac{J(T)\widetilde{K}(T)}{J(T) + \widetilde{K}(T)} \right) \left( \frac{e^* A_\mu}{\hbar c} \right)^2 \right\}. \tag{4.39}$$

The action is now in a form which is amenable to the determination of the helicity modulus via the definition (4.19). We easily identify

$$\Upsilon(T) = \left( \frac{1}{J(T)} + \frac{1}{\widetilde{K}(T)} \right)^{-1}. \tag{4.40}$$

The relationship (4.7) connects the helicity modulus to the superfluid density; applying it here clearly reproduces the Ioffe-Larkin rule above (4.36) where $J(T)$, arising from phase fluctuations alone, is proportional to $\rho_{\rm pf}(T)$ and $\widetilde{K}(T) \propto \rho_{\rm qp}(T)$, having resulted from integration over spinons. From here on in, to avoid any confusion, we will write these contributions as $\Upsilon_{\rm pf}(T)$ and $\Upsilon_{\rm qp}(T)$, respectively.

To make contact with experiment, we can now evaluate the quasiparticle and bosonic superfluid densities separately, and combine them using the rule (4.36). This task will be the focus of the remainder of this chapter. In particular, we will be interested in the relative importance of each contribution to see which is the factor that determines the Uemura scaling.

## 4.7 Quasiparticles

At optimal doping in high temperature superconductors, the quasiparticle fluctuations are expected to be the dominating contribution[6]. The reason for this is that the zero-temperature superfluid density (i. e. phase stiffness) is quite high, so that phase fluctuations are inhibited. At lower dopings, the stiffness is reduced, thus increasing the role of phase fluctuations[26]. To understand the importance of the respective contributions, we consider each in turn, starting here with the quasiparticles.

Before evaluating the superfluid density due to $d$-wave quasiparticles, it is useful to compare with the situation in $s$-wave superconductors. The form of the gap in these two cases is quite different, as shown in Figures 4.1 and 4.2. As we can see, in the $s$-wave case, an isotropic gap forms around the Fermi surface, and an energy proportional to $\Delta_0$ is required to excite quasiparticles. On the other hand, the $d$-wave form displays nodes at special points on the Fermi surface, so only an infinitesimal energy is required to excite quasiparticles in their vicinity. It is natural to speculate, then, that $d$-wave quasiparticles will be much more effective at depleting the superfluid density than their $s$-wave counterparts.

The evaluation of the superfluid density in the $s$-wave case is standard[3] and we'll just quote the result here. The temperature dependence, as a result of the full isotropic gap, has the exponentially activated form[16]

$$\frac{\rho_{\rm qp}(T) - \rho_{\rm qp}(T=0)}{\rho_{\rm qp}(T=0)} \propto \left(\frac{\Delta_0}{T}\right)^{\frac{1}{2}} \exp\left\{-\frac{\Delta_0}{T}\right\}, \tag{4.41}$$

where we have set the Boltzmann constant $k_B = 1$. To recapitulate, at temperatures below the gap energy, thermal fluctuations with sufficient energy to excite quasiparticles are exponentially rare,

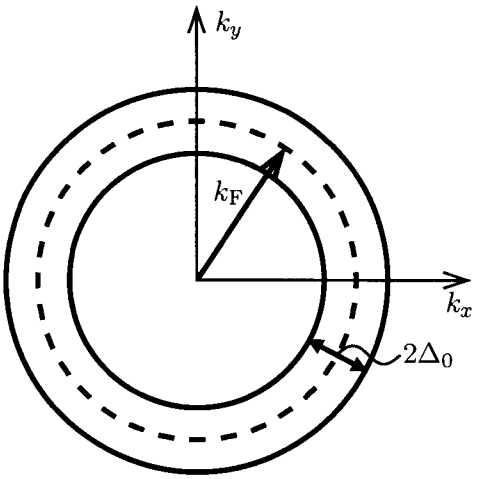


Figure 4.1: Sketch of the $s$-wave gap function. The Fermi wave vector is indicated by $k_F$ and the gap magnitude is $\Delta_0$.

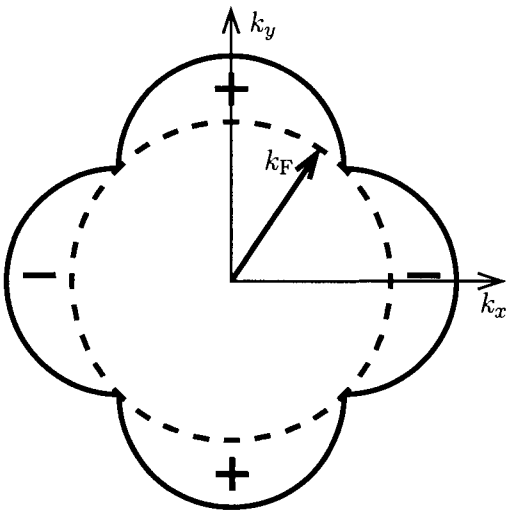


Figure 4.2: Sketch of the $d$-wave gap function. The Fermi wave vector is indicated by $k_F$. The gap function changes sign around the Fermi surface, resulting in the four nodal points.

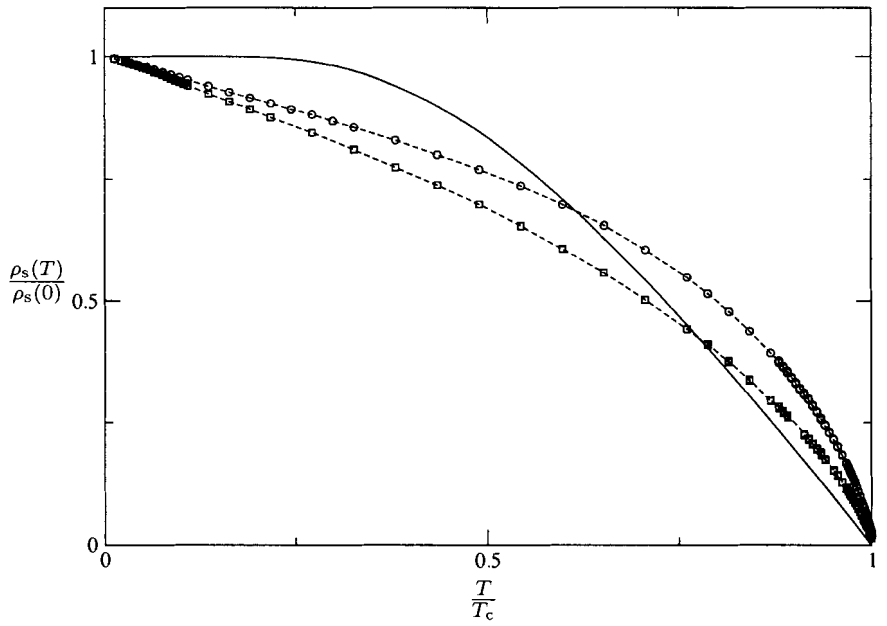


Figure 4.3: The depletion of the superfluid density due to quasiparticles. The solid line is the $s$-wave result[16]. The data points are experimental results on optimally doped YBCO – open circles are from $a$-axis measurements, open squares are $b$-axis, taken from Bonn et al. [88].

and the superfluid density is effectively undepleted. At higher temperatures, quasiparticles are more readily able to overcome the barrier, and the depletion proceeds rapidly. This is displayed by the solid line in Figure 4.3.

Also plotted are experimental results on optimally doped YBCO taken from[88]. The low temperature behaviour is drastically different from the $s$-wave case, indicating that the order parameter of the cuprates must have a different symmetry[6]. Of course, it is now well known that the high temperature superconductors are $d$-wave, as we've been advocating throughout. To see what effect these quasiparticles have on the depletion of the superfluid density, we will now perform the calculation alluded to in the previous section; that is, integrate out fermions in $S_f$.

Upon integration over fermions, the action develops a term $(1/2)\Pi_{\mu\nu}(q)A_\mu(q)A_\nu(-q)$, where $\Pi_{\mu\nu}(q)$ is the fermion polarization or current-current correlation function[52]. It follows, then, from the definition (4.19), that the helicity modulus in the $i$ direction is simply $\Pi_{ii}(q=0)$. To one-loop order, this is represented diagrammatically in Figure 4.4, where the wavy lines represent the electromagnetic gauge field, and the solid lines are fermion propagators. Explicitly, we can write,

for the $x$-direction,

$$\begin{aligned}\Pi_{11}(0) &= -\mathrm{Tr}\, v_f^2 G(\mathbf{k},\omega_n)\gamma_5 G(\mathbf{k},\omega_n)\gamma_5 \\ &= -v_f^2 T\sum_{\omega_n}\int \frac{\mathrm{d}^2\mathbf{k}}{(2\pi)^2}\frac{\mathrm{tr}(\gamma_\mu\gamma_0\gamma_5\gamma_\nu\gamma_0\gamma_5)k_\mu k_\nu}{(\omega_n^2+v_f^2k_x^2+v_\Delta^2k_y^2)^2}.\end{aligned} \tag{4.42}$$

The trace over the $\gamma$ matrices is easily evaluated and yields $8\delta_{\mu,0}\delta_{\nu,0}-4\delta_{\mu,\nu}$. Redefining $v_f k_x \to k_x$ and $v_\Delta k_y \to k_y$, the above integral becomes

$$\Pi_{11}(0) = -\left(\frac{v_f}{v_\Delta}\right) T\sum_{\omega_n}\int \frac{\mathrm{d}^2\mathbf{k}}{(2\pi)^2}\frac{8\omega_n^2-4k^2}{(\omega_n^2+k^2)^2}. \tag{4.43}$$

The Matsubara sums can be done in the usual way[10] and the resulting integral over $k$ is not difficult. The final result is

$$\Upsilon_{\mathrm{qp}}(T)-\Upsilon_{\mathrm{qp}}(0) = \Pi_{11}(0) = -\left(\frac{2\ln 2}{\pi}\frac{v_f}{v_\Delta}\right)T, \tag{4.44}$$

which is well known from other methods[104–106]. We see that the temperature dependence due to nodal quasiparticles is quite different from the fully gapped $s$-wave case: the depletion of the helicity modulus is now *linear* in T.

In the cuprates, the $T = 0$ superfluid density grows continuously as optimal doping is approached. From this fact, we can infer that, at large enough dopings, the quasiparticle contribution to (4.36) will dominate, since the phase fluctuation term would be a very small addition. We would expect, then, experiments on heavily doped materials to display the low temperature behaviour described in (4.44). Indeed, this is exactly what has been seen in materials near optimal doping [6](see Figure 4.3), and, at the time of its discovery, provided the first compelling evidence for $d$-wave pairing in the cuprates.

As the materials are *under*doped, the phase stiffness shrinks, and is proportional to doping at low dopings. It then seems natural that the phase fluctuation contribution becomes dominant near the underdoped quantum critical point, and so must be taken into account.

## 4.8 Phase fluctuations

Our model for phase fluctuations is the XY-model (2.36) of Chapter 2. In the following, we will adopt a slightly more general theory, including interactions and anisotropy, which we refer to as the

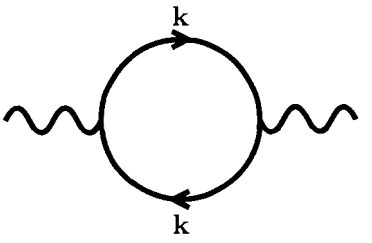


Figure 4.4: The fermion polarization bubble at $q = 0$. Wavy lines represent the gauge field, and solid lines are the fermion propagator. The vertex function is $v_f \gamma_5$.

anisotropic quantum XY model:

$$\begin{aligned} S &= -K \int_0^\beta d\tau \sum_{\mathbf{r},\mu,z} \cos(\phi_{\mathbf{r}+\hat{e}_\mu,z}(\tau) - \phi_{\mathbf{r},z}(\tau)) \\ &\quad -\alpha K \int_0^\beta d\tau \sum_{\mathbf{r},z} \cos(\phi_{\mathbf{r},z+1}(\tau) - \phi_{\mathbf{r},z}(\tau)) \\ &\quad + \frac{1}{2\lambda^2} \int_0^\beta d\tau \sum_{\mathbf{r},\mathbf{r}',z,z'} \dot{\phi}_{\mathbf{r},z}(\tau) V^{-1}(|\mathbf{r}-\mathbf{r}'|, z-z') \dot{\phi}_{\mathbf{r}',z'}(\tau). \end{aligned} \tag{4.45}$$

The first two terms of the action describe the usual classical phase fluctuations, in the plane and out of the plane, respectively. The last term represents the quantum phase fluctuations arising from the number-phase uncertainty relation $\Delta n \cdot \Delta\phi \geq 1$ for the canonically conjugate pair $(n, \phi)$, where $n$ is a density. The interactions tend to fix the density, and so the phase fluctuates due to this quantum mechanical relation. To be specific, we will assume the Coulomb interaction,

$$V(|\mathbf{r}-\mathbf{r}'|, z-z') = \left(\sqrt{|\mathbf{r}-\mathbf{r}'|^2 - (z-z')^2}\right)^{-1}, \tag{4.46}$$

though our results are general. For illustration, we will also consider the short-range repulsion, $V = \delta_{\mathbf{r},\mathbf{r}'}\delta_{z,z'}$. The parameter $\alpha$ tunes the anisotropy between the $x-y$ plane and the $z$ direction. We will mostly be interested in $\alpha = 0$, but will comment on the effect of finite $\alpha$ later.

The temperature dependent helicity modulus can be calculated using the variational Gibbs-Boboliubov-Feynman (GBF) approach[80] introduced in Chapter 3 (3.15). In the present case, we will approximate the theory (4.45) with the trial action,

$$\begin{aligned} S &= \frac{\Upsilon_{\rm pf}(T)}{2} \int_0^\beta d\tau \sum_{\mathbf{r},\mu} \left(\phi_{\mathbf{r}+\hat{e}_\mu}(\tau) - \phi_{\mathbf{r}}(\tau)\right)^2 \\ &\quad + \frac{1}{2\lambda^2} \int_0^\beta d\tau \sum_{\mathbf{r},\mathbf{r}'} \dot{\phi}_{\mathbf{r}}(\tau) V^{-1}(|\mathbf{r}-\mathbf{r}'|) \dot{\phi}_{\mathbf{r}'}(\tau), \end{aligned} \tag{4.47}$$

where we have set $\alpha = 0$. The free energy can be evaluated exactly as before, except now we are working at finite temperature. We can write

$$\begin{aligned} F_{\rm var} &= -\frac{1}{2}\frac{T}{N}\sum_{\mathbf{k},\omega_n}\ln\left(G_0(\mathbf{k},\omega_n)\right) \\ &\quad -K\sum_{\mathbf{R}_{ij}}\exp\left\{-2\frac{T}{N}\sum_{\mathbf{k},\omega_n}G_0(\mathbf{k},\omega_n)\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{R}_{ij}}{2}\right)\right\} \\ &\quad -2\Upsilon_{\rm pf}(T)\frac{T}{N}\sum_{\mathbf{k},\omega_n}G_0(\mathbf{k},\omega_n)\sum_{\mathbf{R}_{ij}}\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{R}_{ij}}{2}\right), \end{aligned} \tag{4.48}$$

where $\mathbf{R}_{ij}$ is a vector pointing to nearest neighbour sites, and we can determine the propagator by Fourier transforming the trial action (4.47):

$$G_0^{-1}(\mathbf{k},\omega_n) = 4\Upsilon_{\rm pf}(T)\sum_{\mathbf{R}_{ij}}\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{R}_{ij}}{2}\right) + \frac{\omega_n^2}{V(k)\lambda^2}. \tag{4.49}$$

Rather than minimizing the variational free energy with respect to the propagator, as we did before, it is now more convenient to minimize with respect to $\Upsilon_{\rm pf}(T)$. This directly yields the self-consistent equation for the helicity modulus

$$\begin{aligned} \Upsilon_{\rm pf}(T) &= K\exp\left\{-\frac{T}{N}\sum_{\mathbf{k},\omega_n}G_0(\mathbf{k},\omega_n)\sum_{\mathbf{R}_{ij}}\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{R}_{ij}}{2}\right)\right\} \\ &= K\exp\left\{-\frac{1}{4}\sqrt{\frac{\lambda^2}{\Upsilon_{\rm pf}(T)}}\int\frac{{\rm d}^2\mathbf{k}}{(2\pi)^2}F(\mathbf{k})\coth\left(\frac{\sqrt{\Upsilon_{\rm pf}(T)\lambda^2}}{T}F(\mathbf{k})\right)\right\}. \end{aligned} \tag{4.50}$$

The second line results from evaluating the Matsubara frequency sum and taking the continuum k limit; we've also defined the function

$$F(\mathbf{k}) = \left[V(\mathbf{k})\sum_{\mathbf{R}_{ij}}\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{R}_{ij}}{2}\right)\right]^{\frac{1}{2}}. \tag{4.51}$$

In the classical limit $\lambda^2 = 0$, we can easily determine the low temperature behaviour of $\Upsilon_{\rm pf}$. The self-consistent equation becomes

$$\Upsilon_{\rm pf}(T) = K\exp\left\{-\frac{T}{4\Upsilon_{\rm pf}(T)}\right\}, \tag{4.52}$$

which, to lowest order in $T$, yields $\Upsilon_{\rm pf}(T) = K - T/4$ which is linear, just like the quasiparticle contribution. This is the known result [107, 108] for the depletion of the helicity modulus due to

spin waves excitations, giving credence to the variational approach. At higher temperatures, transverse vortex excitations must be taken into account, though these are not expected to be important except very close to $T_c$[109]. Here and in the following, we will assume that these excitations serve only to establish the critical temperature through the Nelson-Kosterlitz jump in the stiffness (4.35), $\Upsilon(T_c)/T_c = 2/\pi$. For the non-interacting case in our self-consistent Gaussian approximation, this yields the Uemura ratio $Q(\alpha = 0, \lambda^2 = 0) = (2/\pi + 1/4) \approx 0.89$, which should be compared with the known value, incorporating vortex effects, $Q_{2DXY} \approx 1.11$[107, 108]. Anyway, it is clear that classical phase fluctuations alone cannot account for the observed value of the ratio $Q \sim 2$.

We include quantum phase fluctuations by setting $\lambda^2 \neq 0$. Interaction affects the helicity modulus in two ways: First, its zero temperature value is reduced from the bare value $K$. This can be seen by setting $T = 0$ in the self-consistent equation (4.50),

$$\begin{aligned} \Upsilon_{pf}(0) &= K \exp\left\{ -\frac{1}{4}\sqrt{\frac{\lambda^2}{\Upsilon_{pf}(0)}} \int \frac{d^2\mathbf{k}}{(2\pi)^2} F(\mathbf{k}) \right\} \\ &= K\left[ 1 - \frac{1}{4}\sqrt{\frac{\lambda^2}{K}} \int \frac{d^2\mathbf{k}}{(2\pi)^2} F(\mathbf{k}) + \mathcal{O}(\lambda^3) \right]. \end{aligned} \tag{4.53}$$

In the last line, we have expanded in powers of the interaction strength; the term $\mathcal{O}(\lambda^2)$ cancels when the helicity modulus is reinserted self-consistently. The correction to $\Upsilon_{pf}(0)$ from interactions is negative, and so the helicity modulus is reduced from its bare value.

The second effect of interactions is the introduction of a new quantum energy scale, below which the depletion of $\Upsilon_{pf}(T)$ is suppressed; for example, the energy scale for Coulomb interactions is the plasma frequency[97]. We can see this by determining the temperature dependence from (4.50). For short-ranged interaction, $V(k) = 1$. At low temperatures, the integral is dominated by the region of small $k$, and so we can approximate $\sin^2(\mathbf{k}\cdot\mathbf{R}_{ij}/2) \approx (\mathbf{k}\cdot\mathbf{R}_{ij})^2/4$. Scaling $\Upsilon_{pf}(T)\lambda^2 k^2/T^2 \to k^2$, we find that the helicity modulus has the form,

$$\frac{\Upsilon_{pf}(T)}{\Upsilon_{pf}(0)} = 1 - \text{constant} \times \left(\frac{T}{T_q}\right)^3, \tag{4.54}$$

where the quantum scale is $T_q = \lambda^{2/3}\Upsilon_{pf}^{2/3}(0)$. Performing a similar operation for the case of Coulomb interactions, $V(k) = 2\pi/k$, we find

$$\frac{\Upsilon_{pf}(T)}{\Upsilon_{pf}(0)} = 1 - \text{constant} \times \left(\frac{T}{T_q}\right)^5, \tag{4.55}$$

where $T_q = \lambda^{4/5}\Upsilon_{pf}^{3/5}(0)$. So, we see that the depletion of $\Upsilon_{pf}(T)$ is much more strongly suppressed at low temperatures by quantum fluctuations than by classical fluctuations. This situation is shown in Figure 4.5.

In plotting the helicity modulus, we have estimated the critical temperature, even for the interacting case, from the Nelson-Kosterlitz relation (4.35). It has been shown[110] that for a weak Coulomb interaction, the coupling $\lambda$ is irrelevant (in the renormalization group sense), and so the transition remains in the BKT universality class. In particular, the discontinuous jump in the helicity modulus still holds, and the transition occurs at $T_c = (\pi/2)\Upsilon_{pf}(T_c)$. Using this, we can compute the ratio $Q$ in our self-consistent Gaussian approximation. Expanding in powers of $\lambda$, we find

$$T_c = \frac{\pi}{2} K e^{-\frac{\pi}{8}} \left[ 1 - \frac{1}{6} \frac{\lambda^2 e^{\frac{\pi}{8}}}{\pi K} \int \frac{d^2\mathbf{k}}{(2\pi)^2} F^2(\mathbf{k}) + \mathcal{O}(\lambda^3) \right]. \tag{4.56}$$

Putting this together with the expansion (4.53) for $\Upsilon_{pf}(0)$, we find

$$Q(\alpha = 0, \lambda^2) = \frac{2e^{\frac{\pi}{8}}}{\pi} \left[ 1 - \frac{1}{4}\sqrt{\frac{\lambda^2}{K}} \int \frac{d^2\mathbf{k}}{(2\pi)^2} F(\mathbf{k}) + \frac{1}{6} \frac{\lambda^2 e^{\frac{\pi}{8}}}{\pi K} \int \frac{d^2\mathbf{k}}{(2\pi)^2} F^2(\mathbf{k}) + \mathcal{O}(\lambda^3) \right]. \tag{4.57}$$

We notice that $Q(\alpha = 0, \lambda^2)$ clearly is a decreasing function of $\lambda^2$. This is demonstrated graphically in Figure 4.6. The symbols represent the full solution to the self-consistent equation for short-ranged ($*$) and Coulomb ($\circ$) interactions, the lines are the low order results from (4.57).

Although the preceding self-consistent Gaussian approximation in principle includes only the spin-wave phase fluctuations, it should be a reasonable indicator of the interaction dependence of $Q$. The main contribution to the depletion of the helicity modulus for $T \leq 0.8T_c$ in the classical 2DXY model comes from these spin-waves, and it is the increase of their energies with interaction that should ultimately be responsible for the proposed decrease of $Q$ from the non-interacting case. From these considerations, we conjecture that the universal ratio in 2D is bounded above by the non-interacting value, $Q_{2DXY} \approx 1.11$, well below the Uemura result. Combining this with the lower bound implied by the Nelson-Kosterlitz relation, we find that the class of models described by the action (4.45) in 2D ($\alpha = 0$) generically yields a ratio $Q$ which is restricted to the range

$$\frac{2}{\pi} \leq Q(\alpha = 0, \lambda^2) \leq Q_{2DXY}. \tag{4.58}$$

Incorporating the effect of the third dimension by slowly tuning $\alpha \neq 0$ will also be deleterious to the magnitude of $Q$. Away from the pure two dimensional system, the Nelson-Kosterlitz jump in the helicity modulus ceases to hold, and the curves should be expected to become rounded

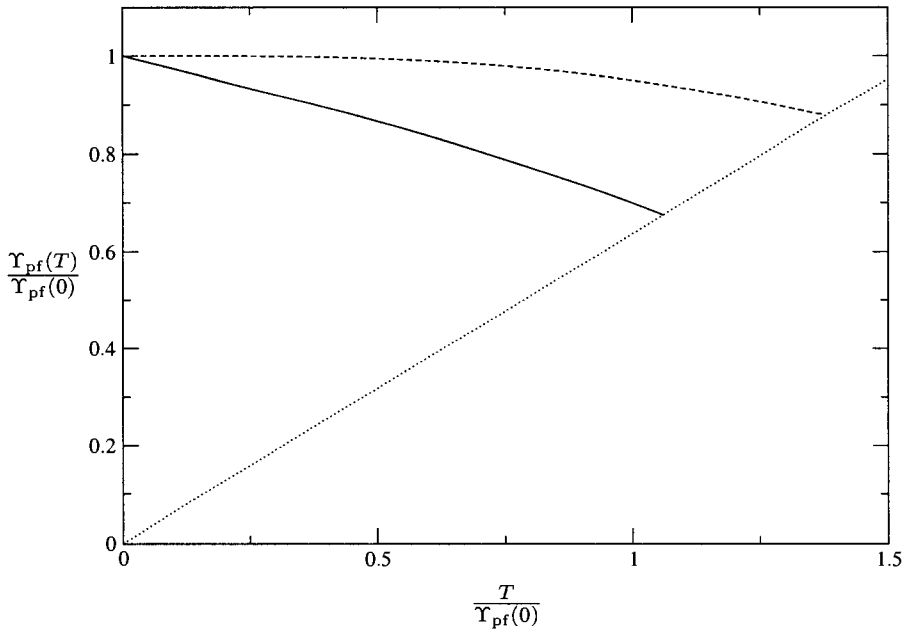


Figure 4.5: Helicity modulus vs. temperature in the self-consistent Gaussian approximation, for the 2D XY model (solid line; $\alpha = e_* = 0$), and with Coulomb interactions (dashed line; $\alpha = 0$, $e_*^2 = 5/\pi$ and $V(k) = 2\pi/k$). The straight dotted line denotes the universal BKT limit, $\Upsilon_{\rm pf}(T_c) = (2/\pi)T_c$, at which the helicity modulus discontinuously vanishes. Note that $T_c$ for fixed $\Upsilon_{\rm pf}(0)$ *increases* with interaction.

near the transition, with higher critical temperatures. Indeed, $Q$ has been computed for the 3DXY model[111] to be $Q_{\rm 3DXY} \approx 0.45$. The upper bound (4.58) proposed above then holds for all values of the anisotropy parameter:

$$Q(\alpha, \lambda^2) \leq Q_{\rm 2DXY}. \tag{4.59}$$

So, it seems that the experimental value for $Q$ is well above the upper bound coming from phase fluctuations alone. To bring the theory in better agreement with the data, we would have to assume a much lower value for the layer thickness $d$. Using $d = 6$Å yields $Q = 1.09$ and just below our upper bound, although this is a much smaller thickness than has been measured. It seems more reasonable, then, that another mechanism must be responsible for reducing the superfluid density. Fortunately, we have one such – quasiparticles.

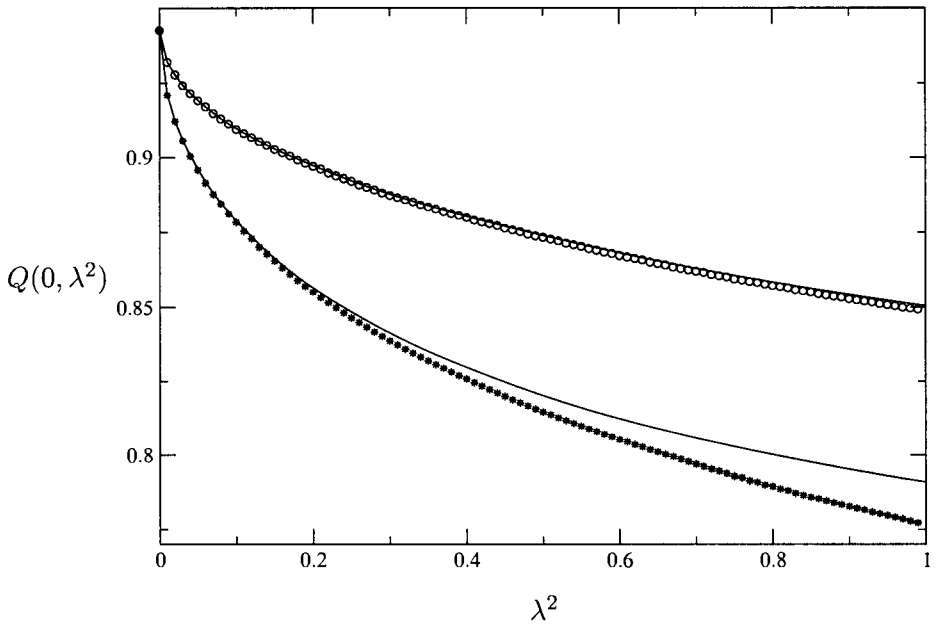


Figure 4.6: The amplitude $Q(\alpha = 0, \lambda^2) = \Upsilon_{\rm pf}(0)/T_c$ in the self-consistent Gaussian approximation for the helicity modulus in 2D, with $T_c$ defined by $\Upsilon_{\rm pf}(T_c) = (2/\pi)T_c$, for long (o) and short-range ($*$) interactions. The lines are the weak-interaction results from (4.57); $K = 1$.

## 4.9 Phase fluctuations + quasiparticles

As we established earlier, the superfluid density resulting from $d$-wave quasiparticles has the form

$$\Delta\Upsilon_{\rm qp}(T) = -\frac{2\ln 2}{\pi}\frac{v_f}{v_\Delta}T + \mathcal{O}(\frac{T^2}{\Delta}). \tag{4.60}$$

Higher order in temperature corrections are negligible in the very underdoped case where $T_{\rm c} \ll \Delta$, and we will drop them hereafter. We can now combine the quasiparticle and phase fluctuation contributions by way of the Ioffe-Larkin rule derived previously (4.36). It must be noted that the temperature dependence of the phase fluctuation part is a much higher power than linear, and can be dropped in comparison to the quasiparticle part. To make this more concrete, notice that the depletion in, for example, the Coulomb case (4.55) is significant only when $T \sim T_{\rm q}$. From the scaling arguments near the quantum critical point (4.33), $\Upsilon_{\rm pf}(0) \sim T_{\rm c}$ and so $T_{\rm q} \gg T_{\rm c}$ in the very underdoped cuprates. The Ioffe-Larkin rule now leads to

$$\Upsilon(T) = \Upsilon(0) - \left(\frac{2\ln 2}{\pi}\frac{v_f}{v_\Delta}z^2\right)T. \tag{4.61}$$

The factor $z^2 \equiv \Upsilon_{\rm pf}(0)^2/(\Upsilon_{\rm pf}(0) + \Upsilon_{\rm qp}(0))^2$ is often referred to as the charge renormalization factor, in analogy with a Landau Fermi liquid parameter accounting for screening by the background. To elucidate, we could have included the effect of phase fluctuations in the spinon action (2.31) through the *phenomenological* replacement of the term $iA_iJ_i \rightarrow izA_iJ_i$. Integrating out fermions as we did earlier in this chapter then leads to exactly the expression above. In the language of that section, the replacement amounts to a renormalization of vertex function $v_f\gamma_5 \rightarrow zv_f\gamma_5$.

Even though we have concluded that it is the quasiparticles which determine the temperature dependence of the helicity modulus, it is still the unbinding of vortices in 2D which leads to the transition, and we expect the critical temperature to be determined by the relation (4.35):

$$\Upsilon(0) - yT_{\rm c} = \frac{2}{\pi}T_{\rm c}, \tag{4.62}$$

where the slope $y$ is the bracketed expression in (4.61). This predicts a particularly simple form for $\Upsilon(T)$, as depicted schematically in Figure 4.7. The slope $y$ of the helicity modulus is expected to change with doping, as plotted; this stems from the factor $z^2v_f/v_\Delta$. The $d$-wave gap velocity $v_\Delta$ is known to increase with underdoping[22], while the Fermi velocity $v_f$ remains approximately constant[17]. The behaviour of the charge renormalization factor is much less well established, but if anything, decreases with underdoping[112]. Altogether, this predicts a decreasing slope as we approach the underdoped critical point. The effect of vortex fluctuations near $T_{\rm c}$ will be to deplete the superfluid density even more, thus lowering the critical temperature. Additionally, with $0 < \alpha \ll 1$, some rounding of the BKT jump should be expected, along with a slight increase in $T_{\rm c}$. These fine effects are being omitted in the Figure, but should be expected in experiments.

The addition of the quasiparticles contribution $y$ in (4.61) now makes the agreement with the old Uemura data much better than with phase fluctuations alone. From heat transport measurements[22], the ratio $v_f/v_\Delta \sim 10$, and the charge renormalization is estimated to be $z \sim 0.5 - 1$ in underdoped materials. Putting this together, we find $y \sim 1$. Using (4.62), we determine the ratio $Q = y + 2/\pi \approx 2$, as has been observed experimentally. More importantly, the doping dependence of $y$, thus, implies a systematic deviation of $T_{\rm c}(\Upsilon(0))$ from the simple linear relation. Even if we conservatively assume $z$ remains constant with doping, the decrease of $v_\Delta$ makes the slope $y$ a decreasing function of doping and, therefore, of $\Upsilon(0)$. The curve $T_{\rm c}$ versus $\Upsilon(0)$ is then a *convex up* function, which seems to have been observed recently.

To be more explicit, note that our expression (4.62) with the relation (4.22) between $\Upsilon$ and $1/\lambda^2$

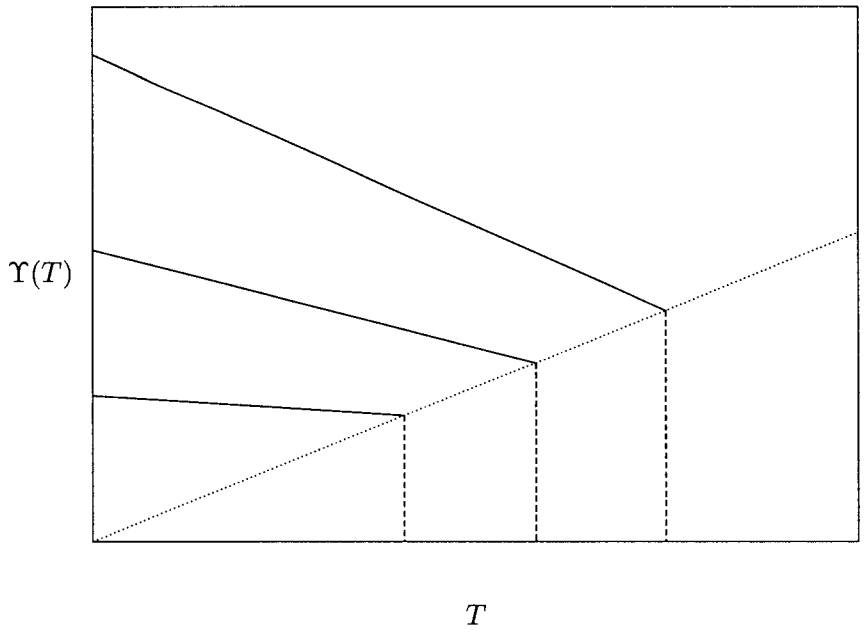


Figure 4.7: Schematic behaviour of the helicity modulus vs. temperature in a low-$T_c$, large gap, quasi-2D d-wave superconductor. Doping is assumed to be decreasing from top to bottom. Note that both $T_c$ and the slope should decrease with underdoping (see the text). Dashed line again represents the universal BKT limit.

yields

$$T_c = \frac{3.1\pi}{1 + \ln(2) z^2 (v_f / v_\Delta)} \left( \frac{1}{\lambda^2(0)} \right). \tag{4.63}$$

We can now use this expression to fit available data. Figure 4.8 shows data taken from [11] where circles(squares) are taken from $a(b)$-axis measurements. Also plotted is the Uemura result, $Q = 2$. To avoid contributions from chains in the $b$ direction, we consider only the $a$ direction, though both sets clearly show a departure from linearity. To fit the data, we need to know the doping dependence of $v_\Delta$, which we can extract from heat transport measurements [22]; we will assume fixed $v_f = 2.5 \times 10^7$ cm/s [22]. The doping dependence of $z$ is more problematic, since it is appears only in combination with $v_f$ and $v_\Delta$ and is, thus, subject to compounded uncertainty. In the Figure, we attempt to fit the data using both a constant $z = 1$ (no charge renormalization; dotted line) and a doping dependent $z = T_c/C$ (solid line) where C is a proportionality constant. The rationale for the doping dependent scaling form comes from the definition of $z$ arising from the Ioffe-Larkin rule, $z \propto \Upsilon_{\rm pf}(0)$ and from the fact that $\Upsilon(0) \propto \Upsilon_{\rm pf}(0)$. Together with the scaling arguments presented

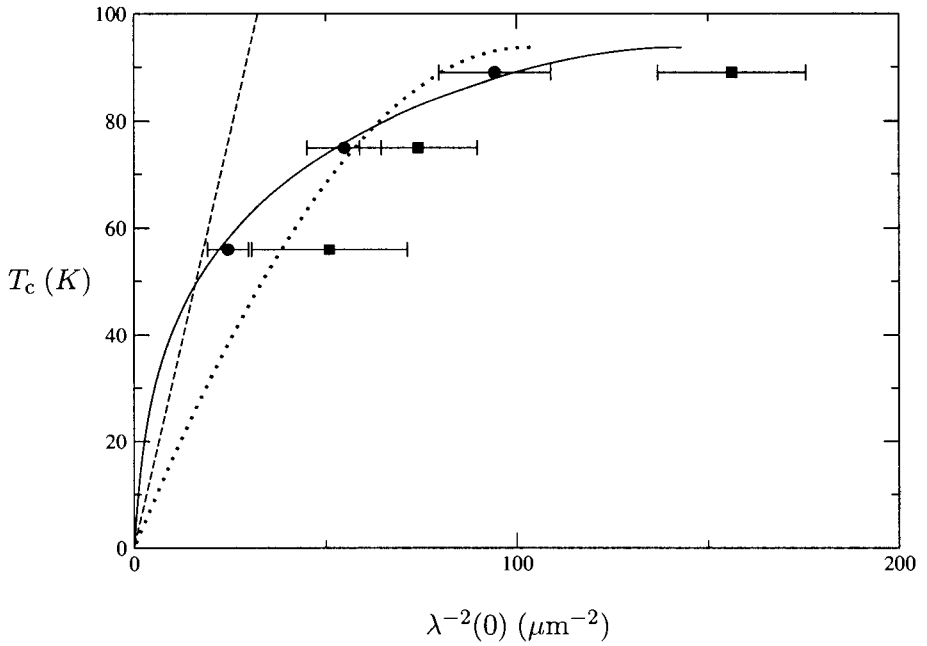


Figure 4.8: Data reported in [11]. The lines are fits to the data as discussed in the text. The convexity of the curves is due to the doping dependence of the quantity $z^2(v_F/v_\Delta)$.

earlier, that $\Upsilon(0) \sim T_c$, this implies the doping dependent $z \propto T_c$ as proposed.

We can see from the Figure that the data do not display the Uemura relation, but are better described by a convex up curve. The doping independent fit clearly shows such convexity, resulting entirely from the dependence of $v_\Delta$ on doping, which however does not fit the data very well. On the other hand, assuming further the suggested doping dependent form for $z$, we can fit the data much better. In that case, the best fitting parameter is $C = 79$ K.

## 4.10 Conclusions

The underdoped cuprates are characterized by a low critical temperature, large $d$-wave pairing gap and quasi-two dimensionality. In this chapter, we examined the role of phase fluctuations on the superfluid density under these assumptions. We found that, while the superfluid density is depleted linearly in the absence of interactions, the inclusion of Coulomb interactions essentially eliminates the longitudinal spin-wave fluctuations responsible for this behaviour, rendering the quasiparticle excitations the dominant contribution at low temperatures. Phase fluctuations are then only relevant

near $T_c$ where vortices control the physics, resulting in a BKT transition and the accompanying Nelson-Kosterlitz jump in stiffness. Using the Ioffe-Larkin rule to combine the effects of both the phase fluctuations and the quasiparticle excitations, we then derived a relation between $\Upsilon(0)$ and $T_c$ which displays a marked deviation from linearity as seen in early experiments, but in quantitative agreement with more recent data.

Our argument differs greatly from another recent attempt to explain the curvature seen in the data[90]. There it was postulated that an additional gap, in competition with superconductivity, opens up in the density of states. The advantage of our argument is that it invokes only very general, phenomenological observations of the underdoped state. In particular, it relies only on the fact that the superconducting gap is increasing as the critical temperature decreases, and on the quasi-two dimensionality.

One drawback of our argument is the prediction that the low-temperature slope of the superfluid density should decrease quite rapidly ($\sim x^2$) with underdoping. This is a common consequence of effective theories of underdoped cuprates[20, 104], but has not been observed experimentally[8, 88]. Indeed, this has been one of the major theoretical obstacles: theories which correctly predict a zero-temperature superfluid density which decreases linearly with underdoping, incorrectly predict a quadratically decreasing slope. It should also be noted that the Nelson-Kosterlitz jump in the superfluid stiffness has not been seen either, except possibly in ultra-thin samples of YBaCuO.

The fact that the experimentally determined slope is only very weakly doping dependent[8], while a theory based on nodal quasiparticles predicts $y \sim x^2$, suggests that phase fluctuations *do* somehow play a role in depleting the superfluid density. We have seen in this Chapter that Coulomb interactions in a strictly two dimensional, or anisotropic three dimensional XY model completely eliminate the spin-wave fluctuations from the problem. To reinstate the importance of phase fluctuations, we need a mechanism whereby interactions can be effectively neglected. In the next Chapter, we'll examine a layered system of Coulomb interacting bosons, and see that the presence of the layers screens the interaction into a short-ranged potential. In the dilute limit, this means that the temperature dependence of the helicity modulus is dominated over a large range by the form of the non-interacting Bose condensate, which is, for all practical purposes, linear. In this way, phase fluctuations become a crucial factor in determining the helicity modulus.

# Chapter 5

# Layered superfluids

In the last Chapter, we saw that the XY model of phase fluctuations, as an effective theory for the depletion of the superfluid density in cuprates, was insufficient to describe what is observed experimentally. This was due to the Coulomb interaction which effectively lifts the energies of the low-lying spin-wave excitations above the plasma energy, modifying the non-interacting linear in temperature behaviour into the $T^5$ power law. Such an argument has been used in the past to criticize the phase fluctuation mechanism of superconductivity[113], leaving only quasiparticles as the relevant contribution in the Ioffe-Larkin rule (4.36).

However, cuprates are strongly anisotropic materials, with the anisotropy between the transport properties in the $ab$-plane[6] and along the $c$-axis[9] increasing with underdoping[102, 114]. We might expect, then, that the plasmon dispersion will become more anisotropic, as well. In fact, this is indeed the case, and it is well known that the large plasmon energy gap becomes replaced by a much smaller one, proportional to the coupling between the two dimensional layers[13, 115, 116]. It is then natural to hope that the plasmon mode will be much less effective at gapping the phase mode, leading to an increased role for phase fluctuations in depleting the superfluid density.

In this Chapter, we will study in greater detail the combined effect of large anisotropy and Coulomb interactions on $\Upsilon(T)$. We will model the phase fluctuation component of the Ioffe-Larkin rule (4.36) by a layered system of bosons, interacting via Coulomb interactions, which provides a rather general representation of a charged, layered superfluid. This work has been published previously in Case and Herbut [43].

## 5.1 The model

The model we will consider is essentially the bosonic theory derived in Chapter 2 (2.49), generalized to arbitrary filling by introducing the chemical potential for the electrons via the shift $A_0 \rightarrow A_0 + ih$. We will neglect the unnecessary complication of two bosonic species, and assume the gauge fields to have been integrated out, giving rise only to renormalized interaction terms. The action for coupled, Coulomb interacting layers can then be written as $S = \int_0^\beta d\tau \int d^2\mathbf{r}\, \mathcal{L}$, where

$$\begin{aligned} \mathcal{L} &= \sum_{i=1}^{N} b_i^*(\mathbf{r},\tau)\left(\partial_\tau - \frac{\hbar^2}{2m}\nabla^2 - \mu\right) b_i(\mathbf{r},\tau) \\ &+ \frac{1}{2}\sum_{i,j}\int d^2\mathbf{r}'|b_i(\mathbf{r},\tau)|^2 V_{ij}(\mathbf{r}-\mathbf{r}')|b_j(\mathbf{r}',\tau)|^2 \\ &- t\sum_i \sum_{j=i\pm 1} b_i^*(\mathbf{r},\tau) b_j(\mathbf{r},\tau), \end{aligned} \tag{5.1}$$

and $h$ has been absorbed into the definition of the fields. Here, the index $i$ runs over the $N$ two dimensional layers and $t$ is a weak Josephson coupling between the layers. The Coulomb interaction is

$$V_{ij}(\mathbf{r}-\mathbf{r}') = e^2/\left(\epsilon\sqrt{|\mathbf{r}-\mathbf{r}'|^2 + |i-j|^2 d^2}\right), \tag{5.2}$$

$d$ being the interlayer separation and $\epsilon$ the static background dielectric constant. We assume the presence of a neutralizing background of density $\rho_0$ equal to the average areal density of bosons at the chemical potential $\mu$.

Of particular note in the above action is the presence of the linear time derivative term, which is known to arise for a general, incommensurate filling of lattice bosons[117]; in the present case, it must be included away from half-filling (i.e. for $x \neq 0$). Fisher et al. [117] have pointed out that the universality class of the model with this term is that of the continuum Bose gas, while, in its absence, the lowest order time derivative term is $\partial_\tau^2$ and the universality class is that of the XY model. This sharp distinction between the commensurate case of the last Chapter, and the incommensurate case of the present (together with the inclusion of the layers), provides the needed mechanism by which phase fluctuations gain importance in the Ioffe-Larkin rule. We begin this discussion by considering the excitations in the layered bosonic system.

## 5.2 The bilayer system

The nature of the excitations in the superfluid system (5.1) is most explicit with only two layers, which differs already from a single layer in a fundamental way. Let us first introduce the usual density-phase variables as

$$b_i(\mathbf{r},\tau) = \sqrt{\rho_0 + \Pi_i(\mathbf{r},\tau)}e^{i\phi_i(\mathbf{r},\tau)}, \tag{5.3}$$

where the $\Pi$ fields represent fluctuations around the average density $\rho_0$, and $\phi$ are the usual phase fluctuations. Expanding the action (5.1) to second order in these fields leads to

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{2}\left(i(\rho_0+\Pi_i)\dot{\phi}_i + \frac{\hbar^2}{8m\rho_0}(\nabla\Pi_i)^2 + \frac{\hbar^2\rho_0}{2m}(\nabla\phi_i)^2\right) \\ &+\frac{1}{2}\sum_{i,j}\int \mathrm{d}^2\mathbf{r}'\Pi_i(\mathbf{r},\tau)V_{ij}(\mathbf{r}-\mathbf{r}')\Pi_j(\mathbf{r}',\tau) \\ &+t\rho_0(\phi_1-\phi_2)^2 + \frac{t}{4\rho_0}(\Pi_1-\Pi_2)^2.\end{aligned} \tag{5.4}$$

The simple transformation to variables $\phi_\pm = (\phi_1 \pm \phi_2)/\sqrt{2}$, $\Pi_\pm = (\Pi_1 \pm \Pi_2)/\sqrt{2}$ diagonalizes the action, which can be written as the sum $\mathcal{L} = \mathcal{L}_+ + \mathcal{L}_-$, where

$$\begin{aligned}\mathcal{L}_+ &= i(\sqrt{2}\rho_0+\Pi_+)\dot{\phi}_+ + \frac{\hbar^2}{8m\rho_0}(\nabla\Pi_+)^2 + \frac{\hbar^2\rho_0}{2m}(\nabla\phi_+)^2 \\ &+\frac{1}{2}\int \mathrm{d}^2\mathbf{r}'\Pi_+(\mathbf{r},\tau)V_+(\mathbf{r}-\mathbf{r}')\Pi_+(\mathbf{r}',\tau)\end{aligned} \tag{5.5}$$

$$\begin{aligned}\mathcal{L}_- &= i\Pi_-\dot{\phi}_- + \frac{\hbar^2}{8m\rho_0}(\nabla\Pi_-)^2 + \frac{\hbar^2\rho_0}{2m}(\nabla\phi_-)^2 \\ &+\frac{1}{2}\int \mathrm{d}^2\mathbf{r}'\Pi_-(\mathbf{r},\tau)V_-(\mathbf{r}-\mathbf{r}')\Pi_-(\mathbf{r}',\tau) \\ &+2t\rho_0\phi_-^2 + \frac{t}{2\rho_0}\Pi_-^2,\end{aligned} \tag{5.6}$$

where $V_\pm(\mathbf{r}) = V_{11}(\mathbf{r}) \pm V_{12}(\mathbf{r})$. The Gaussian integration over $\Pi_\pm$ is now straightforwardly done. After Fourier transforming, we find

$$\mathcal{L}_+ \rightarrow \left(\frac{1}{2}\frac{\omega^2}{V_+(k)+\frac{\hbar^2k^2}{4m\rho_0}} + \frac{\rho_0\hbar^2k^2}{2m}\right)\phi_+^2(\mathbf{k},\omega) \tag{5.7}$$

$$\mathcal{L}_- \rightarrow \left(\frac{1}{2}\frac{\omega^2}{V_-(k)+\frac{\hbar^2k^2}{4m\rho_0}+\frac{t}{\rho_0}} + \frac{\rho_0\hbar^2k^2}{2m} + 2t\rho_0\right)\phi_-^2(\mathbf{k},\omega), \tag{5.8}$$

leading to two branches of excitations with energies,

$$\omega_+ = \frac{\hbar^2 k^2}{2m}\left(2\rho_0 V_+(k) + \frac{\hbar^2 k^2}{2m}\right) \tag{5.9}$$

$$\omega_- = \left(\frac{\hbar^2 k^2}{2m} + 2t\right)\left(2\rho_0 V_-(k) + \frac{\hbar^2 k^2}{2m} + 2t\right), \tag{5.10}$$

with the Fourier transform of the interaction,

$$V_\pm(k) = \frac{2\pi e^2}{\epsilon k}\left(1 \pm e^{kd}\right). \tag{5.11}$$

Note that the low momentum behaviour of $V_+$ and $V_-$ differ drastically. Whereas $V_+ \sim 1/k$ has the usual long-range Coulomb form, $V_- \sim 1$ represents a *screened*, short-range interaction. As a result, the branch $\omega_+$ describes the usual two dimensional plasmon, $\omega_+ \approx \sqrt{4\pi\hbar^2 e^2 \rho_0 k/\epsilon m}$, while, for $t = 0$, the other branch has the linear dispersion $\omega_- \approx (2\pi\hbar^2 e^2 \rho_0/\epsilon m)^{\frac{1}{2}} k$. In terms of the field theory, the upper mode describes the two layers oscillating in phase with one another at no cost in Josephson energy; however, being that the total density fluctuation $\Pi_+$ is the canonically conjugate variable to $\phi_+$, these oscillations represent true charge density (plasma) oscillations and, thus, pay the respective cost in Coulomb energy. On the other hand, the out of phase mode described by $\omega_-$ is subject to the Josephson energy, but the oscillations may occur so as to perfectly screen the Coulomb interaction in one plane by the other, since the canonically conjugate variables are $\phi_-$ and the density difference, $\Pi_-$.

The remarkable feature of the above result is that, in a system with negligible Josephson coupling, the Coulomb interaction becomes effectively short-ranged, as far as the low-energy excitation spectrum is concerned. More precisely, when $t = 0$, $\omega_-$ deviates significantly from $\omega_+$ when $k \ll 1/d$. For a large separation between layers, and with $1/d \ll k \ll (8\pi m e^2 \rho_0/\hbar^2\epsilon)^{\frac{1}{3}}$, $\omega_- \approx \omega_+ \sim \sqrt{k}$. Finally with $(8\pi m e^2 \rho_0/\hbar^2\epsilon)^{\frac{1}{3}} \ll k$, $\omega_- \approx \omega_+ \approx \hbar^2 k^2/2m$. If, however, interactions were weak or the layers were brought close together so that $1/d \gg (8\pi m e^2 \rho_0/\hbar^2\epsilon)^{\frac{1}{3}}$, the dispersion $\omega_-$ would change directly from linear to quadratic, without the intervening $\sqrt{k}$ region. In this regime, $\omega_-$ becomes identical to the phonon spectrum of the weakly interacting Bose gas. This is illustrated in Figure 5.1. With a finite Josephson coupling, $\omega_-$ approaches the Josephson gap for $k \ll \sqrt{4mt}$. If $t$ is large, of course, $\omega_- \gg \omega_+$ and the plasmon would resume its place as the low-energy mode of the system.

Having elucidated the nature of the low-energy excitations in the two layer system, we will now turn to the many layer system, as is relevant to the cuprates. In particular, we will be interested in how the low energy modes affect the depletion of the helicity modulus.

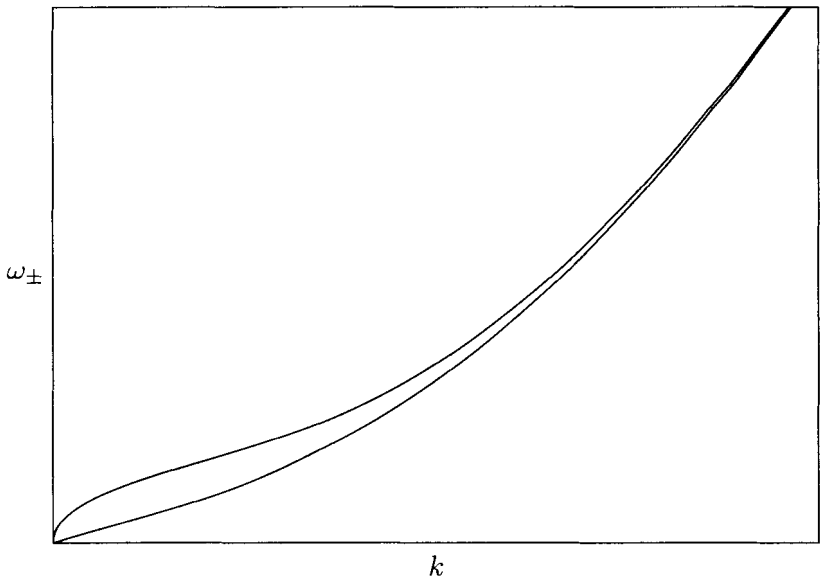


Figure 5.1: The two branches $\omega_\pm$ of the excitation spectrum of the two-layer system, with $t = 0$ and weak Coulomb interaction. The lowest mode crosses over from linear behaviour at low $k$ to $k^2$ behaviour. The plasmon starts out as $\sqrt{k}$ before crossing to $k^2$.

## 5.3 The many layer system

For a system with $N$ layers, the situation is analogous to the bilayer system. There is still just one plasmon mode, but there are now $N-1$ lower energy modes with linear dispersion. To see this, consider the interaction matrix

$$\hat{V}(k) = \frac{2\pi e^2}{\epsilon k} \begin{bmatrix} 1 & e^{-kd} & e^{-2kd} & \dots & e^{-Nkd} \\ e^{-kd} & 1 & e^{-kd} & & \\ e^{-2kd} & e^{-kd} & 1 & & \\ \vdots & & & \ddots & \\ e^{-Nkd} & & & & 1 \end{bmatrix}. \tag{5.12}$$

When $k \to 0$, $V_{ij}(k) = (2\pi e^2/\epsilon k)\,(1 + \mathcal{O}(kd))$, so in this limit

$$\hat{V}(k) \to \frac{2\pi e^2}{\epsilon k}(1, 1, \dots, 1)^T \otimes (1, 1, \dots, 1). \tag{5.13}$$

This matrix has one eigenvector with eigenvalue $2\pi e^2/\epsilon k$, and $N-1$ with eigenvalue zero. The former eigenvector is just the total density which is canonically conjugate to the sum of the phases

and describes the plasmon. The latter $N-1$ modes, being orthogonal to the plasmon, are electrically neutral, and consequently cross over from linear dispersion at low momenta to the Josephson gap at $k = 0$.

We can determine the exact form of the dispersion in the physically relevant case of infinitely many layers. Returning to the Lagrangian (5.4) in the density-phase representation, imposing periodic boundary conditions in the $z$ direction, we find, after a Fourier transformation,

$$\begin{aligned}\mathcal{L} &= \omega\Pi(K)\phi(-K) + \frac{\hbar^2k^2}{8m\rho_0}\Pi(K)\Pi(-K) + \frac{\rho_0\hbar^2k^2}{2m}\phi(K)\phi(-K) \\ &+\frac{1}{2}\Pi(K)V(\mathbf{k},k_z)\Pi(K) + 2t\rho_0\sin^2\left(\frac{k_zd}{2}\right)\phi(K)\phi(-K) \\ &+\frac{t}{2\rho_0}\sin^2\left(\frac{k_zd}{2}\right)\Pi(K)\Pi(-K),\end{aligned} \tag{5.14}$$

with the short-hand $K \equiv (\omega, \mathbf{k}, k_z)$. Now integrating out density fluctuations leads to the excitation spectrum

$$\omega^2(\mathbf{k},k_z) = e(\mathbf{k},k_z)\left(2\rho_0V(\mathbf{k},k_z) + e(\mathbf{k},k_z)\right), \tag{5.15}$$

where $e(\mathbf{k},k_z) = \hbar^2k^2/2m + 2t\sin^2(k_zd/2)$. The Fourier transform of the interaction in the layered system has been calculated previously[118] to be

$$V(\mathbf{k},k_z) = \frac{2\pi e^2}{\epsilon k}\frac{\sinh(kd)}{\cosh(kd)-\cos(k_zd)}. \tag{5.16}$$

The plasmon mode corresponds to $k_z = 0$, and from the spectrum (5.15), we find that it has the usual gapped, three dimensional form $\omega^2(0,0) = \omega_{\rm p}^2 = 4\pi\hbar^2e^2\rho_0/\epsilon md$. The remaining modes with $k_z \neq 0$ have the linear dispersion $\omega(\mathbf{k},k_z) = \omega_{\rm p}kd/2\sin(k_zd/2)$ for $k \ll 1/d$ and with $t = 0$. These become gapped when $t \neq 0$, $\omega(\mathbf{k}\to 0,k_z) \to \omega_{\rm p}\sqrt{d^2mt/\hbar^2}$.

## 5.4 The dilute boson limit

From what we have seen, it seems that the Coulomb interactions are irrelevant at low energies in the layered bosonic systems. Unfortunately, though, the screening due to the other layers merely replaces this long-range potential by the short-range one, and we might still expect the depletion of the superfluid density due to phase fluctuations $\sim T^3$ (as determined in the last Chapter) to be negligible compared to the quasiparticle contribution. While this reasoning is, in principle, correct, it is known that, in the dilute limit[92, 119] and for an incommensurate filling, a temperature region

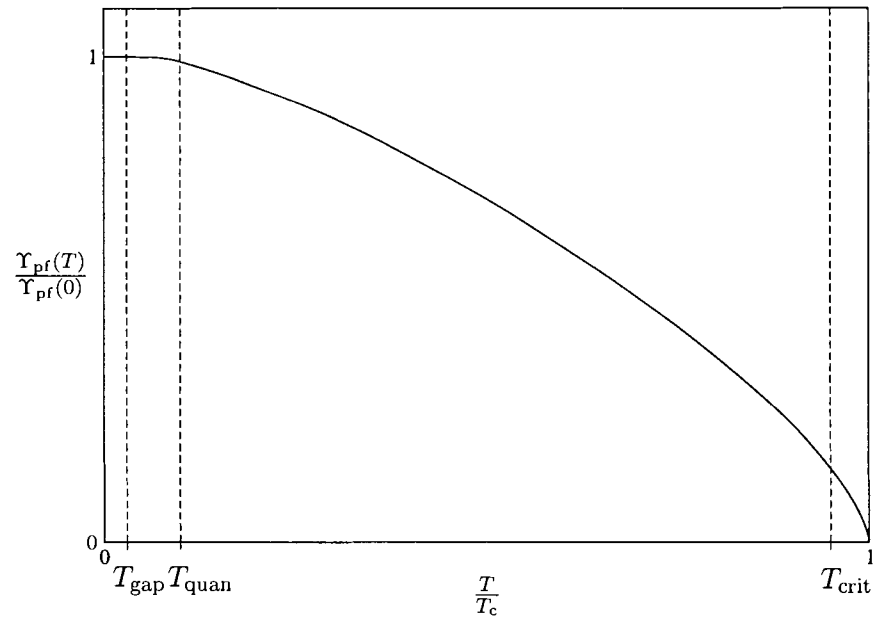


Figure 5.2: Schematic depiction of the helicity modulus in a dilute superfluid. Interacting behaviour occurs below $T_{\rm quan}$, while classical critical fluctuations determine the behaviour above $T_{\rm crit}$. At low enough densities, the system displays ideal behaviour over an appreciable temperature range $T_{\rm quan} < T < T_{\rm crit}$. At extremely low temperatures $T < T_{\rm gap}$, $\Upsilon_{\rm pf}(T) - \Upsilon_{\rm pf}(0)$ is suppressed exponentially.

over which the system behaves as an ideal Bose gas appears, and the regions of interacting critical behaviour near $T = 0$ and $T = T_{\rm c}$ shrink. This is depicted schematically in Figure 5.2.

As a concrete illustration of this, consider a dilute system of bosons in three dimensions, interacting via a short-range interaction of strength $\lambda$. From dimensional analysis, we can write $\Upsilon_{\rm pf}(T) = \Upsilon_{\rm pf}(0)\mathcal{Y}(\lambda/r, Tr^2)$, where $r$ is the average interparticle distance. Evidently, the dilute limit $r \to \infty$ is equivalent to the weak interaction, high temperature regime, and the temperature dependence has the ideal Bose condensate form $\mathcal{Y} \approx 1 - (T/T_{\rm BEC})^{\frac{3}{2}}$, where $T_{\rm BEC}$ is the Bose-Einstein condensate temperature. At low enough temperatures, interactions become important and $\mathcal{Y} \approx 1 - (T/T_{\rm c})^4$; this transition happens at $T_{\rm quan} \sim (\lambda/r)T_{\rm BEC}$. Similarly, the width of the classical critical region is $T_{\rm c} - T_{\rm crit} \sim (\lambda/r)^2 T_{\rm BEC}$. We see that both of these regions shrink much faster than $T_{\rm c} \approx T_{\rm BEC}$ with dilution, as the helicity modulus behaves more and more like the condensate of the ideal Bose gas.

In the language of the renormalization group, the above argument is a result of the irrelevancy

of short-range interactions in three dimensions at the quantum critical point. We can see this by constructing the beta-function for the dimensionless interaction $\hat{\lambda} = \lambda/r$,

$$r\frac{\partial\hat{\lambda}}{\partial r} = -\hat{\lambda}, \tag{5.17}$$

which demonstrates that the flow is towards zero interaction. It is interesting, also, to consider the two dimensional case[119]. The interaction is already dimensionless, and by simple power counting, is marginal. Actually, interactions turn out to the marginally irrelevant, so the above conclusions still hold, except that the interacting critical regions vanish logarithmically slowly with dilution and it is, thus, much harder to reach the dilute limit.

The foregoing discussion highlights the importance of a detailed analysis of the bosonic contribution to the helicity modulus. Applying this to the cuprates gives a greater role to phase fluctuations in determining the temperature dependence of $\Upsilon(T)$, as we will see in the remainder of this Chapter.

## 5.5 The helicity modulus in the layered system

We will derive the helicity modulus using our first definition (4.3)

$$\left\langle \hat{P} \right\rangle = V\rho_{\rm n}\mathbf{v}, \tag{5.18}$$

where $\rho_{\rm n}$ is the normal fluid mass density. The net momentum flux in the plane is

$$\begin{aligned} \left\langle \hat{P}_{\parallel} \right\rangle &= \sum_{\mathbf{k},k_z} \hbar\mathbf{k} n_{\rm b}\Big(\omega(\mathbf{k},k_z) - \hbar\mathbf{k}\cdot\mathbf{v}\Big) \\ &= \frac{\hbar^2\mathbf{v}}{2}\sum_{\mathbf{k},k_z} k^2\left(-\frac{\partial n_{\rm b}(\omega(\mathbf{k},k_z))}{\partial\omega(\mathbf{k},k_z)}\right), \end{aligned} \tag{5.19}$$

where $n_{\rm b}(\omega) = (e^{\omega(\mathbf{k},k_z)/T} - 1)^{-1}$ is the usual boson occupation number, and the last line follows from expanding in $\mathbf{v}$ and performing an angular average over $\mathbf{k}$. This, of course, yields Landau's famous formula for the normal fluid density, which we convert to the helicity modulus using (4.7),

$$\Upsilon_{\rm pf}(T) = \Upsilon_{\rm pf}(0) - \frac{\hbar^4 d}{2m^2}\int\frac{{\rm d}^2\mathbf{k}}{(2\pi)^2}\int_{-\pi/d}^{\pi/d}\frac{{\rm d}k_z}{2\pi}\,k^2\left(-\frac{\partial n_{\rm b}(\omega(\mathbf{k},k_z))}{\partial\omega(\mathbf{k},k_z)}\right); \tag{5.20}$$

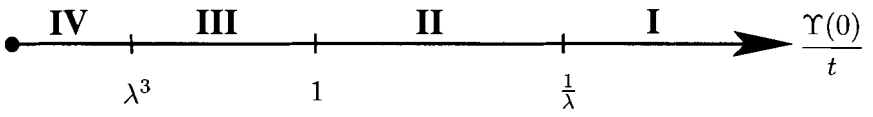


Figure 5.3: Four regimes for the temperature dependence of the helicity modulus $\Upsilon_{\rm pf}(T)$ in a layered bosonic system with weak Coulomb interactions: I) quasi-two-dimensional (2D) regime with weak and effectively short-range (screened) interaction, II) quasi-2D regime with weak long-range interaction, III) three-dimensional (3D) regime with weak long-range interaction, and IV) 3D regime with strong long-range interaction. $t$ is the inter-layer Josephson coupling, $\lambda = (2\pi e^2/(\epsilon d))/(\hbar^2/(md^2)) \ll 1$ is the dimensionless strength of the Coulomb interaction, and $d$ is the inter-layer separation.

we have also taken the continuum limit. Performing the derivative, and rewriting the result in terms of dimensionless quantities, this becomes

$$\widetilde{\Upsilon}_{\rm pf}(T) = \widetilde{\Upsilon}_{\rm pf}(0) - \frac{\widetilde{T}}{8\pi}\int_0^\infty y\mathrm{d}y \int_0^1 \mathrm{d}z\ \sinh^{-2}\left\{\frac{1}{2}\left[f(y,z) \left(f(y,z) + \frac{2\lambda\widetilde{\Upsilon}_{\rm pf}(0)\sinh\sqrt{2y\widetilde{T}}}{\widetilde{T}^{\frac{3}{2}}\sqrt{2y}(\cosh\sqrt{2y\widetilde{T}} - \cos\pi z)}\right)\right]^{\frac{1}{2}}\right\}. \tag{5.21}$$

We have defined $f(y,z) = y + (2\widetilde{t}/\widetilde{T})\sin^2(\pi z/2)$ and dimensionless parameters $\widetilde{X} \equiv X/T_{\rm d}$ where the characteristic energy scale is $T_{\rm d} = \hbar^2/md^2$; the dimensionless interaction is $\lambda = 2\pi e^2/\epsilon d T_{\rm d}$. This expression is expected to be valid for $\lambda \ll 1$ and not too close to $T_{\rm c}$ where critical fluctuations set in.

We can now use this expression (5.21) to determine the temperature dependence of the helicity modulus for various values of the parameters. We will assume $t/\Upsilon_{\rm pf}(0) \ll 1$ as is relevant to the underdoped cuprates, and consider $\Upsilon_{\rm pf}(T)$ as $\Upsilon_{\rm pf}(0)$ is lowered at fixed $t$, roughly corresponding to underdoping. We will see that this leads to four distinct regimes of temperature dependence, summarized in Figure 5.3.

### I: The 2D regime with weak, short-range interactions

For a weak Josephson coupling $\Upsilon_{\rm pf}(0)/t \gg 1$, the system is quasi-two dimensional. Also assuming $\lambda \ll 1$, $\Upsilon_{\rm pf}(T)$ takes on the non-interacting form,

$$\Upsilon_{\rm pf}(T) = \Upsilon_{\rm pf}(0) - \frac{T}{2\pi}\left[\ln\left(\frac{T}{t}\right) + 1.386 + \mathcal{O}(t/T)\right], \tag{5.22}$$

for high temperatures (below $T_{\rm c}$). Deviations from this are significant within the critical region of width $T^{\rm I}_{\rm crit} = \lambda T_{\rm c}$, where the behaviour is $\sim (T_{\rm c} - T)^{\nu}$; the interacting behaviour (4.54), $\Delta\Upsilon_{\rm pf}(T) \sim T^3/\Upsilon^2_{\rm pf}(0)\lambda^2$, becomes significant below $T^{\rm I}_{\rm quan} = \Upsilon_{\rm pf}(0)\lambda$. Due to the presence of the Josephson gap, there exists another low temperature scale $T^{\rm I}_{\rm gap} = \sqrt{2\lambda\Upsilon_{\rm pf}(0)t}$ below which $\Delta\Upsilon_{\rm pf}(T)$ is exponentially suppressed. Estimating the critical temperature from (5.22) to be $T_{\rm c} \approx 2\pi\Upsilon_{\rm pf}(0)/\ln(2\pi\Upsilon_{\rm pf}(0)/t)$, the relative sizes of these low temperature regions are

$$\frac{T^{\rm I}_{\rm quan}}{T_{\rm c}} = \frac{\lambda}{2\pi}\ln\left(\frac{2\pi\Upsilon_{\rm pf}(0)}{t}\right) \tag{5.23}$$

and

$$\frac{T^{\rm I}_{\rm gap}}{T_{\rm c}} = \frac{1}{\pi}\sqrt{\frac{\lambda t}{2\Upsilon_{\rm pf}(0)}}\ln\left(\frac{2\pi\Upsilon_{\rm pf}(0)}{t}\right), \tag{5.24}$$

while $(T_{\rm c} - T^{I}_{\rm crit})/T_{\rm c} = \lambda$. When these temperature scales are well separated,

$$T^{\rm I}_{\rm gap} \ll T^{\rm I}_{\rm quan} \ll T^{\rm I}_{\rm crit}, \tag{5.25}$$

there exists a wide range over which the non-interacting form of the helicity modulus appears. This is satisfied for

$$\frac{t}{\Upsilon_{\rm pf}(0)} \ll \lambda \ll \frac{2\pi}{\ln(2\pi\Upsilon_{\rm pf}(0)/t)}; \tag{5.26}$$

such an interval for $\lambda$ indeed exists for $\Upsilon_{\rm pf}(0)/t > 1$.

To summarize, in the regime with $\Upsilon_{\rm pf}(0)/t \gg 1/\lambda$, marked I in Figure 5.3, the long-range nature of the Coulomb interaction is irrelevant, except at temperatures below $T^{\rm I}_{\rm gap}$. The interactions instead appear effectively short-range, and weak. As a result, quasi-two dimensional non-interacting behaviour can be observed over a wide temperature range. The size of the low temperature interacting region shrinks with underdoping (reducing $\Upsilon_{\rm pf}(0)$), though logarithmically slowly. Again, we can interpret this as the renormalization group flow of the interaction strength to zero, with diluting.

### II: The 2D regime with weak, long-range interactions

As we further decrease $\Upsilon_{\rm pf}(0)$, eventually we reach $1/\lambda \approx \Upsilon_{\rm pf}(0)/t$ where the left hand side of the inequality (5.26) ceases to hold. We are still in the quasi-two dimensional regime since $\Upsilon_{\rm pf}(0)/t \gg 1$, however the two low temperature regions coalesce, $T^{\rm I}_{\rm gap} \approx T^{\rm I}_{\rm quan}$, and the non-interacting behaviour (5.22), therefore, crosses directly over to exponential, with no intermediate short-range behaviour. The interactions, thus, appear effectively long-range in the quantum critical region.

### III: The 3D regime with weak, long-range interactions

When $\Upsilon_{\rm pf}(0)/t \approx 1$, the system reasserts its full three dimensionality, and the non-interacting form of the helicity modulus is

$$\Upsilon_{\rm pf}(T) = \Upsilon_{\rm pf}(0) - \frac{\zeta\left(\frac{3}{2}\right)}{2\pi\sqrt{2\pi t}} T^{\frac{3}{2}}, \tag{5.27}$$

to lowest order in $\lambda$. By contrast, the interacting form has the famous $\sim T^4$ behaviour which sets in below $T^{\rm III}_{\rm quan} = \Upsilon_{\rm pf}(0)\lambda$, while the exponential suppression sets in below $T^{\rm III}_{\rm gap} = \sqrt{2\lambda\Upsilon_{\rm pf}(0)t}$. From (5.27), we estimate $T_{\rm c} \approx (8\pi\Upsilon_{\rm pf}(0)\sqrt{2\pi t}/\zeta(3/2))^{2/3}$ and find the relative sizes of the low temperature regions to be

$$\frac{T^{\rm III}_{\rm quan}}{T_{\rm c}} \propto \lambda \left(\frac{\Upsilon_{\rm pf}(0)}{t}\right)^{\frac{1}{3}} \tag{5.28}$$

and

$$\frac{T^{\rm III}_{\rm gap}}{T_{\rm c}} \propto \sqrt{\lambda} \left(\frac{t}{\Upsilon_{\rm pf}(0)}\right)^{\frac{1}{6}}. \tag{5.29}$$

Also, the size of the classical critical region is again $\lambda$. Evidently, $T^{\rm III}_{\rm quan} < T^{\rm III}_{\rm gap}$ for $\Upsilon_{\rm pf}(0)/t < 1/\lambda$ and $1/\lambda \gg 1$, so when $\Upsilon_{\rm pf}(0)/t \approx 1$, the non-interacting behaviour crosses directly over to the gapped exponential form at low temperatures, without the intervening three dimensional interacting behaviour. The interacting region is, therefore, effectively long-ranged, but the strength of interactions is still effectively weak.

### IV: The 3D regime with strong, long-range interactions

From (5.29), we see that the Josephson gap becomes comparable to $T_{\rm c}$ when

$$\frac{\Upsilon_{\rm pf}(0)}{t} \approx \lambda^3. \tag{5.30}$$

The exponential form of the helicity modulus then dominates the entire temperature range, and the system appears to have strong, long-range interactions. In electronic systems, strong Coulomb interactions are known to lead to a Wigner crystal, which is what we would also expect in this bosonic case, due to the very low density.

These four regimes of temperature dependence are summarized in Figure 5.3. To apply this framework to the underdoped cuprates, we must now try to estimate the values of the relevant parameters, $\Upsilon_{\rm pf}(0)/t$ and $\lambda$.

## 5.6 Application to underdoped cuprates

The superfluid density at zero temperature $\rho_{\rm s}(0)$ is given by the bare density $\rho_0$ only at zero interaction. Otherwise, it is depleted from this value, so we can write

$$\Upsilon_{\rm pf}(0) < \Upsilon(0) = \frac{\hbar^2}{m}\rho_0, \tag{5.31}$$

to relate $\Upsilon_{\rm pf}(0)$ to the unknown mass $m$. The bare density, from our effective theory and also from gauge theories of the $t-J$ model, is just the density of doped holes, so $\rho_0 = x/a^2$, where $a$ is the planar lattice constant. The mass can then be written as

$$m < \frac{\hbar^2 x}{\Upsilon_{\rm pf}(0)a^2}, \tag{5.32}$$

so that the characteristic temperature is subject to the constraint

$$T_{\rm d} > \left(\frac{a}{d}\right)^2 \frac{\Upsilon_{\rm pf}(0)}{x}. \tag{5.33}$$

We can use the Ioffe-Larkin rule (4.36) to estimate $\Upsilon_{\rm pf}(0)$ from the experimentally relevant $\Upsilon(0)$,

$$\Upsilon(0) = \frac{\Upsilon_{\rm pf}(0)\Upsilon_{\rm qp}(0)}{\Upsilon_{\rm pf}(0) + \Upsilon_{\rm qp}(0)}, \tag{5.34}$$

where the quasiparticle and phase fluctuation contributions are related via the charge renormalization factor

$$z = \frac{\Upsilon_{\rm pf}(0)}{\Upsilon_{\rm pf}(0) + \Upsilon_{\rm qp}(0)}. \tag{5.35}$$

This leads to the relation

$$\Upsilon_{\rm pf}(0) = \frac{\Upsilon(0)}{1-z}, \tag{5.36}$$

where $z \approx 0.8$ at $x = 0.1$[22]. Also at this doping in YBaCuO, the zero temperature penetration depth is approximately 50 $\mu\text{m}^{-2}$[11] which we convert to Kelvin units using our expression (4.22) with $d = 12$ Å, so that $\Upsilon_{\text{pf}}(0) \approx 2000$ K. Using this, along with the lattice anisotropy $(a/d)^2 = 0.1$, we find

$$T_{\text{d}} > 2000 \text{ K}. \tag{5.37}$$

The anisotropy between planar and $c$-axis transport properties, as seen in, for example, penetration depth measurements, is characterized by the ratio

$$\frac{\Upsilon_c(0)}{\Upsilon_{ab}(0)} = \frac{t}{2T_{\text{d}}}, \tag{5.38}$$

which is estimated to be approximately $10^{-4}$ in YBCuO and independent of doping[9]. Such a large value is responsible for the approximately two dimensional behaviour seen in the cuprates. We can now estimate the ratio

$$\frac{\Upsilon_{\text{pf}}(0)}{t} \approx 10^4. \tag{5.39}$$

Finally, the interaction $\lambda$ requires knowledge of $\epsilon$ which is measured in experiments on the dielectric constant to be $\epsilon \approx 30$ and doping independent[120]. Using this in the definition of the interaction strength, we find

$$\lambda = \frac{2\pi e^2}{\epsilon d T_{\text{d}}} \approx 1.5. \tag{5.40}$$

Due to interactions, $\Upsilon_{\text{pf}}(0)$ in reality could be well below the non-interacting value $\hbar^2\rho_0/m$ used to establish this upper bound. Additionally, the estimate of $\lambda$ is very sensitive to the precise value of $z$, which, if much larger, would significantly decrease this estimate. We consider (5.40) to be a very conservative upper bound on the strength of the interaction, and believe the actual value most likely to lie much below it.

Comparing these experimental values with the inequalities derived in the last section, we see that the left hand inequality (5.26) $\lambda \ll \Upsilon_{\text{pf}}(0)/t$ is very well satisfied for the doping $x = 0.1$ used to establish these values. Even taking into account that, with underdoping, the value of $\Upsilon(0)$ decreases by about two orders of magnitude from optimally doped YBaCuO with $T_{\text{c}} = 93$ K to underdoped samples with $T_{\text{c}} \approx 5$ K[8], the above inequality still safely holds in the highly underdoped cuprates. This indicates that the high temperature superconductors are well in the two dimensional, short-ranged regime, though with a not quite weak interaction strength, since $\lambda \sim 2\pi/\ln(2\pi\Upsilon_{\text{pf}}(0)/t)$. With such an intermediate coupling $\lambda \approx 1$, the regimes II, III and IV overlap significantly, indicating

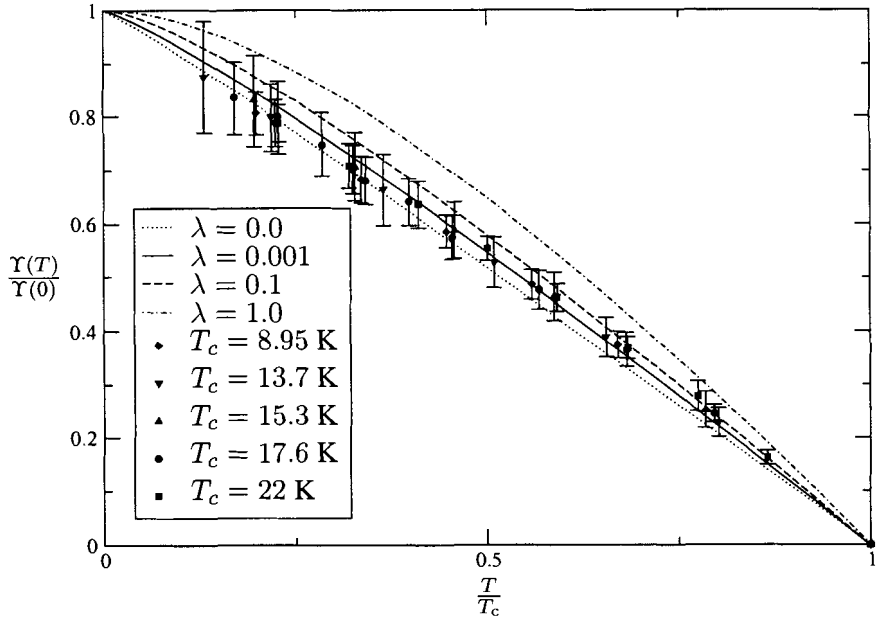


Figure 5.4: The scaled helicity modulus $\Upsilon_{\rm pf}(T)/\Upsilon_{\rm pf}(0)$ as a function of $T/T_c$ for various values of the dimensionless interaction $\lambda$. The value relevant to the cuprates is estimated to be $\lambda \sim 1$. The experimental data are taken from lower critical field data and scaled as in Liang et al. [8].

that the crossover from 2D short-range behaviour goes directly to the 3D long-range regime as we dilute the system.

The effect of a moderate interaction strength on the helicity modulus is shown in Figure 5.4, where we plot $\Upsilon_{\rm pf}(T)$ for various values of $\lambda$. We can see that the approximately linear, non-interacting behaviour becomes slightly more curved as we increase the coupling, though changing this slightly from the estimated value $\lambda \approx 1$ has little effect. Strictly speaking, the helicity modulus obeys the simple scaling used in Figure 5.4 only approximately, though residual dependence on $\Upsilon_{\rm pf}(0)$ at fixed $\lambda$ is extremely weak. Each curve in this Figure, therefore, accurately represents a family of curves for widely different values of $\Upsilon_{\rm pf}(0)$. Absent from the Figure is the critical region of width $\lambda T_{\rm c}$ around $T_{\rm c}$ with temperature dependence $(T - T_{\rm c})^{\nu}$, $\nu \approx 0.67$; this adds further rounding to the curve at high temperatures.

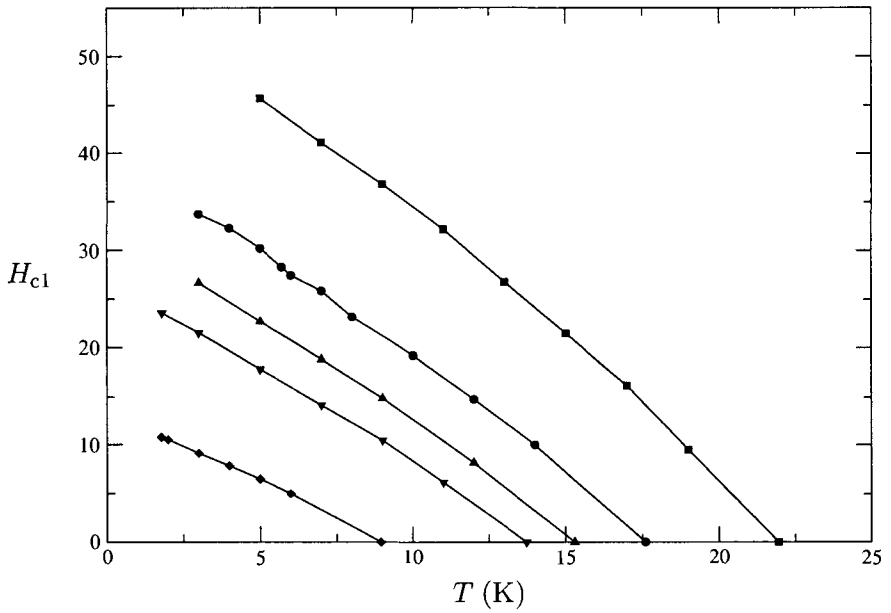


Figure 5.5: Experimental data on $H_{c1}$ measurements, taken from [8]. Symbols are as is Figure 5.4.

## 5.7 Experimental ramifications

Only recently have reliable, high precision measurements of the superfluid density in very underdoped cuprates been performed[8], due in large part to an experimental breakthrough in the production of clean, low $T_c$ samples. The results have been quite unexpected and produced much debate as to their interpretation. The situation is displayed in Figure 5.5, taken from [8]. The measurements are of the lower critical field at various underdopings, which is related to the penetration depth via

$$H_{c1}(T) = \frac{\Phi_0[\ln\kappa + 0.5]}{4\pi\lambda^2(T)} \tag{5.41}$$

or, using our conversion formula (4.22) for the helicity modulus,

$$\Upsilon(T) = \frac{6.2 \times 4\pi}{20.7[\ln\kappa + 0.5]} \left(\frac{d}{10\ \text{Å}}\right) \left(\frac{H_{c1}}{\text{Gauss}}\right). \tag{5.42}$$

The quantity $\kappa$ is the dimensionless Ginzburg-Landau parameter which differentiates between type-I ($\kappa < 1/\sqrt{2}$) and type-II superconductors ($\kappa > 1/\sqrt{2}$). In YBaCuO, $\kappa$ varies only by about a factor of two from optimally doped to very underdoped samples, with $\kappa \approx 55$[8].

The main features of the data can be summarized as follows:

1) approximately linear temperature dependence over almost the entire temperature range;

2) the slope decreasing very little, if at all, with underdoping;

3) the absence of a 3D critical region around $T_c$.

The combination of the first two points defines what has been the major crux in our theoretical understanding of the underdoped cuprates. As we know, a linear temperature dependence is characteristic of the quasiparticle contribution to the helicity modulus with phase fluctuations neglected, but theories which correctly predict this behaviour also predict a highly doping dependent slope, which is clearly not seen. This is equally true of the theory of a phase fluctuating $d$-wave superconductor presented in this thesis: the charge renormalization factor $z^2 \propto \Upsilon_{\rm pf}^2(0) \sim x^2$. The results of this Chapter, however, provide a resolution to this conflict.

We have shown that the Coulomb interactions become effectively short-ranged and are significant only within an ever shrinking region near zero temperature. Over a wide temperature range, the decrease of the helicity modulus is the approximately linear form of the ideal Bose condensate and, thus, the phase fluctuation contribution is on par with that of quasiparticles. The important difference between phase fluctuations and quasiparticles, however, is that the 'slope' due to the former is unaffected by doping while that of the latter is *decreasing*. Therefore, the phase fluctuation contribution becomes more and more dominant as we approach the underdoped critical point. This is clear from the Ioffe-Larkin rule with both fermionic (4.44) and bosonic (5.22) contributions included,

$$\Upsilon(T) \approx \Upsilon(0) - \left[\frac{1}{2\pi}\ln\left(\frac{T}{t}\right) + \frac{1.3826}{2\pi} + \frac{2\ln 2}{\pi}\frac{v_f}{v_\Delta}z^2\right]T, \tag{5.43}$$

to lowest order in interactions and temperature. Recalling that the charge renormalization factor is proportional to doping, the last term in brackets is completely negligible in the extremely underdoped limit. Recently, it was shown[121] that the non-interacting, purely bosonic superfluid density provides an excellent fit to the experimental data. Indeed, the present work can be thought of as the justification for neglecting Coulomb interactions in that article, and the fact that the non-interacting superfluid density describes the data so well bolsters our argument that the true effective interaction strength is much less than the estimated $\lambda \sim 1$. The effect of turning on a small $\lambda$ should only add slight curvature to the fit, making the agreement better. To make this more precise, in Figure 5.4, we have displayed the scaled lower critical field data (as in Liang et al. [8]). We see that the best fit is for $\lambda = 10^{-3}$, although it is clear that any $\lambda < 0.1$ would be almost equally good.

The final point concerns the critical region. In optimally doped samples[32], the width of the critical region is about 10 K or 10%. We might then expect a similarly observable effect in the underdoped samples, of the order of 1 K. However, this is not the case, and in fact, the superfluid

density appears to go to $T_c$ linearly, which happens to be the mean-field behaviour ($\nu = 1$)[30]. This fact can be nicely explained within the present theory, however. As was pointed out earlier, the size of the critical region shrinks with dilution, or with underdoping. For example, in three dimensions, $|T_c - T|/T_c \sim (m\lambda/r)^2$ from the Ginzburg criterion. In two dimensions, interactions are weakly irrelevant, and so the critical region width $\sim \lambda$ should be modified to $\lambda/\ln\ln(r)$[119]. Therefore, as we dilute the system, the critical region slowly vanishes, becoming undetectable in very underdoped samples. In fact, that no critical region is seen at all, even in 20 K materials, is further evidence that the effective interaction strength is much smaller than our predicted value.

# Chapter 6

# Dirty Bosons

This final Chapter contains work previously published in Case and Herbut [44] (see also Case and Herbut [122]) and is somewhat out of line with the previous discussions, not having direct bearing on the underdoped cuprates, and containing work completed before the formulation of the framework presented in Chapter 2. However, as this research applies generally to bosonic systems in which both interactions and disorder play a role, we may hope that our results can be brought to bear on the physically relevant issue of disordered high temperature superconductors. With this in mind, we will proceed by introducing the problem, highlighting some reasons for a general interest in dirty bosons. A discussion of disorder and the underdoped cuprates will wait until the final section, after our results are presented.

## 6.1 Introduction

To motivate an interest in disordered bosonic systems, it is useful to note some facts about their electronic counterparts. It has long been appreciated that non-interacting electrons have all their single-particle states localized by disorder in dimensions two and less[123]. With the addition of interactions, it is less clear whether a metallic state can occur, but if the interactions are attractive, we might expect a superconductor, at least in two dimensions. This follows from the criterion first derived by Harris[124], that weak disorder is irrelevant if the exponents of the pure system in its absence satisfy $\nu D > 2$. In dimension $D = 2$, the pure system undergoes a Berezinskii-Kosterlitz-Thouless (BKT) transition for which $\nu = \infty$[71], so this bound is strongly satisfied. It should, therefore, be possible to observe a transition from the localized (Anderson) insulator to the

superconductor, and the question then arises, how is the superconducting state at zero temperature destroyed?

To address this issue, it is useful to think in terms of the Cooper pairs, which are interacting due to their underlying electronic constituents. By the usual number-phase uncertainty relation, the interactions drive quantum phase fluctuations and are responsible for the loss of superconductivity. The transition can then be understood as the loss of phase coherence of preformed Cooper pairs, rather than being associated with the pairing of the electrons. This suggests that the appropriate effective theory to consider is that of interacting bosonic degrees of freedom in an external random potential[125].

Further motivation for studying the dirty boson problem, at least in $D = 2$, comes from the argument that two dimensions is special with respect to the diffusive properties. The conductivity $\sigma$ has no inherent length scale, and so, at the superconductor-insulator quantum critical point, will be finite (non-zero) and universal[126, 127]. This is surprising since the states on either side of the transition have either $\sigma = 0$ (insulator) or $\sigma = \infty$ (superconductor).

For these reasons, the problem of disordered and interacting bosons has attracted much attention throughout the years [117, 128], but has proven very difficult for a theoretical analysis, since it inextricably combines the effects of interactions and Anderson localization. Just like its fermionic cousin the metal–insulator transition [129], the problem of dirty bosons seems to lack a simple analytic mean-field theory around which to begin a systematic study. Most of the information on the dirty boson quantum phase transitions derive, therefore, from numerical studies [130], and more recently from an expansion around the lower critical dimension [131–133].

In this Chapter we will be concerned with a limited class of the dirty boson models at a commensurate filling, and study the limit where the number of bosonic species $N$ is large[134]. As is well known in this limit the mean-field theory, or the saddle-point approximation, becomes the exact solution. We will see that the potential, though weakened by screening due to interactions, is always random and, thus, the eigenstates of the problem remain localized, eliminating the possibility of the superfluid phase in this limit.

## 6.2 Strongly commensurate dirty bosons

The quantum mechanical action at $T = 0$ that defines our problem is

$$S[\Psi] = \int \mathrm{d}^D\vec{x}\mathrm{d}\tau \left\{ (\partial_\tau \Psi(\vec{x},\tau))^2 + (\nabla\Psi(\vec{x},\tau))^2 + (V(\vec{x}) - \mu)\,\Psi^2(\vec{x},\tau) + \frac{\lambda}{N}\Psi^4(\vec{x},\tau) \right\}, \tag{6.1}$$

where $\Psi(\vec{x},\tau)$ is a real $N$-component bosonic field, $\Psi^2 = \sum_{\alpha=1}^{N} \Psi_\alpha^2$, and $V(\vec{x})$ is a random (in space) external potential. For simplicity, it will be assumed that $V(\vec{x})$ is uncorrelated, so that $\langle V(x)V(y)\rangle = W\delta(x-y)$. We will mostly be interested in two dimensions ($D = 2$), but will leave a general $D$ in the action to comment later on results in other dimensions. Note that disorder is assumed to be a random function only of spatial coordinates, while it is completely correlated in (i.e. independent of) the imaginary time. This is what makes it much stronger than in the corresponding problem in classical mechanics. The theory (1) for $N = 2$ describes the superfluid order parameter in the Bose-Hubbard model, at a density of bosons commensurate with the lattice [117], also known in literature as the random-rod problem [135, 136]. For $N = 3$ the theory may be used to describe disordered quantum rotors, i.e. the magnetic quantum phase transitions in the Heisenberg universality class in the presence of quenched randomness [137]. When $N = 1$ the theory describes a random system with the Ising symmetry. In general the action (1) provides a minimal description of the quantum disordered interacting system, and for $N = \infty$ has been studied by renormalisation group methods in the past [138, 139] with conflicting results. The purpose of this Chapter is to shed some light on the physics implicit in this model, and, in particular, to argue that the model allows no superfluid phase in $D = 2$.

To see what is involved in solving the problem in the (spherical) limit $N = \infty$, perform the standard Hubbard–Stratonovich transformation on the quartic term and integrate out all but one of the bosonic fields. This leaves one with the transformed action:

$$\begin{aligned} S_{\text{eff}}[\chi,\psi] \;=\; & \int \mathrm{d}^D\vec{x}\mathrm{d}\tau \Big\{ -\frac{N}{4\lambda}\chi^2(\vec{x},\tau) + (\partial_\tau\Psi_1(\vec{x},\tau))^2 \\ & + (\nabla\Psi_1(\vec{x},\tau))^2 + (V(\vec{x}) + \chi(\vec{x},\tau) - \mu)\,\Psi_1^2(\vec{x},\tau) \Big\} \\ & + \frac{1}{2}(N-1)\ln\det\left\{-\partial_\tau^2 - \nabla^2 + V(\vec{x}) + \chi(\vec{x},\tau) - \mu\right\}, \end{aligned} \tag{6.2}$$

which is just the original problem rewritten exactly. Assuming that the Hubbard–Stratonovich field at the saddle-point is independent of imaginary time, $\chi(\vec{x},\tau) = \chi(\vec{x})$, and that $\Psi_1(\vec{x},\tau) =$

$N^{\frac{1}{2}} c\phi_0(\vec{x})$, the saddle-point equations become

$$\chi(\vec{x}) = \lambda \left\langle \vec{x}, \tau \left| \frac{1}{-\partial_\tau^2 - \nabla^2 + V(\vec{x}) + \chi(\vec{x}) - \mu} \right| \vec{x}, \tau \right\rangle + c^2 \phi_0^2(\vec{x}), \tag{6.3}$$

$$\varepsilon_0 c = 0, \tag{6.4}$$

where $\phi_\alpha(\vec{x})$ are the random eigenstates, and $\varepsilon_\alpha$ the random eigenvalues of the susceptibility matrix

$$\left(-\nabla^2 + V(\vec{x}) + \chi(\vec{x}) - \mu\right) \phi_\alpha(\vec{x}) = \varepsilon_\alpha \phi_\alpha(\vec{x}), \tag{6.5}$$

with $\varepsilon_0$ being the lowest eigenvalue. The Eqs. (6.3)-(6.5) are completely standard, and the only novelty compared to the case without disorder [140] is the random spectrum instead of the usual plane waves.

From the second saddle-point equation (6.4), we see that either $\varepsilon_0 = 0$ or $c = 0$, which suggests the following possible phases. When there is a gap, $\varepsilon_0 \neq 0$, the system is a Mott insulator (MI) with localized eigenstates, and $c = 0$ simply reflects that no condensation into the lowest state has occurred. As we tune the chemical potential, and the gap decreases, we may reach a point where $\varepsilon_0 = 0$. At this point, we no longer have a Mott insulator, and $c$ may become non-zero. However, even with finite $c$, the state may not be superfluid if the lowest eigenstate remains localized. This gapless, non-superfluid state has been dubbed the Bose glass (BG)[117], and is further characterized by finite compressibility, unlike the incompressible Mott insulator. The superfluid state (SF) arises only when the lowest (gapless) state finally becomes extended. We will argue below that, in the large-$N$ model considered here, only the Mott insulator and superfluid states are possible, though in the incommensurate case, the Bose glass phase should be expected.

In the pure case with $V(\vec{x}) = 0$, the solution to the saddle-point equations is uniform, $\chi(\vec{x}) = \chi_0$, and the model leads to the well-known large-$N$ critical behavior in $D + 1$ dimensions [140]. The correlation length exponent, for example, in the pure case is $\nu = 1/(D-1)$, and in $D = 2$ the Harris criterion [124] (that says that disorder is irrelevant if $\nu D > 2$) implies that disorder is precisely *marginal*. When $V(\vec{x}) \neq 0$, in the MI phase the saddle-point Eq. (6.3), after integration over the frequency, can be written in the basis $\{\phi_\alpha\}$ as:

$$\chi(\vec{x}) = \lambda \sum_\alpha \frac{|\phi_\alpha(\vec{x})|^2}{\sqrt{\varepsilon_\alpha}}. \tag{6.6}$$

The functions $\phi_\alpha$ are the eigenstates of the *screened*, but nevertheless random, potential $V(\vec{x}) + \chi(\vec{x})$ and would, therefore, naively all be expected to be localized in $D = 2$ [123]. In particular, for the

localized ground state the first term in the sum in Eq. (6.6) becomes large as $\varepsilon_0 \to 0$ precisely in the region of localization, which by self-consistency implies that $\chi(\vec{x})$ is also large there. That, on the other hand, then implies $\varepsilon_0$ is large, and not small as assumed, and one runs into a contradiction. Evidently, for the spectrum to extend all the way to zero, the discrete sum in the last equation must be able to be approximated by an integral, so that the infrared singularity becomes integrable. For this to occur the weight of each of the terms corresponding to the low-energy states in Eq. (6.6) must vanish in the thermodynamic limit as the inverse of the system size, which is tantamount to delocalization of the low-energy eigenstates. Put differently, the collapse of the gap must be accompanied by the simultaneous delocalization of the ground state, so that the gapless phase is necessarily a SF. There can be no intermediate localized BG in the model at $N = \infty$.

With this picture in mind the appearance of the superfluid phase in the large-$N$ model in $D = 2$ appears rather counterintuitive: although screening introduces correlations into the effective random potential, the states should nevertheless always remain localized. In the rest of the Chapter we first show that although to the lowest order screening does reduce the random potential, it does not make it completely smooth and consequently the MI gap cannot close. This conclusion is further corroborated by the numerical solution of the self-consistent equations on a lattice and absence of the finite-size scaling of the gap and the ground state participation ratio. In concluding, we compare our result with other studies and speculate on the implications for physical cases $N = 1, 2, 3$.

## 6.3 Weak-disorder expansion

For a given random configuration the self-consistent equations can not be solved analytically, and one has to resort to numerical computations. For weak disorder, however, we can expand the matrix element in (6.3) in powers of the *screened* potential. To that end write $\chi(\vec{x}) = \chi_0 + \chi_1(\vec{x})$, where $\int \chi_1(\vec{x}) \mathrm{d}^D\vec{x} = 0$. The uniform part $\chi_0$ is just the renormalisation of the chemical potential, while $\widetilde{V}(\vec{x}) \equiv V(\vec{x}) + \chi_1(\vec{x})$ is the screened potential, which should vanish with vanishing randomness. Expanding the right hand side of (6.3) in the MI phase in $\widetilde{V}(\vec{x})$ and taking the Fourier transform, we get (for $q \neq 0$)

$$\begin{aligned} \widetilde{V}(\vec{q}) &= V(\vec{q}) - \lambda\Pi(\vec{q})\widetilde{V}(\vec{q}) + \lambda \int \mathrm{d}\vec{k} I_1(\vec{k},\vec{q})\widetilde{V}(\vec{k})\widetilde{V}(\vec{q}-\vec{k}) \\ &\quad -\lambda \int \mathrm{d}\vec{k}\mathrm{d}\vec{l} I_2(\vec{k},\vec{l},\vec{q})\widetilde{V}(\vec{k})\widetilde{V}(\vec{l})\widetilde{V}(\vec{q}-\vec{k}-\vec{l}) + \mathcal{O}(\widetilde{V}^4), \end{aligned} \tag{6.7}$$

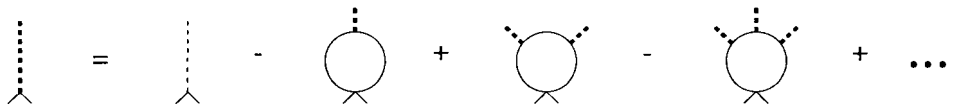

Figure 6.1: Diagrammatic representation of (6.7). The heavy dashed line represents the self-consistently screened random potential, while the thin dashed line is the bare random potential.

where

$$\Pi(\vec{q}) \equiv \int \mathrm{d}\vec{p}\mathrm{d}\omega G_0(\omega,\vec{p})G_0(\omega,\vec{p}+\vec{q}), \tag{6.8}$$

is the standard polarization bubble and

$$I_1(\vec{k},\vec{q}) \equiv \int \mathrm{d}\vec{p}\mathrm{d}\omega G_0(\omega,\vec{p})G_0(\omega,\vec{p}+\vec{k})G_0(\omega,\vec{p}+\vec{q}) \tag{6.9}$$

and

$$I_2(\vec{k},\vec{l},\vec{q}) \equiv \int \mathrm{d}\vec{p}\mathrm{d}\omega G_0(\omega,\vec{p})G_0(\omega,\vec{p}+\vec{k})G_0(\omega,\vec{p}+\vec{k}+\vec{l})G_0(\omega,\vec{p}+\vec{q}). \tag{6.10}$$

The propagator for the clean case is given by $G_0^{-1}(\omega,\vec{p}) = \omega^2 + p^2 + \Omega^2$, where $\Omega^2 \equiv \chi_0 - \mu > 0$ and is the MI gap. Eq. (6.7) can be represented diagrammatically as in Fig. 6.1.

We next introduce the two point correlator $\widetilde{W}(\vec{q})\delta(\vec{r}) = \left\langle \widetilde{V}(\vec{q})\widetilde{V}(-\vec{q}+\vec{r}) \right\rangle$, where $\langle\cdots\rangle$ represents disorder averaging, as a measure of the screened disorder. From Eq. (6.7) it follows that

$$\begin{aligned}
\widetilde{W}(\vec{q})\left\{1+\lambda\Pi(\vec{q})\right\}^2 &= W(\vec{q}) + 2\lambda\int \mathrm{d}\vec{k} I_1(\vec{k},\vec{q}) \left\langle V(-\vec{q})\widetilde{V}(\vec{k})\widetilde{V}(\vec{q}-\vec{k})\right\rangle \\
&\quad -2\lambda\int \mathrm{d}\vec{k}\mathrm{d}\vec{l} I_2(\vec{k},\vec{l},\vec{q}) \left\langle V(-\vec{q})\widetilde{V}(\vec{k})\widetilde{V}(\vec{l})\widetilde{V}(\vec{q}-\vec{k}-\vec{l})\right\rangle \\
&\quad +\lambda^2\int \mathrm{d}\vec{k}\mathrm{d}\vec{k}' I_1(\vec{k},\vec{q})I_1(\vec{k}',-\vec{q}) \left\langle \widetilde{V}(\vec{k})\widetilde{V}(\vec{q}-\vec{k})\widetilde{V}(\vec{k}')\widetilde{V}(-\vec{q}-\vec{k}')\right\rangle \\
&\quad +\mathcal{O}(W^3).
\end{aligned} \tag{6.11}$$

Diagrammatically, the second-order contributions may be represented as in Fig. 6.2. In the Appendix A.3, we compute the above averages in $D = 2$. Note that although the random potential is assumed uncorrelated in space, the screened potential develops correlations and $\widetilde{W}(\vec{q})$ becomes a non-trivial function of the wave-vector. For the low-energy states one expects the localization properties to be determined by $\widetilde{W}(\vec{q})$ at small $\vec{q}$, so we focus on the limit $\vec{q} \to 0$ and denote $\widetilde{W}(\vec{q}\to 0) = \widetilde{W}$. To the second order in $W$ in the limit $\Omega \to 0$ and in $D = 2$ one then finds

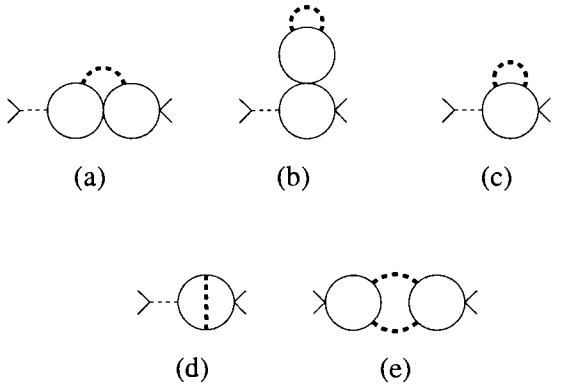


Figure 6.2: Diagrams corresponding to the second order terms in the expansion (6.11).

(see the Appendix for details):

$$\widetilde{W} = \frac{W}{\lambda^2 c^2}\Omega^2 + \left(\frac{W}{\lambda^2 c^2}\right)^2 \Omega^2 \left\{ \frac{1}{2\pi^5}\left(\left(\frac{\Lambda}{\Omega}\right)^2 + \frac{32}{\pi}\left(\frac{\Lambda}{\Omega}\right)\right) + \frac{4}{\pi^4}\left(\frac{\Lambda}{\Omega}\right) + \mathcal{O}\left(\ln\left(\frac{\Lambda}{\Omega}\right)\right)\right\} + O(W^3), \tag{6.12}$$

where the constant $c = 1/(8\pi)$, and $\Lambda$ is the ultraviolet cutoff implicit in (6.7).

The last equation is our central result, and several remarks are in order. To the first order in $W$, one finds that as $\Omega \to 0$, $\widetilde{W} \to 0$, which one may be tempted to interpret as a sign of delocalization of the ground state. This is a consequence of the screening of the random potential by the medium, which to the zeroth order in disorder is pure and, thus, screens perfectly at $q = 0$. Also, recognizing the combination $\widetilde{W}/\Omega^2$ as a dimensionless measure of screened disorder, to the lowest order Eq. (6.12) agrees with the Harris criterion: disorder is marginal in $D = 2$. The fate of disorder is, therefore, determined by the higher-order terms in the expansion. To the second order in disorder we find that

$$\widetilde{W} \to \frac{\Lambda^2}{2\pi^5}\left(\frac{W}{\lambda^2 c^2}\right)^2, \quad \text{as} \quad \Omega \to 0, \tag{6.13}$$

i.e. goes to a non-universal finite constant as the gap decreases. If the bare disorder is weak the screened disorder will be even weaker, but always finite. The consequence is that the ground state and the excited states in $D = 2$ should remain localized [123], so that our qualitative argument from the introduction would imply that the gap can not close. This is in agreement with the direct numerical solution at strong disorder to which we turn next.

## 6.4 Numerical solution

We begin by introducing a discrete version of our theory where the continuous variable $\vec{x}$ is replaced by a lattice-site index $i$ on a quadratic lattice of linear size $L$. The kinetic energy term $\nabla^2$ is replaced by the nearest-neighbour hopping measured by $t$, the random potential is chosen from a uniform distribution of width $W$ and the interaction strength is given by $\lambda$. In our calculations we set $W/t = 4$ and $\lambda/t = 8$, which corresponds to strong disorder and interactions. After the integration over frequency, the self-consistent Eq. (6.3) becomes

$$\chi_i = \lambda \sum_{\alpha=1}^{N} \frac{|\phi_\alpha(i)|^2}{\sqrt{\varepsilon_\alpha}}, \tag{6.14}$$

in the $\{\phi_\alpha(i)\}$ basis where these wave functions are now eigenvectors of the matrix

$$\sum_j \{-t_{ij} + (V_i + \chi_i - \mu)\,\delta_{i,j}\}\,\phi_\alpha(j) = \varepsilon_\alpha \phi_\alpha(i), \tag{6.15}$$

where $t_{ij}$ is non-zero for nearest-neighbour $i, j$ only.

We solve the set of $L \times L$ equations using the Newton-Raphson algorithm. We gradually increase the chemical potential $\mu$, using the last found solution as the initial guess at the next $\mu$. Finally, we average over many disorder configurations. Of course, for finite $L$ the gap is always finite, so to infer the result in the thermodynamic limit we make the standard finite-size scaling *ansatz* for the average ground state energy

$$\varepsilon_0 = L^{-z} F\left[L^{\frac{1}{\nu}}(\mu - \mu_c)\right], \tag{6.16}$$

where $z$ is the *dynamical* critical exponent, $\nu$ is the correlation length exponent, and $\mu_c$ is the critical point in the thermodynamic limit; $F(x)$ is a universal scaling function. The values of $z$ and $\mu_c$ are determined by scaling the $\varepsilon_0$-axis until all curves cross at a single point. The exponent $\nu$ is found by scaling the $\mu$-axis so that a reasonable collapse of all the data onto a single curve is achieved.

Such an attempt of finite size scaling of our data is shown in Fig. 6.3 for systems of linear size $L = 6, 8, 10, 12$. We display the result for the value $z = 0.9$, but the picture remains qualitatively the same for all $0.5 < z < 1.0$. The gap continuously decreases with $\mu$, but the failure of the finite size scaling suggests it does not vanish in the thermodynamic limit.

We have also argued that at the point of collapse of the gap, the ground state would be expected to become delocalized. A useful measure of the degree of localization of the wavefunctions at a

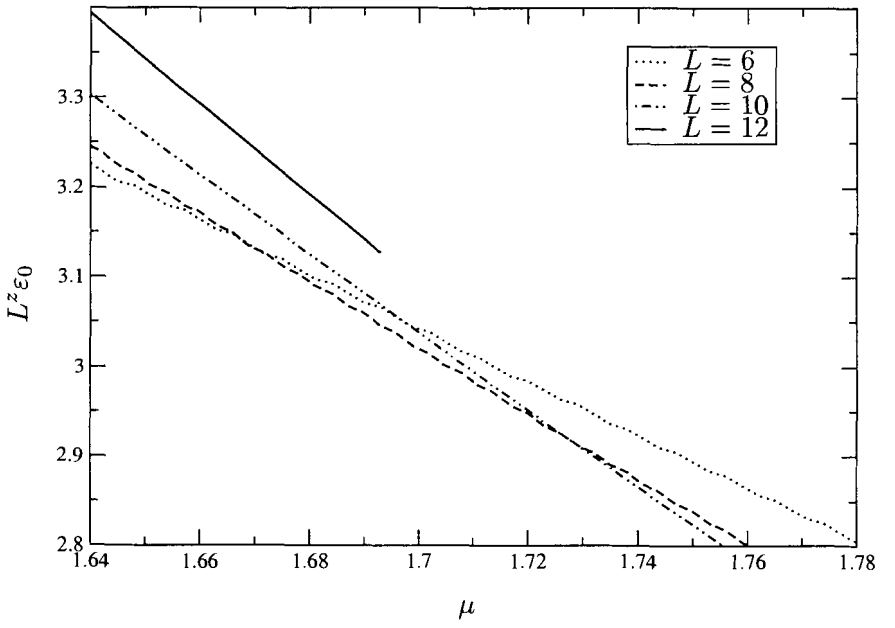


Figure 6.3: Finite size scaling attempt of the ground state energy $\varepsilon_0$ with $z = 0.9$ demonstrating the lack of a transition in our numerical calculations. The disorder averaging was done over 500 configurations for $L = 6$, 1200 for $L = 8$, 1000 for $L = 10$ and 1000 for $L = 12$.

given energy is provided by the participation ratio

$$P(\varepsilon) = \sum_{\alpha} \frac{\delta(\varepsilon_\alpha - \varepsilon)}{L^2 \sum_{i=1}^{L^2} |\phi_\alpha(i)|^4}, \tag{6.17}$$

which is proportional to $1/L^2$ for the localized states and approaches a constant for the extended ones. In the critical region, one expects the participation ratio to assume a similar finite-size scaling form:

$$P(\varepsilon_0) = L^{-(D-D_f)} \Phi\left[L^{\frac{1}{\nu}}(\mu - \mu_c)\right], \tag{6.18}$$

where $D_f$ is the fractal dimension of the ground state wavefunction and $\Phi(x)$ another scaling function. Our data for the participation ratio are shown in Fig. 6.4 for the sizes $L = 8, 10, 12$. Again, attempts to find the common crossing point by tuning $D_f$ fail. We see that the participation ratio of the ground state grows as $\mu$ is increased, but conclude that the ground state nevertheless seems to remain localized. This is consistent with the data for the ground state energy.

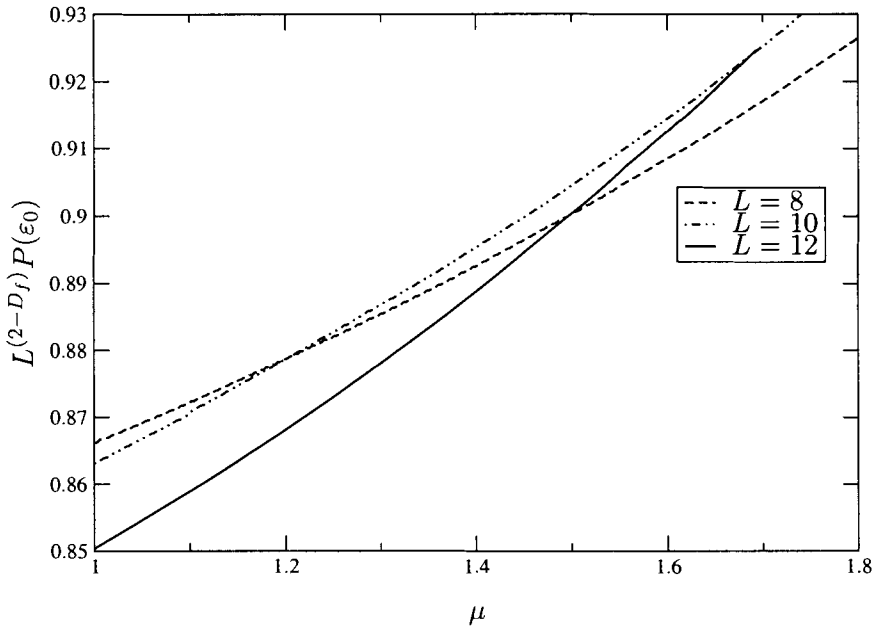


Figure 6.4: Finite size scaling attempt of the ground state participation ratio with $D_f = 0.5$. Again, the inability to cross these curves at a common point indicates the lack of the transition. Disorder averaging was done over the same configurations as in Fig. 6.3.

Our Newton-Raphson algorithm has difficulties converging as $\mu$ is increased and the problem becomes more non-linear. It is possible we simply have not been able to reach the critical point in our numerical calculation. When taken together with the weak-disorder expansion and the physical arguments, however, we believe a more likely interpretation is that there is no SF phase in the model.

## 6.5 Conclusions

To summarize, we studied the large-$N$ limit of the commensurate dirty boson theory, and argued that at weak disorder screening does not delocalize the ground state, and consequently, that there is no MI–SF transition in $D = 2$. Numerical results for the ground state energy and the participation ratio that support this conclusion were provided.

Our conclusion agrees with the results of Kim and Wen [138] who found that disorder is always relevant for $D \geq 2$ and could not find any stable critical points within their renormalisation scheme. The latter point may in principle be interpreted in three ways: as a failure of the renormalisation

procedure, as that the transition is discontinuous, or that there is no transition. Our findings support the third conclusion. On the other hand, we are in disagreement with the recent study of Hastings [139], who considered the closely related random spherical model, and found the disorder to be marginally *irrelevant* in $D = 2$. At the moment we do not fully understand what is the resolution of this disagreement, nor how the ground state becomes extended in Hastings' theory.

While we were mostly concerned with $D = 2$, the same perturbative procedure can be repeated in $D = 3$. We found that the same diagram in Fig. 6.2(e) that led to the finite term for $\widetilde{W}$ in $D = 3$ vanishes logarithmically as the gap decreases. More importantly, in $D = 3$ the Anderson localization problem allows a mobility edge, so the screened disorder need not go all the way to zero for the ground state to delocalize. We would, therefore, expect that the theory (1) would have a MI–SF transition in $D = 3$, as apparently has been found in earlier numerical calculations [141].

An important question is what our considerations imply for the physical cases $N = 1, 2, 3$ mentioned in the introduction. We believe that in $D = 2$, for $N = 2$ the theory (1) does have a transition and which is in the BG–SF universality class. This has been found in the dual theory for the commensurate dirty-bosons [125], in both $D = 1$ and $D = 2$, and in detailed numerical calculations [130, 142]. The BG–SF transition is best understood in terms of disorder-induced proliferation of topological defects and, thus, is very specific to having a complex ($N = 2$) order parameter. The same topological mechanism will not apply to the case of a random quantum ferromagnet $N = 3$, and we conjecture that for $N = 3$ there may not be a gapless phase in $D = 2$. On the same grounds, we expect that for the Ising case $N = 1$ the transition again will exist [143].

Finally, we note the similarity between our problem and the problem of interacting disordered fermions in $D = 2$ [144]. In the large-$N$ limit the metallic phase in the fermionic problem would correspond to an extended state at the Fermi level, as opposed to the extended ground state in our problem. Nevertheless, one can show [145] that already to the lowest order in disorder, screened disorder remains finite and, thus, the state should remain localized. We would, therefore, expect that the fermionic version of the action (6.1) also should have only the localized phase in $D = 2$, at least in the large-$N$ limit.

## 6.6 Disorder and the cuprates

At the beginning of this Chapter, we motivated the study of bosons in a random potential through consideration of the superconductor to insulator transition, under the assumptions of preformed

Cooper pairs and a phase fluctuation induced transition. In conventional materials, this is not necessarily relevant since the loss of phase coherence occurs with the collapse of the gap. In underdoped cuprates, on the other hand, we interpret the pseudogap as a phase fluctuating superconductor, and so the dirty boson problem is directly applicable to the study of disorder in these materials. With this reasoning, the transition from the superconductor to the pseudogap is in the BG–SF universality class, where the pseudogap is a compressible Bose glass.

The concept of the pseudogap as a compressible, non-superfluid phase has been considered before[146], though without any obvious motivation. In that study, the uniform susceptibility and electronic specific heat were calculated using a model of Dirac fermions coupled to a compressible boson current via a U(1) gauge field, and the results were shown to be in good qualitative agreement with experiments on the pseudogap state.

Stronger evidence for a BG–SF transition comes from a variety of experiments (for a review, see [102], Chap. 7) near the underdoped critical point. These seem to indicate that the critical exponents are $z \approx 1$ and $\nu \approx 1$, in agreement with the epsilon expansion near the lower critical dimension for the BG–SF transition[132]. It is clear, however, that more work is needed, both experimentally and theoretically, to establish the exact critical behaviour of the transition into the pseudogap state.

# Chapter 7

# Directions for future consideration

## 7.1 On confinement

In Chapter 3, we established that the anomalous sine-Gordon theory, believed to be dual to compact $QED_3$, has no transition, and we discussed implications of this result for gauge theories of the cuprates. While such compact theories arise often in studies of high temperature superconductivity and other strongly correlated systems, we noted that, in the framework presented in Chapter 2, the gauge fields are actually non-compact, and we briefly mentioned what is known about this brand of QED.

Several open problems immediately surface, however, from our discussion. First of all, it should be determined whether the anomalous sine-Gordon theory is indeed dual to $cQED_3$, or are monopoles irrelevant at the transition, as suggested by Hermele et al. [84]. Extending the powerful numerical methods previously used on the non-compact theory[85] to $cQED_3$ would circumvent this question, determining directly whether a deconfinement transition occurs or not.

With regards to non-compact $QED_3$, while it is believed that the deconfinement transition does occur for small coupling, the problem still remains to determine if this transition remains for all finite couplings. In terms of the effective temperature $\widetilde{T} \sim 1/N_f$, does a transition occur at $\widetilde{T} = 0$ as we tune the coupling?

Finally, we assumed that, in the presence of matter, confinement occurs due to monopole condensation, as is the case for the pure gauge theory[64]. We also assumed that chiral symmetry follows the same transition. While current evidence seems to support these assumptions[81], further work is still needed.

## 7.2 On the superfluid density

In Chapters 4 and 5, we discussed the superfluid density, primarily the effect of phase fluctuations thereupon, focusing on the in-plane value. However, much research is also dedicated to the $c$-axis superfluid density[9], which has a completely different functional form than its $ab$-plane counterpart. It would be very interesting to extend the considerations presented in this work to these $c$-axis measurements.

A theory of the $c$-axis superfluid density would also have to include the quasiparticle contribution arising due to tunneling of electrons from one layer to the next. One attempt at such a calculation can be found in Sheehy et al. [147]. The tunneling process may, of course, depend on the particular pseudogap model used, and in this way, the $c$-axis superfluid density may provide a testing ground for various different theories. For example, in the RVB theory, a spin singlet must be broken in order to hop an electron out of the plane[14], and so this excitation energy would have to be included in the calculation of the superfluid density.

For the in-plane superfluid density, we made extensive use of the Ioffe-Larkin rule to incorporate the effects of both phase fluctuations and quasiparticles. We derived this rule in the Gaussian approximation, and it is not obvious that it holds beyond this. In addition, we might wonder if this rule can be applied generally to superfluid systems when combining superfluid densities coming from different contributions. For example, when applied to conventional superconductors, the Ioffe-Larkin rule provides a simple explanation of why phase fluctuations can be ignored when determining the superfluid density: the phase stiffness is extremely large (on the order of the Fermi energy) compared to the gap energy, and so $1/\rho_{\mathrm{pf}} << 1/\rho_{\mathrm{qp}}$. It would, thus, be interesting to study a general superfluid system which has both types of contributions, to see how they combine to determine the physical superfluid density.

## 7.3 On dirty bosons

In Chapter 6, we saw that the large-$N$ theory of strongly commensurate dirty bosons has only the Mott insulating phase, though we conjectured that the cases with $N = 1, 2$ should indeed have a superfluid phase. The next obvious step is to consider the incommensurate theory, which includes the linear time derivative term. In this case, it can easily be shown that a gapless phase exists, though whether the eigenstates become delocalized or not is a much more difficult question. This problem

must necessarily be approached numerically.

It would also be interesting to calculate the superfluid density in the presence of disorder. Experimentally, sample inhomogeneity reveals itself in a modified low-temperature behaviour of the superfluid density, so this complication must be included in the evaluation of the total $\rho_s$.

## 7.4 Generally

The general framework presented in Chapter 2 provides a complete effective theory for the underdoped cuprate superconductors. Many questions have already been addressed with respect to this theory, though many more remain untouched. One problem of current interest is whether antiferromagnetism can coexist with superconductivity. There are conflicting theoretical claims on this problem[61, 62], and some recent experimental data in favour[148], so a timely solution would be quite helpful.

# Appendix A

## A.1 Proof of (3.60)

Here we will show that $F_<^{(n)}$ indeed satisfies Eqn. (3.60). To this end, first let us define $\Delta F^{(n)} \equiv F^{(n+1)} - F^{(n)}$. Then we may equivalently show that

$$F_<^{(n)} = F_0 + \Delta F^{(1)} + 2\Delta F^{(2)} + \cdots + n\Delta F^{(n)}. \tag{A.1}$$

Let us also denote the path integral over the field $\varphi(\mathbf{q})$ by Tr and define, for a real variable $t$,

$$\mathcal{F}(t) \equiv -\ln \mathrm{Tr}\left(e^{-S_0} e^{-t\Delta S}\right). \tag{A.2}$$

Then, $\mathcal{F}(1) = -\ln \mathrm{Tr}\ \exp(-S) = F_{\mathrm{ASG}}$ and

$$\left.\frac{\mathrm{d}\mathcal{F}(t)}{\mathrm{d}t}\right|_{t=1} = \frac{\mathrm{Tr}\left(\Delta S\, e^{-S}\right)}{\mathrm{Tr}\left(e^{-S}\right)} = \langle \Delta S \rangle. \tag{A.3}$$

On the other hand, we may expand the RHS of Eqn. (A.2) in powers of $\Delta S$ as

$$\mathcal{F}(t) = F_0 + \sum_{i=1}^{\infty} \Delta\mathcal{F}^{(i)}(t) \tag{A.4}$$

where $\Delta\mathcal{F}^{(i)}(t) = t^i \Delta\mathcal{F}^{(i)}(1) = t^i \Delta F^{(i)}$. Thus

$$\langle \Delta S \rangle = \left.\frac{\mathrm{d}\mathcal{F}(t)}{\mathrm{d}t}\right|_{t=1} = \sum_{i=1}^{\infty} i\ \Delta F^{(i)}. \tag{A.5}$$

Upon insertion of Eqn. (A.5) into the definition of $F_<$ in Eqn. (3.48) and truncating the expansion at $i = n$ we find (A.1).

## A.2 Proof of claim concerning $F_{\rm var}^{(n)}$

In this Appendix, we will give the proof for our claim that the extremum of $F_{\rm var}^{(n)}$ as defined in (3.52) is given by the expansion of Eqn. (3.49) to order $(\Delta S)^n$, i. e.

$$\frac{\delta F_{\rm var}^{(n)}}{\delta G_0(\mathbf{k})} = 0 \iff \langle\varphi(-\mathbf{k})\varphi(\mathbf{k})\rangle_0 = \langle\varphi(-\mathbf{k})\varphi(\mathbf{k})\rangle^{(n)}. \tag{A.6}$$

The calculations are, for general $n$, cumbersome and not very instructive so we will first present the case for $n = 2$ which is also the one with which we are concerned in Section 3.9.

Setting $n = 2$, we see that Eqn. (3.57) follows from an expansion of the RHS of Eqn. (3.50). To show that the same result arises from extremising $F_{\rm var}^{(n)}$, it is first useful to establish

$$\frac{\delta F_0}{\delta G_0(\mathbf{q})} = \left\langle \frac{\delta S_0}{\delta G_0(\mathbf{q})} \right\rangle_0 = \frac{-1}{2(2\pi)^3[G_0(\mathbf{q})]^2}\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0, \tag{A.7}$$

$$\frac{\delta\langle g\rangle_0}{\delta G_0(\mathbf{q})} = \frac{\delta F_0}{\delta G_0(\mathbf{q})}\langle g\rangle_0 + \left\langle \frac{\delta g}{\delta G_0(\mathbf{q})} - g\frac{\delta S_0}{\delta G_0(\mathbf{q})} \right\rangle_0, \tag{A.8}$$

where $g = g(S_0)$ is an arbitrary function of $S_0$. Thus choosing appropriate forms of $g$ for $F^{(1)} = F_0 + \langle\Delta S\rangle_0$ and $F^{(2)} = F^{(1)} - \frac{1}{2}\langle(\Delta S)^2\rangle_0 + \frac{1}{2}\langle\Delta S\rangle_0^2$ we find

$$\frac{\delta F^{(1)}}{\delta G_0(\mathbf{q})} = \frac{\delta F_0}{\delta G_0(\mathbf{q})}\langle\Delta S\rangle_0 - \left\langle \Delta S\frac{\delta S_0}{\delta G_0(\mathbf{q})} \right\rangle_0, \tag{A.9}$$

$$\begin{aligned}\frac{\delta F^{(2)}}{\delta G_0(\mathbf{q})} &= \frac{\delta F_0}{\delta G_0(\mathbf{q})}\left[-\frac{1}{2}\langle\Delta S^2\rangle_0 + \langle\Delta S\rangle_0^2\right] \\ &+ \frac{1}{2}\left\langle \frac{\delta S_0}{\delta G_0(\mathbf{q})}(\Delta S)^2 \right\rangle_0 - \left\langle \frac{\delta S_0}{\delta G_0(\mathbf{q})}\Delta S \right\rangle_0 \langle\Delta S\rangle_0.\end{aligned} \tag{A.10}$$

Inserting Eqn. (A.7) into Eqns. (A.9, A.10) and adding them together, we find that the restriction $\delta F_{\rm var}^{(2)}/\delta G_0(\mathbf{q}) = 0$ leads to the same equation as Eqn. (3.57).

The proof for arbitrary $n$ goes along essentially the same steps as above. Various truncated expansions we have defined can be read off the Taylor expansion identity

$$-\ln \mathrm{Tr}\left\{e^{-S_{\rm b}-V}\right\} = F_{\rm b} + \sum_{i=1}^{\infty}\sum_{l=1}^{i}\frac{(-1)^{i+l}}{l}{\sum_{\{k_\alpha\}}}'\frac{\langle V^{k_1}\rangle_{\rm b}\cdots\langle V^{k_l}\rangle_{\rm b}}{k_1!\cdots k_l!}, \tag{A.11}$$

by setting $i$ to the desired order. In (A.11) $S_{\rm b}$ and $V$ give an arbitrary splitting of the action into a bare and potential part respectively and

$${\sum_{\{k_\alpha\}}}' \equiv \sum_{k_1=1}^{i}\cdots\sum_{k_l=1}^{i}\delta_{k_1+\cdots+k_l,i}.$$

Notice that in Eqns. (A.9) and (A.10) all the terms are to the same order of $\Delta S$, which is also the largest in the corresponding expansion of the free energy. By choosing $S_{\rm b} = S_0$ and $V = \Delta S$ and setting $i = n$ in (A.11) one can see, after some lengthy algebra, that the same is true for arbitrary $n$:

$$\begin{aligned}\frac{\delta F^{(n)}}{\delta G_0(\mathbf{q})} &= \sum_{l=1}^{n}(-1)^{n+l}{\sum_{\{k_\alpha\}}}' \frac{\langle(\Delta S)^{k_1}\rangle_0 \cdots \langle(\Delta S)^{k_l}\rangle_0}{k_1!\cdots k_l!}\frac{\delta F_0}{\delta G_0(\mathbf{q})} \\ &- \sum_{l=1}^{n}(-1)^{n+l}{\sum_{\{k_\alpha\}}}' \frac{\left\langle \frac{\delta S_0}{\delta G_0(\mathbf{q})}(\Delta S)^{k_1}\right\rangle_0 \langle(\Delta S)^{k_2}\rangle_0 \cdots \langle(\Delta S)^{k_l}\rangle_0}{k_1!k_2!\cdots k_l!}. \end{aligned} \tag{A.12}$$

Let us now define, for a real variable $t$,

$$\mathcal{G}(\mathbf{k},t) \equiv -\ln \mathrm{Tr}\left\{e^{-S_0-\Delta S-t\varphi(-\mathbf{k})\varphi(\mathbf{k})}\right\}, \tag{A.13}$$

so that $\partial\mathcal{G}(\mathbf{k},t)/\partial t|_{t=0} = \langle\varphi(-\mathbf{k})\varphi(\mathbf{k})\rangle$. Then, taking $S_{\rm b} = S_0$ and $V = \Delta S + t\varphi(-\mathbf{k})\varphi(\mathbf{k})$ in Eqn. (A.11) to compute this derivative, it can be shown through additional tedious but straightforward algebra that

$$\langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle^{(n)} - \langle\varphi(-\mathbf{q})\varphi(\mathbf{q})\rangle_0 = -2n(2\pi)^3[G_0(\mathbf{q})]^2\frac{\delta F_{\rm var}^{(n)}}{\delta G_0(\mathbf{q})}, \tag{A.14}$$

where we have also made use of Eqn. (A.12). Thus, the requirement that $F_{\rm var}^{(n)}$ be an extremum implies Eqn. (3.49) truncated at $n$th order, and *vice versa*, proving our claim (A.6).

## A.3 Calculation of the two point correlator (6.11)

In this Appendix, we provide the details of the calculations leading up to our main analytic result (6.12). We begin by calculating the integrals (6.8)–(6.10). Using the standard Feynman parameters [140] the integrals can be rewritten as

$$\begin{aligned}\Pi(\vec{q}) &= \frac{\Gamma\left(\frac{3-D}{2}\right)}{(4\pi)^{\frac{D+1}{2}}}\int_0^1 \mathrm{d}t \frac{1}{\left[t(1-t)q^2+\Omega^2\right]^{\frac{3-D}{2}}} \\ &\to \begin{cases} \frac{c}{\Omega}\left[\pi\left(\frac{\Omega}{q}\right)-4\left(\frac{\Omega}{q}\right)^2+\mathcal{O}\left(\left(\frac{\Omega}{q}\right)^3\right)\right], & \frac{q}{\Omega}\to\infty \\ \frac{c}{\Omega}, & q\to 0, \end{cases}\end{aligned} \tag{A.15}$$

$$\begin{aligned}I_1(\vec{k},0) &= \frac{1}{2}\frac{\Gamma(\frac{5-D}{2})}{(4\pi)^{\frac{D+1}{2}}}\int_0^1 \mathrm{d}t \frac{1}{\left[t(1-t)k^2+\Omega^2\right]^{\frac{5-D}{2}}} \\ &\to \begin{cases} \frac{c}{4\Omega^3}\left[4\left(\frac{\Omega}{k}\right)^2-16\left(\frac{\Omega}{k}\right)^4+\mathcal{O}\left(\left(\frac{\Omega}{k}\right)^6\right)\right], & \frac{k}{\Omega}\to\infty \\ \frac{c}{4\Omega^3}, & k\to 0. \end{cases}\end{aligned} \tag{A.16}$$

where we assumed $D = 2$ in evaluating the limits. The diagrams in Figs. 6.2(c),(d) will require the evaluation of the following two limits of $I_2$:

$$\begin{aligned}I_2(\vec{k},0,0) &= \frac{\Gamma\left(\frac{7-D}{2}\right)}{(4\pi)^{\frac{D+1}{2}}}\int_0^1 \mathrm{d}t \frac{t(1-t)}{\left[t(1-t)k^2+\Omega^2\right]^{\frac{7-D}{2}}} \\ &\to \frac{3c}{4\Omega^5}\left[\frac{8}{3}\left(\frac{\Omega}{k}\right)^4+\mathcal{O}\left(\left(\frac{\Omega}{k}\right)^6\right)\right], \quad \frac{k}{\Omega}\to\infty\end{aligned} \tag{A.17}$$

$$\begin{aligned}I_2(0,\vec{l},0) &= \frac{1}{2}\frac{\Gamma\left(\frac{7-D}{2}\right)}{(4\pi)^{\frac{D+1}{2}}}\int_0^1 \mathrm{d}t \frac{(1-t)^2}{\left[t(1-t)l^2+\Omega^2\right]^{\frac{7-D}{2}}} \\ &\to \frac{3c}{8\Omega^5}\left[\frac{2}{3}\left(\frac{\Omega}{l}\right)^2+\mathcal{O}\left(\left(\frac{\Omega}{l}\right)^6\right)\right], \quad \frac{l}{\Omega}\to\infty\end{aligned} \tag{A.18}$$

where the limits $\frac{p}{\Omega}\to\infty$ and $p\to 0$ are taken with fixed $p$ and $\Omega$, respectively, and in $D = 2$. We also define $c \equiv \frac{\Gamma(\frac{1}{2})}{(4\pi)^{\frac{3}{2}}} = \frac{1}{8\pi}$.

We can now evaluate the series (6.11) term-by-term. From the first order term, in the limit $\vec{q}\to 0$, we get

$$\begin{aligned}\widetilde{W}_1(\vec{q}\to 0) &\equiv \frac{W}{(1+\lambda\Pi(0))^2} \\ &= \frac{W}{\lambda^2 c^2}\Omega^2\end{aligned} \tag{A.19}$$

The contributions of order $\mathcal{O}(W^2)$ are shown diagrammatically in Fig. 6.2. Referring to this figure, we label the corresponding terms generated in the expansion accordingly. To illustrate our procedure, we will explicitly calculate the diagram shown in Fig. 6.2(e) arising from the final term in (6.11). This term is

$$\widetilde{W}_{2(e)}(\vec{q}) = \frac{\lambda^2}{\{1+\lambda\Pi(\vec{q})\}^2}\int \mathrm{d}\vec{k}\mathrm{d}\vec{k}' I_1(\vec{k},\vec{q})I_1(\vec{k}',-\vec{q})\left\langle \widetilde{V}(\vec{k})\widetilde{V}(\vec{q}-\vec{k})\widetilde{V}(\vec{k}')\widetilde{V}(-\vec{q}-\vec{k}')\right\rangle, \tag{A.20}$$

where

$$\left\langle \widetilde{V}(\vec{k})\widetilde{V}(\vec{q}-\vec{k})\widetilde{V}(\vec{k}')\widetilde{V}(-\vec{q}-\vec{k}')\right\rangle = 2\left\langle \widetilde{V}(\vec{k})\widetilde{V}(\vec{k}')\right\rangle\left\langle \widetilde{V}(\vec{q}-\vec{k})\widetilde{V}(-\vec{q}-\vec{k}')\right\rangle \tag{A.21}$$

are the contractions which contribute for $\vec{q}\neq 0$. Using the definition

$$\widetilde{W}(\vec{q})\delta(r) \equiv \left\langle \widetilde{V}(\vec{q})\widetilde{V}(-\vec{q}+\vec{r})\right\rangle, \tag{A.22}$$

and integrating over $\vec{k}'$, (A.20) becomes

$$\begin{aligned}\widetilde{W}_{2(e)}(\vec{q}) &= \frac{2\lambda^2}{\{1+\lambda\Pi(\vec{q})\}^2}\int \mathrm{d}\vec{k} I_1(\vec{k},\vec{q})I_1(-\vec{k},-\vec{q})\widetilde{W}(\vec{k})\widetilde{W}(\vec{q}-\vec{k}) \\ &\to \frac{2\lambda^2}{\{1+\lambda\Pi(0)\}^2}\int \mathrm{d}\vec{k} I_1^2(\vec{k},0)\widetilde{W}^2(\vec{k}), \quad \vec{q}\to 0.\end{aligned} \tag{A.23}$$

We now replace $\widetilde{W}(\vec{k})$ in (A.23) to first order in $W$ to get

$$\begin{aligned}\widetilde{W}_{2(e)}(\vec{q}\to 0) &= \frac{2\lambda^2 W^2}{\{1+\lambda\Pi(0)\}^2}\int \mathrm{d}\vec{k}\frac{I_1^2(\vec{k},0)}{\left\{1+\lambda\Pi(\vec{k})\right\}^4} \\ &= \frac{1}{\pi^5}\left(\frac{W}{\lambda^2 c^2}\right)^2\Omega^2\int_0^{\frac{\Lambda}{\Omega}} x\mathrm{d}x\frac{\left[1-4\left(\frac{1}{x^2}\right)+\mathcal{O}\left(\frac{1}{x^4}\right)\right]}{\left[1-\frac{4}{\pi}\left(\frac{1}{x}\right)+\mathcal{O}\left(\frac{1}{x^2}\right)\right]},\end{aligned} \tag{A.24}$$

where the last line follows from substituting (A.15) and (A.16) into the previous line and making the change of variable $x = \dfrac{k}{\Omega}$; $\Lambda$ is the usual ultraviolet cutoff imposed by the lattice. Expanding the denominator in (A.24) and integrating over $x$ now yields the result

$$\widetilde{W}_{2(e)}(\vec{q}\to 0) = \frac{1}{2\pi^5}\left(\frac{W}{\lambda^2 c^2}\right)\Omega^2\left[\left(\frac{\Lambda}{\Omega}\right)^2 + \frac{32}{\pi}\left(\frac{\Lambda}{\Omega}\right) + \mathcal{O}\left(\ln\left(\frac{\Lambda}{\Omega}\right)\right)\right]. \tag{A.25}$$

It is important to note that (A.25) does not vanish as $\Omega\to 0$.

The calculation of the remaining terms now follows in a similar way. The diagrams arising from the second term on the RHS of (6.11) are those shown in Figs. 6.2(a),(b). These give

$$\widetilde{W}_{2(a)}(\vec{q}\to 0)=\frac{4}{\pi^4}\left(\frac{W}{\lambda^2c^2}\right)^2\Omega^2\left[\left(\frac{\Lambda}{\Omega}\right)+\mathcal{O}\left(\ln\left(\frac{\Omega}{\Lambda}\right)\right)\right], \tag{A.26}$$

and

$$\widetilde{W}_{2(b)}(\vec{q}\to 0)=\frac{1}{4\pi^3}\left(\frac{W}{\lambda^2c^2}\right)^2\Omega^2\left[\left(\frac{\Lambda}{\Omega}\right)^2+\frac{16}{\pi}\left(\frac{\Lambda}{\Omega}\right)+\mathcal{O}\left(\ln\left(\frac{\Lambda}{\Omega}\right)\right)\right]. \tag{A.27}$$

The third term on the RHS of (6.11) gives rise to the diagrams 6.2(c),(d). These give

$$\widetilde{W}_{2(c)}(\vec{q}\to 0)=-\frac{1}{4\pi^3}\left(\frac{W}{\lambda^2c^2}\right)^2\Omega^2\left[\left(\frac{\Lambda}{\Omega}\right)^2+\frac{16}{\pi}\left(\frac{\Lambda}{\Omega}\right)+\mathcal{O}\left(\ln\left(\frac{\Lambda}{\Omega}\right)\right)\right], \tag{A.28}$$

and

$$\widetilde{W}_{2(d)}(\vec{q}\to 0)=-\frac{2}{\pi^3}\left(\frac{W}{\lambda^2c^2}\right)^2\Omega^2\left[\mathcal{O}\left(\ln\left(\frac{\Lambda}{\Omega}\right)\right)\right]. \tag{A.29}$$

Note that the two highest order terms in (A.28) cancel exactly with those in (A.27). Summing the contributions (A.19) and (A.25)-(A.29), we then get the result quoted in (6.12). As mentioned, the second order term that remains constant when $\Omega \to 0$ comes entirely from the diagram 6.2(e).

# Bibliography


[1] M. Buchanan, Nature **409**, 8 (2001).

[2] J. G. Bednorz and K. A. Muller, Zeitschrift fur Physik B **64**, 189 (1986).

[3] J. R. Schrieffer, *Theory of Superconductivity* (Perseus Books, 1999).

[4] B. A. Scott, E. Y. Suarda, C. C. Tsuei, D. B. Mitzi, T. R. McGuire, B. H. Chen, and D. Walker, Physica C **230**, 239 (1994).

[5] C. N. Jiang, A. R. Baldwin, G. A. Levin, T. Stein, C. C. Almasan, D. A. Gajewski, S. H. Han, and M. B. Maple, Physical Review B **55**, R3390 (1997).

[6] W. N. Hardy, D. A. Bonn, D. C. Morgan, R. Liang, and K. Zhang, Physical Review Letters **70**, 3999 (1993).

[7] T. Ishida, K. Okuda, H. Asaoka, Y. Kazumata, and K. Noda, Physical Review B **56**, 11897 (1997).

[8] R. Liang, D. A. Bonn, W. N. Hardy, and D. Broun, Physical Review Letters **94**, 117001 (2005).

[9] A. Hosseini, D. M. Broun, D. E. Sheehy, T. P. Davis, M. Franz, W. N. Hardy, R. Liang, and D. A. Bonn, Physical Review Letters **93**, 107003 (2004).

[10] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Perseus Books, 1998).

[11] T. Pereg-Barnea, P. J. Turner, R. Harris, G. K. Mullins, J. S. Bobowski, M. Raudsepp, R. Liang, D. A. Bonn, and W. N. Hardy, Physical Review B **69**, 184513 (2004).

[12] B. Batlogg, R. J. Cava, A. Jayaraman, R. B. van Dover, G. A. Kourouklis, S. Sunshine, D. W. Murphy, L. W. Rupp, H. S. Chen, A. White, et al., Physical Review Letters **58**, 2333 (1987).

[13] A. Paramekanti, M. Randeria, T. V. Ramakrishnan, and S. S. Mandal, Physical Review B **62**, 6786 (2000).

[14] P. A. Lee, N. Nagaosa, and X.-G. Wen, cond-mat/0410445 pp. 1–69 (2004).

[15] P. A. Lee, N. Nagaosa, T.-K. Ng, and X.-G. Wen, Physical Review B **57**, 6003 (1998).

[16] B. Muhlschlegel, Zeitschrift fur Physik **155**, 313 (1959).

[17] A. Damascelli, Z. Hussain, and Z.-X. Shen, Reviews of Modern Physics **75**, 473 (2003).

[18] D. J. V. Harlingen, Reviews of Modern Physics **67**, 515 (1995).

[19] D. A. Wollman, D. J. V. Harlingen, J. Giapintzakis, and D. M. Ginsberg, Physical Review Letters **74**, 797 (1995).

[20] I. F. Herbut, Physical Review B **66**, 094504 (2002).

[21] I. F. Herbut, Physical Review Letters **88**, 047006 (2002).

[22] M. Sutherland, D. G. Hawthorn, R. W. Hill, F. Ronning, S. Wakimoto, H. Zhang, C. Proust, E. Boaknin, C. Lupien, L. Taillefer, et al., Physical Review B **67**, 174520 (2003).

[23] Y. J. Uemura, G. M. Luke, B. J. Sternlieb, J. H. Brewer, J. F. Carolan, W. N. Hardy, R. Kadono, J. R. Kempton, R. F. Kiefl, S. R. Kreitzman, et al., Physical Review Letters **62**, 2317 (1989).

[24] Y. J. Uemura, L. P. Le, G. M. Luke, B. J. Sternlieb, W. D. Wu, J. H. Brewer, T. M. Riseman, C. L. Seaman, M. B. Maple, M. Ishikawa, et al., Physical Review Letters **66**, 2665 (1991).

[25] J. E. Sonier, R. F. Kiefl, J. H. Brewer, D. A. Bonn, J. F. Carolan, K. H. Chow, P. Dosanjh, W. N. Hardy, R. Liang, W. A. MacFarlane, et al., Physical Review Letters **72**, 744 (1994).

[26] V. J. Emery and S. A. Kivelson, Nature (London) **374**, 434 (1995).

[27] Z. A. Xu, N. P. Ong, Y. Wang, T. Kakeshita, and S. Uchida, Nature **406**, 486 (2000).

[28] Y. Wang, Z. A. Xu, T. Kakeshita, S. Uchida, S. Ono, Y. Ando, and N. P. Ong, Physical Review B **64**, 224519 (2001).

[29] P. A. Lee, cond-mat/9812226 pp. 1–9 (1998).

[30] C. J. Lobb, Physical Review B **36**, 3930 (1987).

[31] D. S. Fisher, M. P. A. Fisher, and D. A. Huse, Physical Review B **43**, 130 (1991).

[32] S. Kamal, D. A. Bonn, N. Goldenfeld, P. J. Hirschfeld, R. Liang, and W. N. Hardy, Physical Review Letters **73**, 1845 (1994).

[33] L. B. Ioffe and A. I. Larkin, Physical Review B **39**, 8988 (1989).

[34] P. W. Anderson, Science **235**, 1196 (1987).

[35] M. U. Ubbens and P. A. Lee, Physical Review B **46**, 8434 (1992).

[36] L. Balents, M. P. A. Fisher, and C. Nayak, Physical Review B **60**, 1654 (1999).

[37] L. Balents, M. P. A. Fisher, and C. Nayak, Physical Review B **61**, 6307 (2000).

[38] M. Franz, Z. Tesanovic, and O. Vafek, Physical Review B **66**, 054535 (2002).

[39] M. Franz and Z. Tesanovic, Physical Review Letters **87**, 257003 (2001).

[40] A. M. Polyakov, Nuclear Physics B **120**, 429 (1977).

[41] M. J. Case, B. H. Seradjeh, and I. F. Herbut, Nuclear Physics B **676**, 572 (2004).

[42] I. F. Herbut and M. J. Case, Physical Review B **70**, 094516 (2004).

[43] M. J. Case and I. F. Herbut, cond-mat/0504266 pp. 1–6 (2005).

[44] M. J. Case and I. F. Herbut, Journal of Physics A: Mathematical and General **34**, 7739 (2001).

[45] A. Hosseini, R. Harris, S. Kamal, P. Dosanjh, J. Preston, R. Liang, W. N. Hardy, and D. A. Bonn, Physical Review B **60**, 1349 (1999).

[46] A. Kaminski, J. Mesot, H. Fretwell, J. C. Campuzano, M. R. Norman, M. Randeria, H. Ding, T. Sato, T. Takahashi, T. Mochiku, et al., Physical Review Letters **84**, 1788 (2000).

[47] I. F. Herbut, Physical Review B **70**, 184507 (2004).

[48] M. Franz and Z. Tesanovic, Physical Review Letters **84**, 554 (2000).

[49] J. Villain, Journal de Physique **36**, 581 (1975).

[50] J. Jose, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, Physical Review B (Solid State) **16**, 1217 (1977).

[51] H. Kleinert, *Gauge Fields in Condensed Matter* (World Scientific, 1989).

[52] M. E. Peskin and D. V. Schroeder, *Introduction to Quantum Field Theory* (Westview Press, 1995).

[53] R. D. Pisarski, Physical Review D **29**, 2423 (1984).

[54] T. W. Appelquist, M. Bowick, D. Karabali, and L. C. R. Wijewardhana, Physical Review D **33**, 3704 (1986).

[55] P. Maris, Physical Review D **54**, 4049 (1996).

[56] E. Dagotto, J. B. Kogut, and A. Kocic, Physical Review Letters **62**, 1083 (1989).

[57] Z. Tesanovic, O. Vafek, and M. Franz, Physical Review B **65**, 180511 (2002).

[58] B. H. Seradjeh and I. F. Herbut, Physical Review B **66**, 184507 (2002).

[59] D. J. Lee and I. F. Herbut, Physical Review B **66**, 094512 (2002).

[60] D. J. Lee and I. F. Herbut, Physical Review B **67**, 174512 (2003).

[61] I. F. Herbut and D. J. Lee, Physical Review B **68**, 104518 (2003).

[62] T. Pereg-Barnea and M. Franz, Physical Review B **67**, 060503 (2003).

[63] D. H. Kim and P. A. Lee, Annals of Physics (New York) **272**, 130 (1999).

[64] A. M. Polyakov, *Gauge Fields and Strings* (Harwood Academic, 1987).

[65] W. Rantner and X.-G. Wen, Physical Review Letters **86**, 3871 (2001).

[66] X.-G. Wen, Physical Review B **65**, 165113 (2002).

[67] H. Kleinert, F. S. Nogueira, and A. Sudbo, Physical Review Letters **88**, 232001 (2002).

[68] H. Kleinert, F. S. Nogueira, and A. Sudbø, Nuclear Physics B **666**, 361 (2003).

[69] G.-Z. Liu, cond-mat/0406662 pp. 1–4 (2004).

[70] V. L. Berezinskii, Soviet Physics JETP **34**, 610 (1972).

[71] J. M. Kosterlitz and D. J. Thouless, Journal of Physics C: Solid State Physics **6**, 1181 (1973).

[72] J. M. Kosterlitz, Journal of Physics C: Solid State Physics **7**, 1046 (1974).

[73] J. M. Kosterlitz, Journal of Physics C: Solid State Physics **10**, 3753 (1977).

[74] I. F. Herbut and B. H. Seradjeh, Physical Review Letters **91**, 171601 (2003).

[75] I. F. Herbut, B. H. Seradjeh, S. Sachdev, and G. Murthy, Physical Review B **68**, 195110 (2003).

[76] G. Murthy, Physical Review Letters **67**, 911 (1991).

[77] S. Sachdev and K. Park, Annals of Physics (New York) **298**, 58 (2002).

[78] H. J. Rothe, *Lattice Gauge Theories: An Introduction* (World Scientific, 1997).

[79] R. P. Feynman, *Statistical Mechanics, A Set of Lectures* (Perseus Books, 1998).

[80] T. Giamarchi and P. L. Doussal, Physical Review B **53**, 15206 (1996).

[81] H. R. Fiebig and R. M. Woloshyn, Physical Review D **42**, 3520 (1990).

[82] A. Diehl, M. C. Barbosa, and Y. Levin, Physical Review E **56**, 619 (1997).

[83] V. Azcoiti and X. Q. Luo, Modern Physics Letters A **8**, 3635 (1993).

[84] M. Hermele, T. Senthil, M. P. A. Fisher, P. A. Lee, N. Nagaosa, and X.-G. Wen, Physical Review B **70**, 214437 (2004).

[85] S. J. Hands, J. B. Kogut, and C. G. Strouthos, Nuclear Physics B **645**, 321 (2002).

[86] E. L. Andronikashvili, Journal of Physics (USSR) **10**, 201 (1946).

[87] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, 1971).

[88] D. Bonn, S. Kamal, A. Bonakdarpour, R. Liang, W. Hardy, C. Homes, D. Basov, and T. Timusk, Czechoslovak Journal of Physics **46**, 3195 (1996).

[89] J. Annett, N. Goldenfeld, and S. R. Renn, Physical Review B **43**, 2778 (1991).

[90] J. L. Tallon, J. W. Loram, J. R. Cooper, C. Panagopoulos, and C. Bernhard, Physical Review B **68**, 180501 (2003).

[91] M. E. Fisher, M. N. Barber, and D. Jasnow, Physical Review A **8**, 1111 (1973).

[92] P. B. Weichman, Physical Review B **38**, 8739 (1988).

[93] J. Rudnick and D. Jasnow, Physical Review B **16**, 2032 (1977).

[94] J. R. Banavar and D. Jasnow, Journal of Physics A: Mathematical and General **11**, 1361 (1978).

[95] W. Meissner and R. Oschenfeld, Die Naturwissenschaften **21**, 787 (1933).

[96] H. K. Onnes, Commun. Phys. Lab. Univ. Leiden, Suppl. **34b** (1913).

[97] J. D. Jackson, *Classical Electrodynamics* (John Wiley and Sons, Inc., 1975).

[98] F. London, *Superfluids* (Dover, 1961).

[99] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Perseus Books, 1992).

[100] J. A. Hertz, Physical Review B **14**, 1165 (1976).

[101] D. R. Nelson and J. M. Kosterlitz, Physical Review Letters **39**, 1201 (1977).

[102] T. Schneider and J. M. Singer, *Phase Transition Approach to High Temperature Superconductivity* (Imperial College Press, 2000).

[103] R. G. Goodrich, P. W. Adams, D. H. Lowndes, and D. P. Norton, Physical Review B **56**, R14299 (1997).

[104] P. A. Lee and X.-G. Wen, Physical Review Letters **78**, 4111 (1997).

[105] A. Paramekanti and M. Randeria, Physical Review B **66**, 214517 (2002).

[106] J. Stajic, A. Iyengar, K. Levin, B. R. Boyce, and T. R. Lemberger, Physical Review B **68**, 024520 (2003).

[107] R. Gupta, J. DeLapp, G. G. Batrouni, G. C. Fox, C. F. Baillie, and J. Apostolakis, Physical Review Letters **61**, 1996 (1988).

[108] R. Gupta and C. F. Baillie, Physical Review B **45**, 2883 (1992).

[109] P. Minnhagen, Reviews of Modern Physics **59**, 1001 (1987).

[110] I. F. Herbut and Z. Tesanovic, Physical Review Letters **76**, 4588 (1996).

[111] J. Adler, C. Holm, and W. Janke, Physica A **201**, 581 (1993).

[112] L. B. Ioffe and A. J. Millis, cond-mat/0112509 pp. 1–12 (2001).

[113] A. J. Millis, S. M. Girvin, L. B. Ioffe, and A. I. Larkin, J. Phys. Chem. Solids **59** (1998).

[114] T. Schneider, cond-mat/0210478 pp. 1–15 (2002).

[115] S. Doniach and M. Inui, Physical Review B **41**, 6668 (1990).

[116] H. A. Fertig and S. D. Sarma, Physical Review B **44**, 4480 (1991).

[117] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Physical Review B **40**, 546 (1989).

[118] A. L. Fetter, Annals of Physics **88**, 1 (1974).

[119] D. S. Fisher and P. C. Hohenberg, Physical Review B **37**, 4936 (1988).

[120] M. A. Kastner, R. J. Birgeneau, G. Shirane, and Y. Endoh, Reviews of Modern Physics **70**, 897 (1998).

[121] I. F. Herbut, Physical Review Letters **94**, 237001 (2005).

[122] M. J. Case and I. F. Herbut, Journal of Physics A: Mathematical and General Physics **35**, 2523 (2002).

[123] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Physical Review Letters **42**, 673 (1979).

[124] A. B. Harris, Journal of Physics C: Solid State Physics **7**, 1671 (1974).

[125] I. F. Herbut, Physical Review B **57**, 13729 (1998).

[126] M. P. A. Fisher, G. Grinstein, and S. M. Girvin, Physical Review Letters **64**, 587 (1990).

[127] I. F. Herbut, Physical Review Letters **79**, 3502 (1997).

[128] J. A. Hertz, L. Fleishman, and P. W. Anderson, Physical Review Letters **43**, 942 (1979).

[129] D. Belitz and T. R. Kirkpatrick, Reviews of Modern Physics **66**, 261 (1994).

[130] M. Wallin, E. S. Sørensen, S. M. Girvin, and A. P. Young, Physical Review B **49**, 12115 (1994).

[131] I. F. Herbut, Physical Review Letters **81**, 3916 (1998).

[132] I. F. Herbut, Physical Review B **58**, 971 (1998).

[133] I. F. Herbut, Physical Review B **61**, 14723 (2000).

[134] Y. Tu and P. B. Weichman, Physical Review Letters **73**, 6 (1994).

[135] D. Boyanovsky and J. L. Cardy, Physical Review B **26**, 154 (1982).

[136] I. D. Lawrie and V. V. Prudnikov, Journal of Physics C: Solid State Physics **17**, 1655 (1984).

[137] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, 1999).

[138] Y. B. Kim and X.-G. Wen, Physical Review B **49**, 4043 (1994).

[139] M. B. Hastings, Physical Review B **60**, 9755 (1999).

[140] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Cambridge University Press, 1996).

[141] J. W. Hartman and P. B. Weichman, Physical Review Letters **74**, 4584 (1995).

[142] S. Rapsch, U. Schollwöck, and W. Zwerger, Europhysics Letters **46**, 559 (1999).

[143] D. S. Fisher, Physical Review B **51**, 6411 (1995).

[144] E. Abrahams, S. V. Kravchenko, and M. P. Sarachik, Reviews of Modern Physics **73**, 251 (2001).

[145] I. F. Herbut, Physical Review B **63**, 113102 (2001).

[146] D. H. Kim, P. A. Lee, and X.-G. Wen, Physical Review Letters **79**, 2109 (1997).

[147] D. E. Sheehy, T. P. Davis, and M. Franz, Physical Review B **70**, 054510 (2004).

[148] Y. Saiga and M. Oshikawa, Physical Review B **68**, 094511 (2003).