Abstract

This is a LaTeX template intended for academic theses, and was put together by Jabir Ali Ouassou while preparing his PhD dissertation. The template itself is released under a Creative Commons Attribution licence (CC BY 4.0). This basically means that you are free to use the template for any purpose as long as you give appropriate credit.

The template bundles the Libertinus fonts, which is used for all regular text and mathematics, and the URW classico fonts, which are used for chapter and section headings. The former is available under the Open Font Licence (SIL OFL 1.1), and is free for both private and commercial use. The latter is available under the Aladdin Free Public Licence (AFPL), and is only free for non-commercial use. If commercial use is of importance, a suitable replacement for URW classico would be the Libertinus Sans fonts, which are also bundled with the template.

Note that this template relies on LuaLaTeX for e.g. font customization, and on BibTeX for reference handling. For command-line users, the easiest way to compile the document is to run `latexmk -lualatex thesis.tex`. If using an IDE, please check the program settings for how to enable compilation with LuaLaTeX and BibTeX. The template is based on the KOMA-Script book class (`scrbook`), so for further customization of the template, please check out their documentation.

The template does not include a title page. This is because the style requirements typically varies between universities, and many institutions will anyway autogenerate a titlepage upon thesis submission.

Preface

This would be a natural place to specify what kind of thesis this is, acknowledge your supervisor and coworkers, and so on.

Contents

Chapter 1

Introduction

In the world today we see an ever increasing need to not waste, but use energy in the form of electricity as efficiently as possible. There is also a continuous drive to create ever more advanced computers to process the ever increasing amounts of data that we produce and to help us calculate and come up with further innovation. At a fundamental level, the limit of how much advancement we can make in these areas is dependent on the materials available to build solutions out of and our understanding of how these materials behave. In the stone age, the material understanding allowed technology to be made out of wood, bone and stone. In the bronze age, technology superiour to this could be made because we had discovered how to extract a metal from its natural stony oxide. When it comes to solutions using electricity then the electrical properties of materials is naturally especially important. The current age has been called the information age, and most of the infrastructure upon which this information is stored and communicated through is possible because of our understanding of the behaviour of electricity, or more precicely electrons in materials. Electricity is understood as a stream of electrons and thus we often talk in this respect of materials electronic properties.

In 1911, the electronic property called superconductivity was discovered in Netherlands by Heike Onnes [1]. Mercury was cooled to $-268.99\,^{\circ}\mathrm{C}$ and then Dr. Onnes measured the resistivity, which is a measure of how much energy is used to push electrons through a ma-

terial. He found that the resistivity suddenly dissapeared at this temperature and lower. It was expected that the resistivity would become lower with temperature but the extreme suddenness of the dissapearance was surprising. Not only did resistivity dissapear so that electrons could flow completely freely through the material as if there was no material there at all, but in 1933 it was discovered that at this low temperature, its inside was devoid of any magnetic field, even if a magnetic field had previously penetrated the material at the higher temperatures. When a material changes its properties in such a sudden manner with respect to a controlling variable, in this case temperature, we call it a phase transition. Just as when the mechanical properties of water changes from making it easily fill the shape of a cup in its liquid form, to making it being stuck on top and impossible to drink after it has frozen into ice, so the electronic properties of mercury changes in a fundamental way as well during a phase transition. The phase transition that happens to water when it freezes into ice can be described by classical mechanics which is built on Newtons theories, however because of the Meissner effect, classical mechanics is not sufficient to describe the phase-transition that happens to mercury at $-268.99\,^{\circ}\mathrm{C}$. For that we need quantum mechanics. Thus the new low-temperature phase of mercury is not called a perfectly conducting phase, but a superconducting phase.

1.1 Climate Change and its solutions

Arguably one of modern industrialized civilization's largest future challenges is finding a way to reduce the amount of climate gases like carbon dioxide in the atmosphere.

Write an INTRODUCTION to the field field here.

Chapter 2

Statistical Mechanics

In statistical mechanics we attempt to describe an ensamble of particles that may be interacting to extract not precise information about what each and every particle is doing, but statistical information about what most of the particles are doing. We zoom out and look upon the bunch of particles and try to answer the question of what is this bunch doing. What is the most significant behavior of this bunch as a whole.

2.1 Canonical ensamble and the partition function

Most of the business of statistical mechanics is about calculating what is known as the partition- function. Once this function is known, all the heavy lifting is done since most important statistical quantities can be extracted from it following already established systematic steps. To calculate the partition function is theoretically very simple: we sum the quantity $e^{-\beta E_i}$ over all the possible states of the system. Every state is a particular configuration of things in the system and since all things in the system has a certain energy, if we sum the energy of all the things in the system for each state, we can say that each state has a certain energy. If we label each state of the system with the index i, then we can denote the energy of each state E_i. The definition of the partition function Z in the canonical ensamble can then be written

$$Z = \sum_i e^{-\beta E_i}, \tag{2.1}$$

where $\beta = 1/(k_{\mathrm{B}}T)$ and k_{B} is the Boltzmann constant given by $k_{\mathrm{B}} \approx 1.380\,649 \times 10^{-23}\,\mathrm{JK}^{-1}$.

As we can see, the essential ingrediants needed to calculate the partition function is to be able to enumerate all possible states i of the system and also calculate their corresponding energy E_i. Since we have used the summation sign $\sum_i$ in Eq. (2.1), we have implicitely assumed that there exists a finite number of different states. However if there is one thing in the system that can change in a continuous fashion, which we would measure using the set of real numbers $\mathbb{R}$ and some unit, then the number of states is infinite. In this case we sum over the different numbers of states by simply integrating over the things that are continuous and the unit of the partition function becomes the product of the units of the continuous variables (things). In most cases, it is the position and momentum of particles in the system that are continuous, hence the definition of the partition function becomes[1]

$$Z \sim \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; e^{-\beta E(\boldsymbol{r},\boldsymbol{p})}. \tag{2.2}$$

2.2 Calculating observables

An observable in statistical mechanics is anything that we want to measure. In quantum mechanics observables are restricted to operators that have real eigenvalues which makes sense considering that we don't really have an intuition for complex numbers, which is the alternative, and thus we want to restrict things we can observe to things we can understand in terms of a point on a single line (the real axis).

Since we are interested again in ensambles of many particles we are restricting our attention to statistical information about this ensamble. To get this information we need some kind of probability distribution of the particular states of the system. We are imagining that we for each such state (indexed by j) can calulate a real number for the thing (observable) we are interested in measuring. Let's call this observable O. Then O is a statistical variable but takes a particular value o_j in the state j of the system. If we now let P_j be the probability distribution,

1. The reason why there is a $\sim$ sign in Eq. (2.2) is that specifically there is a factor of Planck's constant h in the denominator for each $\mathrm{d}r\,\mathrm{d}p$ in the integral measure since this makes the partition function dimensionless and thus consistent with the definition in terms of finite number of states.

i.e. the probability that the system exists in state j, then we know from probability theory that the mean of the observaible is

$$\langle O \rangle = \sum_j o_j P_j. \tag{2.3}$$

The probability distribution P_j is given by

$$P_j = e^{-\beta E_j}/Z, \tag{2.4}$$

in terms of the partition function Z. Inserting this we get

$$\langle O \rangle = \sum_j \frac{o_j e^{-\beta E_j}}{Z}. \tag{2.5}$$

2.3 Ginzburg-Landau model

The experimental discovery of superconductivity was a surprise to the scientists at the time. No theoretical model had so far predicted the properties that the experimentalists were measuring. On the contrary, they predicted very different results no matter how they were twisted and turned, and thus superconductivity seemed to demand a radically different understanding of how electrons moved in a material.

2.3.1 Landau Model

Before such an understanding had been developed, Landau took a shortcut and came up with a theory that could describe the phenomenon of superconductivity without knowing its microscopic origin. In other words he treated superconductivity as a black box and instead of asking what was inside to give the boxes output, he used the output to determine a small set of *material parameters* which could then be used to predict how the box would react to a large range of stimuli or conditions. The merit of the theory was that he used symmetry arguments to reduce this set as much as possible.

The Ginzburg-Landu theory of superconductivity is based on Landau's previous work on a theory for general second order phase-transitions[2]

2. 2nd order phase-transitions are phase transitions of systems whose free energy has a discontinuous second order derivative at the transition point, but is continuous for lower orders. Since the specific heat is given by the second order derivative, then the specific heat is discontinuous in this case.

The approach is given by two ideas. The first is simply that the phase transition should be able to be characterized by the appearence of some kind of measurable order that can be described by a function ψ which we call the *order parameter*. In the liquid water to ice transition, it is the position of the molecules that become ordered in a lattice,[3] in the magnetization of a metal it is the individual spins that become ordered along a particular direction.

The second idea is that at the phase transition, it is the appearence of this order that should dominate the behaviour of the system, to the exclusion of all other effects. Thus the system should be described in terms of the order-parameter and since this is infinitesimally small close to the transition, the free energy[4] can be expanded in a Maclaurin-series with respect to this parameter as

$$F = F_0 + c_1\psi + c_2\psi^2 + c_3\psi^3 + c_4\psi^4 + c_5\psi^5 + ... \tag{2.6}$$

The real constants F_0 and c_i constitute the set of material parameters of the theory and this set can then be reduced by any symmetries that we suspect should be inherent in the underlying theory. For example, if ψ should represent the order parameter of magnetization of a system of ising spins, which can point either up or down, then the free energy should be invariant of this global choice, i.e. we need to enforce that the free energy be invariant with respect to the transformation $\psi \mapsto -\psi$. Then all the constants c_j for odd j vanish.

In the case of superconductivity the order parameter Ψ represents the probability amplitude of the collective state of the superfluid of Cooper paired electrons such that $|\Psi|$ can be interpreted as the density of such electron pairs. Since Ψ is complex it has to be combined with its complex conjugate Ψ^* in ways that yield real numbers to produce terms that are valid in the free energy since F itself should be a real number. Furthermore, the phenomenon of superconductivity is produced as a result of the breaking of $U(1)$ symmetry, so F needs

3. The astute reader might have noticed that this example is a first order phase transition because of the existance of latent heat. Actually first order phase transitions is also able to be described by a modified Landau theory, however we will here focus on the second order kind.
4. In this case we are talking about the Helmholtz free energy which is related to the partition function through a logarithm while at the same time the Legendre transformation of the internal energy of the thermodynamic system.

also to be $U(1)$ symmetric, i.e. it has to be invariant under the transformation $\Psi \mapsto e^{i\phi}\Psi$ for $\phi \in \mathbb{R}$. These restrictions result in the free energy

$$F = F_0 - a|\Psi|^2 + b|\Psi|^4, \tag{2.7}$$

when keeping the lowest order terms that produce a phase transition.

Thermodynamic equilibrium is reached at the minimum of free energy. This restricts $b \geq 0$ since negative b yields a free energy with no definite minimum. The minimum is then found by the condition

$$\frac{\partial F}{\partial \Psi^*} = (-a + 2b|\Psi|^2)\Psi = 0, \tag{2.8}$$

which yields the possibilities $|\Psi| = 0$ or $|\Psi| = \sqrt{a/2b}$. The first case gives the energy $F = F_0$, while the second gives $F = F_0 - a^2/(4b)$. We see that the second case is energetically favorable, but only exists and is different from the first case when $a > 0$. Furthermore, the second case represents the ordered state since in this case the order-parameter $|\Psi| \neq 0$, in the conventional Landau theory.[5]

In superconductivity it is the parameter of temperature which conventionally determines when the system enters the superconducting regime. Looking at the free energy in Eq. (2.7), the order parameter Ψ is the dynamical variable of the theory while the explicit temperature dependence lies in the material parameters a and b. Denoting the temperature at which the phase transition happens — the critical temperature: T_c, the dimensionless parameter $t = (T - T_c)/T_c$ is small close to the critical point which means it can be used to expand the temperature dependence of the material parameters such that

$$\begin{aligned} a(T) &= a_0 + a_1 t + ... \\ b(T) &= b_0 + b_1 t + ... \end{aligned} \tag{2.9}$$

Now we argue for what terms to keep in these expansions. Since $a(T)$ should change sign at $t = 0$ then we only keep odd terms. Since

5. Actually this only represents when Cooper-pairs are forming and the real onset of superconductivity is determined by the point in parameter-space where the gauge-mass becomes non-zero, which is closely related but not exactly the same as where the density of Cooper-pairs becomes non-zero. The real onset of superconductivity is thus more related to when the phase of the wave-function settles on a value.

we need $b(T) > 0$ for the theory to be thermodynamically stable it seems that b_0 is the important term which need to be larger than any negative contributions from the other terms. Keeping only lowest order terms then the expansions reduce to $a(T) = a_1(T - T_c)/T_c$ and $b(T) = b_0$. Since the ordered state is the solution of the theory when $a > 0$ and this ordered state exists at temperatures $T < T_c$ then $a_1 < 0$ and the final temperature dependence of a becomes $a(T) = -|a_1|(T - T_c)/T_c$. From this temperature dependence, it is straight forward to derive critical exponents, the specific heat etc.

2.3.2 Gradient Terms

The simple Landau theory described above is a type of mean field theory in that there is no spatial dependence in the solution, and thus gives a simplified picture that can only be valid far away from any defects or boundaries. This simple approach can be extended to include spatial variation by allowing terms with gradients of the order parameter in the free energy through a gradient expansion

$$F = \int \mathrm{d}^3 r \; f(\Psi, \nabla\Psi, \nabla^2\Psi, \nabla^3\Psi, ...). \tag{2.10}$$

Keeping only the lowest order in this expansion that is invariant under $U(1)$ symmetry we get the term $|\nabla\Psi|^2$.

Perhaps the single most important phenomenon of superconductivity from a theoretical stand-point is the fact that it expells magnetic fields, hence it is clear that any theory that attempts to explain superconductivity needs to have some way the superconducting order interacts with magnetic fields. The standard way to acheive this is through the recipe of *minimal coupling*, where the vector potential $\boldsymbol{A}$ times a constant is subtracted with from any momentum in the previously neutral theory. Specifically $\boldsymbol{p} \mapsto \boldsymbol{p} - q/c\boldsymbol{A}$ where q is the charge of the particle and c is the speed of light.

$$f = f_0 - a|\Psi|^2 + b|\Psi|^4 + K|(\nabla + ig\boldsymbol{A})\Psi|^2 - \int_0^{B_a} \boldsymbol{M} \cdot \mathrm{d}\boldsymbol{B}_a \tag{2.11}$$

<++>

Chapter 3

Field Theory Methods

In this chapter we will give a short introduction to the use of grassmann variables and complex numbers in the calculation of the field-integrals on the statistical-mechanical partition function and how they can be used to transform the expression for the action through the Hubbard-Stratonovich transformation.

A field theoretic expression for the quantum mechanical partition function $\mathcal{Z}$ is obtained by using a coherent state basis. A coherent state is the eigen-state of an annihilation operator, thus it produces an eigenvalue when operated on by the annihilation operator. Letting $\hat{H}$ be the quantum mechanical Hamiltonian of the system for which we are interested in calculating the parition function, μ the the chemical potential and $\hat{N}$ is the number operator then the partition function

$$\mathcal{Z} = \mathrm{Tr}(e^{-\beta(\hat{H}-\mu\hat{N})}). \tag{3.1}$$

Inserting a basis of coherent states $\{|\xi\rangle\}$ when calculating the trace, we obtain a functional integral over the coherent state eigenvalues ξ_α and ξ^*_α which substitute the annihilation and creation operators c_α and $c^\dagger_\alpha$ respectively. Here α symbolize the collection of quantum-numbers needed to specify a state. The functional integral then takes the form

$$\mathcal{Z} = \int \mathcal{D}[\xi^*\,\xi] e^{-\int_0^\beta \mathrm{d}\tau \sum_\alpha [\xi^*_\alpha(\partial_\tau - \mu)\xi_\alpha + H(\xi_\alpha, \xi^*_\alpha)]}. \tag{3.2}$$

The integration variable τ is the imaginary time and a τ dependence is implicit in the notation such that $\xi_\alpha = \xi_\alpha(\tau)$. This path-integral no-

tation is a shorthand for a more involved expression where the imaginary time-dependence of τ is split into a collection of time-indexed coherent state eigenfunctions $\xi_{\alpha,\tau}$ and the integration measure is a product over these indices and the quantum-state indices α. For further detail we refer to Ref. [2] which we will follow for a large part of this chapter.

3.1 Quadratic Fermionic Field Integrals

Because of the anti-commuting property of the fermion annihilation operators, any coherent state[1] has to have eigenvalues that anti-commute as well. This leads to the partition function written in the convenient basis of the coherent states being constructed with Graßmann numbers as the central variables.

3.1.1 Graßmann algebras

A Graßmann algebra is constructed on a set of generators $\{\xi_\alpha\}$ such that a specific product of the generators $\xi_{\alpha_1}\xi_{\alpha_2}\cdots\xi_{\alpha_n}$ together with a complex coefficient ϕ constitute a number in the algebra and the generators anti-commute such that $\xi_\alpha\xi_\beta = -\xi_\beta\xi_\alpha$. On such an algebra, differentiation can be defined such that

$$\frac{\mathrm{d}}{\mathrm{d}\xi_{\alpha_m}}\,\phi\,\xi_{\alpha_1}\cdots\xi_{\alpha_n} = (-1)^m\phi\,\xi_{\alpha_1}\cdots\xi_{\alpha_{m-1}}\xi_{\alpha_{m+1}}\cdots\xi_{\alpha_n}, \tag{3.3}$$

provided the generator ξ_{α_m} is in the number and 0 otherwise. The factors of (-1) comes from anti-commuting the generator ξ_{α_m} such that it is next to the differentiation operator. In Graßmann algebra, integration can be (perhaps a little non-intuitively) be defined such that it acts in the same way as differentiation, i.e. generators have to be anti-commuted until they are next to the symbolic differential symbol $\mathrm{d}\xi_\alpha$, and then

$$\int \mathrm{d}\xi\;\xi = 1, \tag{3.4}$$

while

$$\int \mathrm{d}\xi\;1 = 0. \tag{3.5}$$

1. An eigenstate of the annihilation operator

On an algebra consisting of $2n$ generators we define conjugation as a map from the first half of the generators $\{\xi_{\alpha_i}\}_{i=1}^{n}$ to the other half $\{\xi^*_{\alpha_i}\}_{i=1}^{n}$ and in such a way that when applied to a particular number

$$(\phi\xi_\alpha\xi_\beta)^* = \phi^*\xi^*_\beta\xi^*_\alpha. \tag{3.6}$$

3.1.2 Nambu Spinor

In the Nambu notation we group spin-dependent Graßmann numbers $\xi_\uparrow$ and $\xi^*_\downarrow$, which correspond to the annihilation- and creation-operators $\hat{c}^\dagger_\uparrow$ and $\hat{c}_\downarrow$, in a vector called a Nambu spinor

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_\uparrow \\ \xi^*_\downarrow \end{pmatrix}. \tag{3.7}$$

A sesquilinear form can then be created with this vector and its adjoint such that

$$\boldsymbol{\xi}^\dagger S\boldsymbol{\xi} = S_{11}\xi^*_\uparrow\xi_\uparrow + S_{22}\xi^*_\downarrow\xi_\downarrow + S_{12}\xi^*_\uparrow\xi^*_\downarrow + S_{21}\xi_\uparrow\xi_\downarrow. \tag{3.8}$$

This allows any action that contains spin-dependent terms of the form of the right hand side of Eq. (3.8) to be put on sesquilinear form. Assuming this is the case, then the partition function in the field-integral representation takes the form

$$\mathcal{Z} = \int \mathcal{D}[\xi^*\,\xi]\; e^{-\int_0^\beta \mathrm{d}\tau\, \boldsymbol{\xi}^\dagger_\gamma S_{\gamma\delta}\boldsymbol{\xi}_\delta}. \tag{3.9}$$

In this equation, the indices γ and δ is an arbitrary collection of quantum numbers needed to specify a state other than spin, for example they could be momentum indices $\gamma = \boldsymbol{k}, \delta = \boldsymbol{k}'$, and summation over these repeated indices is implicitly understood.

Splitting the integral over τ into M imaginary time-slices and expanding the path integral measure into a product of individual integrals over specific quantum numbered and time-sliced Graßmann variables such that

$$\int \mathcal{D}[\xi^*\,\xi] \;\propto\; \lim_{M\to\infty} \int \prod_{\tau=1}^{M}\prod_{\alpha} \mathrm{d}\xi^*_{\alpha,\tau}\mathrm{d}\xi_{\alpha,\tau}, \tag{3.10}$$

the path-integral in Eq. (3.9) can be evaluated by the Gaussian Graßmann integral identity

$$\int \prod_i (\mathrm{d}\xi_i^* \mathrm{d}\xi_i)\; e^{-\xi_i^* S_{ij} \xi_j} = \det S, \tag{3.11}$$

for which a derivation is found in Ref. [2]. This identity holds for any Hermitian matrix S, even if it is not positive definite. The result is then that the partition function in Eq. (3.9) becomes $\mathcal{Z} = \det S$. To calculate this determinant one has to consider the matrix S as also a matrix with time-slice indices. This is perhaps most easily accomplished using the Matsubara formalism in which the τ dependence is substituted with a dependence on Matsubara frequencies through a Fourier-like transform.

3.1.3 Extended Nambu Spinor

From Eq. (3.8) we see that the Nambu spinor sesquilinear product fails to accomodate terms in a Hamiltonian that mix creation and annihilation operators of differing spins, e.g. a term $\propto \hat{c}_\uparrow^\dagger \hat{c}_\downarrow$. In general, a quadratic Hamiltonian can contain any combination of spin-indices of the form $\hat{c}_{s_1}\hat{c}_{s_2}$, $\hat{c}_{s_1}^\dagger \hat{c}_{s_2}$, $\hat{c}_{s_1}\hat{c}_{s_2}^\dagger$ and $\hat{c}_{s_1}^\dagger \hat{c}_{s_2}^\dagger$. This gives in total 16 different combinations and to accomodate them all we thus need a 4×4 matrix. Exchanging to Graßmann numbers we define the vector

$$\boldsymbol{\xi}_\gamma = \begin{pmatrix} \xi_{\gamma,\uparrow}^* \\ \xi_{\gamma,\uparrow} \\ \xi_{\gamma,\downarrow}^* \\ \xi_{\gamma,\downarrow} \end{pmatrix}, \tag{3.12}$$

where all quantum numbers except spin is included in the index γ. Writing the elements of this vector $(\boldsymbol{\xi}_\gamma)_i = \tilde{\xi}_{\gamma,i}$ regardless of whether it is a conjugate or not, we can write all quadratic terms of a Hamiltonian on the bilinear form

$$\boldsymbol{\xi}_\gamma^\mathsf{T} S_{\gamma\delta} \boldsymbol{\xi}_\delta = \tilde{\xi}_{\gamma,i} S_{\gamma i;\delta j} \tilde{\xi}_{\delta,j}, \tag{3.13}$$

where $S_{\gamma\delta}$ is a 4×4 anti-symmetric[2] matrix, and $S_{\gamma i;\delta j}$ denotes its elements. Let there be n number of different quantum numbers, now

2. To see why this matrix can always be said to be anti-symmetric lets first simplify the notation and write the bilinear product as $\tilde{\xi}_i S_{ij} \tilde{\xi}_j$. Then the matrix $S =$

including spin. Then there must be $2n$ different Graßmann generators $\tilde{\xi}_{\gamma,i}$. All of these are integrated over in the discrete version of the partition function field integral

$$\mathcal{Z} = \int \mathcal{D}[\xi^*\, \xi]\; e^{-\int_0^\beta \mathrm{d}\tau\, \tilde{\xi}_{\gamma,i} S_{\gamma i;\delta j} \tilde{\xi}_{\delta,j}}. \tag{3.14}$$

Even though this superficially looks like the field integral in Eq. (3.9), we now have a bilinear and not a sesquilinear form, and S is now a $2n \times 2n$ matrix and not an $n \times n$ matrix. This means that we can not use the integral in Eq. (3.11) to evaluate the integral but instead have to rely on the more general Gaussian Graßmann integral

$$\int \prod_i (\mathrm{d}\tilde{\xi}_i)\; e^{-\frac{1}{2}\tilde{\xi}_i S_{ij} \tilde{\xi}_j} = \mathrm{Pf}(S), \tag{3.15}$$

which applies for any anti-symmetric matrix S. The right hand side is called the Pfaffian $\mathrm{Pf}(S)$ of the matrix S and is defined for any anti-symmetric matrix to be given by

$$\mathrm{Pf}[S] = \frac{1}{2^n n!} \sum_{P \in S_n} (-1)^P S_{P_1 P_2} \cdots S_{P_{n-1} P_n}, \tag{3.16}$$

where P is a permutation in the finite group S_n of all possible permutations of n numbers. This matrix function is related to the determinant by the relation $\mathrm{Pf}(S)^2 = \det(S)$.

Applying the integral identity in Eq. (3.15) to the partition function[3] in Eq. (3.14) after applying the proper discretization of the imaginary

---

$(S + S^\mathsf{T})/2 + (S - S^\mathsf{T})/2$, such that we can write it as a symmetric matrix $\mathcal{S} = (S + S^\mathsf{T})/2$ and an anti-symmetric matrix $\mathcal{A} = (S - S^\mathsf{T})/2$. Considering only the symmetric part of the bilinear form we get

$$\tilde{\xi}_i \mathcal{S}_{ij} \tilde{x} i_j = -\tilde{\xi}_j \mathcal{S}_{ij} \tilde{\xi}_i = -\tilde{\xi}_i \mathcal{S}_{ji} \tilde{\xi}_j = -\tilde{\xi}_i \mathcal{S}_{ij} \tilde{\xi}_j.$$

Hence $\tilde{\xi}_i \mathcal{S}_{ij} \tilde{\xi}_j = 0$ and all that remains is the antisymmetric bilinear form.

3. In relating the discrete version of Eq. (3.15) to (3.14) we have to make sure that the spinor elements $\tilde{\xi}_i$ are defined in terms of ξ_i and ξ_i^* in such a way as to get a correspondence to the sequence of Graßmann generators $\mathrm{d}\xi_i^* \mathrm{d}\xi_i$ in the measure to avoid any sign errors. One solution is to set $\xi_i^* = \tilde{\xi}_{2i-1}$ and $\xi_i = \tilde{\xi}_{2i}$ as we have done in Eq. (3.12). With this definition, then the measure $\int \prod_i \mathrm{d}\xi_i^* \mathrm{d}\xi_i$, which results from the discretized version of the field-integral measure, becomes equal to $\int \prod_{i=1}^{2n} \mathrm{d}\tilde{\xi}_i$ such that Eq. (3.15) can be directly applied.

time then yields the result

$$\mathcal{Z} = \int \mathcal{D}[\xi^*\,\xi]\; e^{-\int_0^\beta \mathrm{d}\tau\, \tilde{\xi}_{\gamma,i} S_{\gamma i;\delta j} \tilde{\xi}_{\delta,j}} = \sqrt{\det(S)}. \tag{3.17}$$

We have chosen the positive result in $\mathrm{Pf}(S) = \pm\sqrt{\det(S)}$ since the partition function $\mathcal{Z}$ needs to be positive on physical grounds. The matrix S to be taken the derminant of on the right hand side of Eq. (3.17) is the full matrix one gets after discretizing the imaginary-time in slices which we usually have done through the Matsubara-frequency formalism.

Now that we know the partition function is given in terms of the determinant of the action-matrix S we can use this information to manipulate the definition of $\tilde{\xi}_{\gamma,i}$ so that we can still write the action as a sesquilinear form. In particular, switching the position of $\xi^*_{\gamma,s}$ and $\xi_{\gamma,s}$ for both spins in the transposed vector on the left of the bilinear form $\boldsymbol{\xi}^{\mathsf{T}}_\gamma S_{\gamma\delta}\boldsymbol{\xi}$, the transposed vector becomes the adjoint vector. This affects the matrix S by switching two pairs of rows. Denoting the matrix where the rows are switched S' then we have rewritten the bilinear form such that

$$\boldsymbol{\xi}^{\mathsf{T}}_\gamma S_{\gamma\delta}\boldsymbol{\xi}_\delta = \boldsymbol{\xi}^\dagger_\gamma S'_{\gamma\delta}\boldsymbol{\xi}_\delta. \tag{3.18}$$

Now the integral over the exponent has not changed since all we have done is simply re-ordering its terms. However, since exchange of rows in a determinant at most produces a minus sign and we do this twice we get that $\det S = \det S'$ and we can write

$$\mathcal{Z} = \int \mathcal{D}[\xi^*\,\xi]\; e^{-\int_0^\beta \mathrm{d}\tau\, \boldsymbol{\xi}^\dagger_\gamma S'_{\gamma\delta}\boldsymbol{\xi}_\delta} = \sqrt{\det(S')}. \tag{3.19}$$

where we need to remember that S' is the row switched matrix of an anti-symmetric matrix S.

3.2 Matsubara formalism

The Matsubara formalism[4] is a way of handling the imaginary time τ dependence of the coherent state eigenvalue fields $\xi_\alpha(\tau)$, where α denotes a collection of quantum numbers sufficient to specify a state, without having to go back to the time-sliced path-integral. It also lets

4. Named after the Japanese physicst Matsubara, Takeo.

us automatically satisfy the imaginary-time boundary conditions $\xi_\alpha(0) = \zeta\xi_\alpha(\beta)$, where $\zeta = +1$ for bosons and $\zeta = -1$ for fermions. Imagining that τ is a continuous variable as suggested in the path-integral notation, we define two countable infinite sets of new field-variables through the Fourier-transforms

$$\xi_{\alpha,n} = \frac{1}{\sqrt{\beta}} \int_0^\beta \mathrm{d}\tau \, e^{i\omega_n \tau} \xi_\alpha(\tau), \tag{3.20a}$$

$$\xi^*_{\alpha,n} = \frac{1}{\sqrt{\beta}} \int_0^\beta \mathrm{d}\tau \, e^{-i\omega_n \tau} \xi^*_\alpha(\tau). \tag{3.20b}$$

The frequencies ω_n are called Matsubara frequencies and are defined by $\omega_n = (2n+1)\pi/\beta$ with $n \in \mathbb{Z}$ for fermions. For bosons we use instead the notation ν_n where $\nu_n = 2n\pi/\beta$. The inverse relations are given by

$$\xi_\alpha(\tau) = \sum_{n\in\mathbb{Z}} e^{-i\omega_n \tau} \xi_{\alpha,n}, \tag{3.21a}$$

$$\xi^*_\alpha(\tau) = \sum_{n\in\mathbb{Z}} e^{i\omega_n \tau} \xi^*_{\alpha,n}. \tag{3.21b}$$

3.2.1 Matsubara sums

When the Matsubara formalism $\xi_{\alpha,n}$ is used for the field variables in the action of a partition-function field integral we will often need to evaluate infinite sums of Matsubara frequencies of the form

$$\sum_n h(if_n), \tag{3.22}$$

where f_n is either a fermionic or bosonic Matsubara frequency, to evaluate the field integral. A useful strategy in such evaluations is to transform the sum to a complex integral by using reverse residue integration. We consider the complex contour integral along a path $\mathcal{C}$ running counter-clockwise around the complex plane infinitesimally close to the imaginary axis as shown in Fig. 3.1. The integrand we consider is given by the product of the summand and the complex continuation of the Fermi-Dirac- or Bose-Einstein-distribution function

$$n_\zeta(z) = (e^{\beta z} - \zeta)^{-1}, \quad z \in \mathbb{C}, \tag{3.23}$$

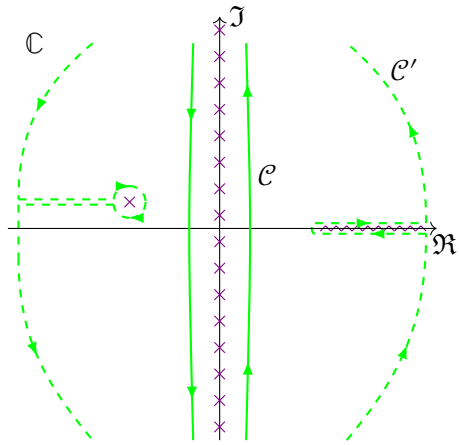

Figure 3.1: Integration contour for the Matsubara sum $\sum_n h(if_n)$. The contour is marked by a solid line and is imagined to continue to $\pm i\infty$. The crosses along the imaginary axis symbolizes the simple poles of the Fermi-Dirac distribution-function. A deformed integration contour is shown with dashed lines that is imagined to cross the real axis at $\pm\infty$. This contour then encloses a simple pole on the left and a branch cut on the right belonging to the summand.

depending on whether the Matsubara frequency in the sum is of fermionic ($\zeta = +1$) or bosonic ($\zeta = -1$) nature. This function has simple poles[5] at $z = if_n$ and thus integration around the contour results in a sum of residues of the integrand at these poles such that we get

$$\sum_n h(if_n) = \frac{\zeta\beta}{2\pi i}\oint_{\mathcal{C}} \mathrm{d}z\; h(z)n_\zeta(z), \tag{3.24}$$

given that $h(z)$ does not contain any poles at these points. The contour can now be continuously deformed at will as long as it does not cross any singularities which can greatly facilitate the calculation of the integral. The default approach is to see if the integrand vanishes as $|z| \to \infty$, in which case it is usually useful to expand the contour as much as possible as illustrated by the deformed contour $\mathcal{C}'$ in Figure 3.1.

5. That the poles are simple, i.e. 1st order, is easily seen by expanding the exponential around if_n to leading order.

Using the method outlined above, we may calculate the sums

$$\sum_{n\in\mathbb{Z}} \frac{1}{if_n - x} = -\zeta\beta n_\zeta(x), \tag{3.25a}$$

$$\sum_{n\in\mathbb{Z}} \frac{1}{(i\omega_n - x)(i\omega_n - y)} = \frac{\beta}{x-y}\left(\frac{1}{1+e^{\beta x}} - \frac{1}{1+e^{\beta y}}\right), \tag{3.25b}$$

$$\sum_{n\in\mathbb{Z}} \ln[\beta(i\omega_n + x)] = \ln(1+e^{-\beta x}). \tag{3.25c}$$

3.3 Hubbard-Stratonovich transformation

The Hubbard-Stratonovich transformation is a transformation in the fields of a theory where a new complex field is introduced in order to covert a term that is square and thus non-linear in an existing field variable, into a linear term in this variable that is coupled to the new field. This is particularly useful when the existing field is Fermionic and thus a Graßmann variable since it makes it possible to consider low energy exitations of the theory using e.g. a saddle-point approximation. It is however important to point out that the transformation itself is not in any way approximative, but is an exact transformation that maintains all information in the original theory.

In technical terms, the Hubbard-Stratonovich transformation can be viewed simply as the solution of a complex multivariate integral. Let A have a strictly positive Hermitian part and $\boldsymbol{J}$ be a vector of coefficients that could contain Graßmann variables or complex variables. Then

$$e^{\boldsymbol{J}^\dagger A \boldsymbol{J}} = \det A^{-1} \int_{\mathbb{C}} \prod_i \left[\frac{\mathrm{d}z_i^* \mathrm{d}z_i}{2\pi i}\right] e^{-(\boldsymbol{z}^\dagger A^{-1}\boldsymbol{z} + \boldsymbol{z}^\dagger \boldsymbol{J} + \boldsymbol{J}^\dagger \boldsymbol{z})}, \tag{3.26}$$

exchanges a quadratic term in $\boldsymbol{J}$ with a integration over the complex $\boldsymbol{z}$ variable. Since $\boldsymbol{J}$ usually represents some field in a field theory, the new $\boldsymbol{z}$ is called the auxhillary field or the conjugate field because of its linear coupling to $\boldsymbol{J}$. In the less general case that A is an Hermitian matrix, this formula is proved simply by completing the square, then diagonalizing A by a unitary tranformation and calculating the resulting integrals by the formula $\int_{\mathbb{C}} \mathrm{d}z^* \mathrm{d}z\, e^{-azz^*} = 2\pi i/a$.

From Eq. (3.26) we see that what we have to do to perform the Hubbard-Stratonovich transformation is first to make a choice for what to interpret as part of the matrix A and what to interpret as part of $\boldsymbol{J}$.

We then have to check that this definition of A leads to its Hermitian part having only positive eigenvalues. Finally we need to know an analytical expression for its inverse. It is usually the first step that is the most difficult since this dictates the low energy excitation a subsequent saddle point approximation or a stationary phase approximation will produce. Typically we are interested in transforming a Fermionic interaction potential of the form

$$V = \frac{1}{2} \sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta} \xi^*_\alpha \xi^*_\beta \xi_\delta \xi_\gamma, \tag{3.27}$$

where ξ_α are Graßmann variables, which can be sketched in the way of single-vertex diagram in Figure 3.2. The HS-transformation is clas-

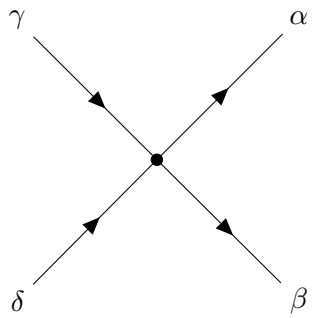

Figure 3.2: Generic two-body interaction.

sified into being done in a specific *channel* depending on which pair of Graßmann variables are considered to be part of $\boldsymbol{J}$ and consequently $\boldsymbol{J}^\dagger$. The direct channel[6] is given by the identification $J_i \sim \xi^*_\alpha \xi_\gamma$, the Cooper channel[7] is defined by the identification $J_i \sim \xi_\delta \xi_\gamma$ while the exchange channel is given by the identification $J_i \sim \xi^*_\alpha \xi_\delta$. Depending on exactly how $\boldsymbol{J}$ is chosen, the Gaussian integral in Eq. (3.26) might have to be modified. For example in the case of the direct- and exchange-channel, the exponential argument on the left side will have the form $\boldsymbol{J}^\mathsf{T} A \boldsymbol{J}$ which necessitates the Gaussian integral identity

$$e^{-\frac{1}{2}\boldsymbol{J}^\mathsf{T} A \boldsymbol{J}} = \sqrt{\det A^{-1}} \int_{\mathbb{R}} \prod_i \left[\frac{\mathrm{d}x_i}{\sqrt{2\pi}} \right] e^{-\frac{1}{2}\boldsymbol{x}^\mathsf{T} A^{-1} \boldsymbol{x} - i \boldsymbol{J}^\mathsf{T} \boldsymbol{x}}, \tag{3.28}$$

6. Also known as the density-density channel.
7. Also known as the particle-particle channel.

where the auxhillary field $\boldsymbol{x}$ now is a real conjugate field.

3.3.1 Transformation in symmetry channels

In the Cooper-channel of the Hubbard-Stratonovich transformation, the complex field $\boldsymbol{z}$ is conjugate to some combination of pairs of annihilation operators $\hat{c}_\delta \hat{c}_\gamma$ (or their corresponding Graßmann variables). The symmetry of the specific combination in turn then determines the symmetry of any low energy field theory obtained through a subsequent stationary phase approximation. By diagonalizing the interaction potential $\widehat{V}$ into its particular irreducible representations as we did in Section 4.9 then a Hubbard-Stratonovich transformation in a specific symmetry channel is done by identifying $\boldsymbol{J}$ with irrep. basis functions.

Lets take the case of a BCS theory of superconductivity where the interaction can be written in terms of basis functions $d^{(b),m}_{s_1 s_2}(\boldsymbol{k})$ such that the diagonalized interaction takes the form

$$\hat{V} = \sum d^{(b),m}_{s_1 s_2}(\boldsymbol{k})^* v_{(b)} d^{(b),m}_{s_1' s_2'}(\boldsymbol{k}') c^\dagger_{\frac{\boldsymbol{q}}{2}+\boldsymbol{k}\, s_1} c^\dagger_{\frac{\boldsymbol{q}}{2}-\boldsymbol{k}\, s_2} c_{\frac{\boldsymbol{q}}{2}-\boldsymbol{k}' s_2'} c_{\frac{\boldsymbol{q}}{2}+\boldsymbol{k}' s_1'}, \tag{3.29}$$

where $\sum$ indicates the sum over the indices, $\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{q}, s_1, s_2, s_1', s_2', b$ and m. Here b specifies the irreducible representation while m enumerates the representation basis. Identifying

$$\hat{J}^{(b_m)}_{\boldsymbol{q}} = \sum_{\boldsymbol{k} s_1 s_2} d^{(b),m}_{s_1 s_2}(\boldsymbol{k}) \hat{c}_{\frac{\boldsymbol{q}}{2}-\boldsymbol{k}, s_1} \hat{c}_{\frac{\boldsymbol{q}}{2}+\boldsymbol{k}, s_2}, \tag{3.30}$$

the interaction potential is simply written

$$\hat{V} = \sum_{\boldsymbol{q}, b, m} \hat{J}^{(b_m)\,\dagger}_{\boldsymbol{q}} v^{(b)} \hat{J}^{(b_m)}_{\boldsymbol{q}}. \tag{3.31}$$

In the path-integral representation of the partition function, the annihilation operators become Graßmann variables which we denote by writing J instead of $\hat{J}$ such that the contribution from the interaction potential results in the exponential

$$\mathcal{Z}_I = e^{-\int_0^\beta \mathrm{d}\tau \sum_{\boldsymbol{q}, b, m} J^{(b_m)\,\dagger}_{\boldsymbol{q}} v^{(b)} J^{(b_m)}_{\boldsymbol{q}}}. \tag{3.32}$$

Now it is straight forward to use the Hubbard-Statonovich formula

$$e^{\int_0^\beta \mathrm{d}\tau \sum_{ij} J_i^* A_{ij} J_j} = \int \mathcal{D}[\eta_i^* \eta_i] e^{-\int_0^\beta \mathrm{d}\tau \left(\eta_i^* A^{-1}_{ij} \eta_j + J_i^* \eta_i + J_i \eta_i^*\right)}, \tag{3.33}$$

which is a path integral version of Eq. (3.26), to transform each pair of irreducible representation basis vectors to individual conjugate fields. In the notation of Eq. (3.33) implicit summation over repeated indices is used and each index i is a collection $i = (b, m, \boldsymbol{q})$ of indices. Comparing Eq. (3.33) and (3.32) we gather that

$$A_{ij} = A_{b,m,\boldsymbol{q};b',m',\boldsymbol{q}'} = -\delta_{\boldsymbol{q}\boldsymbol{q}'}\delta_{mm'}\delta_{bb'}v^{(b)}, \tag{3.34}$$

which is trivially Hermetic and positive definite provided $v^{(b)} < 0$. In this case we say that the irreducible representation b is an attractive channel. A is in this case also trivially invertible with $A^{-1}_{ij} = -\delta_{ij}/v^{(b)}$. Writing out all the indices we finally arrive at the Hubbard-Statonovich transformation of the interaction potential in individually attractive symmetry channels

$$\mathcal{Z}_I = \int \mathcal{D}[\eta^{(b_m)*}_{\boldsymbol{q}}\eta^{(b_m)}_{\boldsymbol{q}}] e^{\int_0^\beta \mathrm{d}\tau \sum_{\boldsymbol{q}bm} \left[\frac{|\eta^{(b_m)}_{\boldsymbol{q}}|}{v^{(b)}} - \left(J^{(b_m)*}_{\boldsymbol{q}}\eta^{(b_m)}_{\boldsymbol{q}} + J^{(b_m)}_{\boldsymbol{q}}\eta^{(b_m)}_{\boldsymbol{q}}\right)\right]}, \tag{3.35}$$

where

$$J^{(b_m)}_{\boldsymbol{q}} = \sum_{\boldsymbol{k}s_1s_2} d^{(b),m}_{s_1s_2}(\boldsymbol{k})\xi_{\frac{\boldsymbol{q}}{2}-\boldsymbol{k},s_1}\xi_{\frac{\boldsymbol{q}}{2}+\boldsymbol{k},s_2} \tag{3.36}$$

in terms of Graßmann variables ξ. We note that this derivation does not assume either odd or even basis functions for the irreducible representations and works just as well for either.

3.4 Field theory approximations

3.4.1 Stationary phase and the one loop expansion

Let Z be the partition function

$$Z(l) = \int \mathcal{D}[\eta^*_\alpha\, \eta_\alpha]\; e^{-lS(\eta^*_\alpha,\eta_\alpha)}, \tag{3.37}$$

given in terms of bosonic fields $\eta^*_\alpha, \eta_\alpha$ and where l is some large parameter $l \gg 1$. In what is called the stationary phase approximation[8]

8. The name is misleading in the case of the many-particle partition function since we do not have a strict phase in the exponent but in general a complex function because $\eta_\alpha \in \mathbb{C}$. Technically it is the method of steepest descent in the case of an exponent with stationary points that is used in this case.

of a bosonic field integral we create an expansion of the free energy around the field configuration $\{\eta_\alpha^c\}$ where the action S is stationary and a minimum. This configuration is the main contribution to the integral in Eq. (3.37) since it provides the maximum of the exponent, and determines the leading order asymptotic behavior as $l \to \infty$. It also corresponds in a sense to a classical solution and gives in the case of the feynman path integral for the evolution operator of a single particle in an external potential, the classical Euler-Lagrange equations. The configuration is found in the path-integral notation by varying the fields in the action such that

$$\frac{\delta S}{\delta \eta_\alpha} = 0 \quad \wedge \quad \frac{\delta S}{\delta \eta_\alpha^*} = 0. \tag{3.38}$$

As an example, the Hubbard-model can be expressed by a bosonic field integral through a Hubbard-Statonovich transformation. The stationary field configuration $\{\eta_q^c(\tau)\}$ is in this case given by a imaginary-time- and spatialy-independent field-configuration η^c. Assuming such a solution, the action reduces to

$$S(\eta^{c\,*}, \eta^c) = \frac{\beta N}{U}|\eta^c|^2 - \sum_{\boldsymbol{q}} \ln\left[\left(1 + e^{\beta E_{\boldsymbol{q}}}\right)\left(1 + e^{-\beta E_{\boldsymbol{q}}}\right)\right], \tag{3.39}$$

where $E_{\boldsymbol{q}} = \sqrt{(\epsilon(\boldsymbol{q}) - \mu)^2 + |\eta^c|^2}$, N is the number of hopping sites, $\epsilon(\boldsymbol{q})$ is the Fourier transformed kinetic hopping energy, μ is the chemical potential, and U is the repulsive Hubbard interaction strength. The stationary action condition in Eq. (3.38) then yields the equation

$$|\eta^c|\left(\frac{1}{U} - \frac{1}{N}\sum_{\boldsymbol{q}} \frac{1}{2E_{\boldsymbol{q}}} \tanh\frac{\beta E_{\boldsymbol{q}}}{2}\right) = 0, \tag{3.40}$$

which has two solution: one given by $\eta^c = 0$ and one given by setting the terms in parenthesis equal to 0. The last non-trivial solution is a form of the BCS solution and represents the order-parameter in an s-wave superconductor. Because of the parameter β, this solution will be temperature dependent and dissapears at some critical temperature at which the two solutions for η_c converge.

The stationary phase expansion is then the expansion resulting from expanding the action around the stationary solution. Setting $\tilde{\eta}_\alpha =$

$\sqrt{l}(\eta_\alpha - \eta_\alpha^c)$ we get in general the expansion

$$\begin{aligned} S(\tilde{\eta}_\alpha, \tilde{\eta}_\alpha^*) &= S_c + \frac{1}{l}\Bigg[\frac{1}{2!}\frac{\delta^2 S}{\delta\eta_\alpha \delta\eta_\beta}\Big|_{\eta_c} \tilde{\eta}_\alpha \tilde{\eta}_\beta \\ &\quad + \frac{1}{2!}\frac{\delta^2 S}{\delta\eta_\alpha^* \delta\eta_\beta^*}\Big|_{\eta_c} \tilde{\eta}_\alpha^* \tilde{\eta}_\beta^* + \frac{\delta^2 S}{\delta\eta_\alpha \delta\eta_\beta^*}\Big|_{\eta_c} \tilde{\eta}_\alpha \tilde{\eta}_\beta^*\Bigg] + \mathcal{O}\Big(\frac{1}{l\sqrt{l}}\Big) \\ &= \sum_{n_1+n_2\geq 2} \frac{1}{n_1! n_2!} \frac{\delta^{n_1+n_2} S}{\delta\tilde{\eta}_\alpha^{n_1} \delta\tilde{\eta}_\beta^{*\, n_2}} \tilde{\eta}_\alpha^{n_1} \tilde{\eta}_\beta^{*\, n_2}\Big|_{\eta_c} \frac{1}{l^{(n_1+n_2)/2}}. \end{aligned} \tag{3.41}$$

We have used implicit summation over repeated indices and on the last line we have assumed that new indices should be introduced and summed for each n_1 and n_2. This is simply the multivariate Taylor expansion around η_c where we have treated η and η^* as independent variables.

A simpler expansion can be found when the bosonic fields correspond to order parameters in a system close to a phase transition. In this case we can assume the fields η_α to be small in general such that S simply can be expanded about 0. If the system is fermionic, then the bosonic field integral of the order-parameters will have resulted from a Hubbard-Stratonovich transformation, in which case we will have a contribution to the integral of the form

$$\sqrt{\det G^{-1}(\eta_\alpha, \eta_\alpha^*)} = e^{\frac{1}{2}\operatorname{Tr}\ln G^{-1}}, \tag{3.42}$$

where G^{-1} is the result of the integration of a quadratic fermionic action, and will in general depend on the auxhillary bosonic fields η_α and η_α^* in a linear way. Then the matrix G^{-1} can be decomposed in a matrix G_0^{-1} that results purely from fermionic integration, and a matrix ϕ that is dependent on the auxhillary fields η_α and importantly vanishes as $\eta_\alpha \to 0$, such that $G^{-1} = G_0^{-1} + \phi$. When the system is close the the phase transition such that the auxhillary fields are small, then this allows for the expansion of the logarithm in Eq. (3.42) such that

$$\begin{aligned} \frac{1}{2}\operatorname{Tr}\ln(G_0^{-1} + \phi) &= \frac{1}{2}\Big(\operatorname{Tr}\ln G_0^{-1} + \operatorname{Tr}\ln(1 + G_0\phi)\Big) \\ &= \frac{1}{2}\Big(\operatorname{Tr}\ln G_0^{-1} - \sum_{n=1}^{\infty} \frac{\operatorname{Tr}(-G_0\phi)^n}{n}\Big). \end{aligned} \tag{3.43}$$

This is known as the one-loop expansion since in terms of perturbation theory, G_0 is the fermionic propagator such that the sum in the last line of Eq. (3.43) corresponds to a series of propagators and interactions that are connected in a closed loop by the trace.

3.4.2 Gradient expansion

The gradient expansion rests on the assumption that the fields are sufficiently smooth such that progressively higher order derivatives with respect to the field parameters, are progressively smaller, i.e. we assume

$$l_\alpha^n |\partial_\alpha^n \eta_\alpha| \gg l_\alpha^{n+1} |\partial_\alpha^{n+1} \eta_\alpha|, \tag{3.44}$$

where l_α is some appropriate length scale such as to make $l_\alpha \partial_\alpha$ dimensionless. In practice this usually means that given a momentum-dependent action density, $S(\eta^*_{\boldsymbol{q},\alpha}, \eta_{\boldsymbol{q},\alpha}; \boldsymbol{q})$, we assume q_i small compared to the lattice spacing[9] a_i. Then we expand the explicit $\boldsymbol{q}$ dependence in the action density in the small parameters $a_i q_i$ as a McLaurien series. In the following notation, we will assume the length parameter a^i is present in the notation q^i where appropriate. The momentum dependence of the fields themselves should not be expanded as our goal is terms of the form $(q^i)^n (\tilde{\eta}_{\boldsymbol{q},\alpha})^m$. If we then let partial momentum-derivatives only act on the explicite momentum-dependence in S, and neglecting the field dependence in the notation for S, the expanded action can be written as the series

$$S(\boldsymbol{q}) = S(0) + \frac{\partial S}{\partial q^i}(0)\, q^i + \frac{1}{2!} \frac{\partial^2 S}{\partial q^i \partial q^j}(0)\, q^i q^j + \mathcal{O}(q^3). \tag{3.45}$$

Usually, the linear term cancels by symmetry of the underlying lattice. In general it is smart to here check for terms that cancel by considering any internal momentum sums that may be included in the coefficients.

Terms with products between fields $\eta_{\boldsymbol{q},\alpha}$ and q^i lead to gradients of the spatially dependent fields $\eta_{\boldsymbol{R},\alpha}$, which is why the expansion in Eq. (3.45) can be called a gradient expansion. The spatially dependent fields are defined as the coefficients in the inverse Fourier transform

$$\eta_{\boldsymbol{q},\alpha} = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}} e^{i\boldsymbol{q}\cdot\boldsymbol{R}} \eta_\alpha(\boldsymbol{R}), \tag{3.46}$$

9. This could be the spacing between hopping sites, i.e. stationary ions in an electron model.

where N is the number of terms in the sums $\sum_{\boldsymbol{q}}$ and $\sum_{\boldsymbol{R}}$. One way of now obtaining gradients of the spatial fields is to realize that since q^i is small we can set

$$q^i \approx \sin(q^i) = \frac{1}{2i}\Big(e^{iq^i} - e^{-iq^i}\Big). \tag{3.47}$$

With this identification, all the momentum dependence in $S(\boldsymbol{q})$ exists as phases such that the sum $\sum_{\boldsymbol{q}} S(\boldsymbol{q})$ results in a series of Kronicker-delta functions which we evaluate by the $\sum_{\boldsymbol{R}}$ sums coming from the inverse Fourier transforms in Eq. (3.46). Grouping terms of displaced spatial fields, we can identify derivaties, such that

$$\eta_\alpha(\boldsymbol{R} + a\hat{e}_i) - \eta_\alpha(\boldsymbol{R}) \approx a\frac{\partial}{\partial R^i}\eta_\alpha(\boldsymbol{R}). \tag{3.48}$$

These identifications are justified by going to the continuum limit where $a \to 0$.

As an example, consider the sum

$$S = \sum_{\boldsymbol{q}} K_{\alpha\beta ij} q^i q^j \eta^*_{\boldsymbol{q},\alpha}\eta_{\boldsymbol{q},\beta}, \tag{3.49}$$

where there is an implicit summation over repeated indices i, j, α and β. Fourier transforming the fields according to Eq. (3.46) and writing q^i as $\sin q^i$, we can group terms such that

$$\begin{aligned} S &= \frac{1}{N}\sum_{\boldsymbol{R}_1\boldsymbol{R}_2} K_{\alpha\beta ij}\eta_\alpha(\boldsymbol{R}_1)^*\eta_\beta(\boldsymbol{R}_2)\frac{1}{4}\sum_{\boldsymbol{q}}\Big[e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_2-\boldsymbol{R}_1+a\hat{e}_i-a\hat{e}_j)} \\ &+e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_2-\boldsymbol{R}_1+a\hat{e}_j-a\hat{e}_i)} - e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_2-\boldsymbol{R}_1+a\hat{e}_i+a\hat{e}_j)} - e^{i\boldsymbol{q}\cdot(\boldsymbol{R}_2-\boldsymbol{R}_1-a\hat{e}_i-a\hat{e}_j)}\Big] \\ &= \sum_{\boldsymbol{R}} K_{\alpha\beta ij}a^2\frac{\partial}{\partial R_i}\eta_\alpha(\boldsymbol{R})^*\frac{\partial}{\partial R_j}\eta_\beta(\boldsymbol{R}). \end{aligned} \tag{3.50}$$

A more conventional way of converting to gradients, is to use integration by parts. Then the product rule of partial derivatives is used to obtain

$$q^j\eta_\alpha(\boldsymbol{R})e^{i\boldsymbol{q}\cdot\boldsymbol{R}} = i\frac{\partial\eta_\alpha}{\partial R^i}e^{i\boldsymbol{q}\cdot\boldsymbol{R}} - i\frac{\partial}{\partial R^i}\Big[\eta_\alpha(\boldsymbol{R})e^{i\boldsymbol{q}\cdot\boldsymbol{R}}\Big]. \tag{3.51}$$

Summing on both sides and arguing that the bounary term vanishes because $\eta_\alpha(\boldsymbol{R}) \to 0$ as $R^i \to \infty$, then

$$\sum_{\boldsymbol{R}} q^j\eta_\alpha(\boldsymbol{R})e^{i\boldsymbol{q}\cdot\boldsymbol{R}} = i\sum_{\boldsymbol{R}} e^{i\boldsymbol{q}\cdot\boldsymbol{R}}\nabla_i\eta_\alpha(\boldsymbol{R}). \tag{3.52}$$

Chapter 4

Group Theory

In this chapter we will introduce the minimum mathematical framework needed to understand the phrase "This free energy is the Γ_{5u}^{-} irreducible representation of the D_{4h} symmetry group".

A few words about notation. We will use the semicolon ';' in equations as notation for the words 'such that', e.g. when defining sets. A colon ': ' is used when defining maps where the symbol representing the mapping itself should be on the left while the the sets being related or how the elements of the sets are related is on the right of the colon. The colon ':' is also used as a shortcut for the words 'applied through its representation to' for when group elements are applied to vectors, where the correct representation to use for this application should be implicitely understood.

4.1 Discrete groups

4.2 Irreducible representations

To know what an irreducible representation is, let's start with what we mean by a reducible representation.

Def. 4.1. *A matrix representation is **reducible** if there exists a non-trivial invariant subspace of the vector space of the representation.*

The intuition is then that the vector space of the representation is reducible if a "smaller" representation is contained within it. Since

there is a smaller vector space within the vector space of the original representation and this vector space is invariant, it is possible to define another representation on this smaller vector space, i.e. *reduce* the original representation. We have now used the word "invariant" a couple of times, so let's define what it means more precisely.

Def. 4.2. *Let $D(g)$ a representation of the group G on the vector space V such that $D(g) : V \to V$. Then a subspace $U \subseteq V$ is **invariant** if*

$$\forall g \in G \quad u \in U \implies D(g)u \in U. \tag{4.1}$$

In other words: a vector space is invariant if it is not possible for any vector in it to escape using a representation of any group element. All representations applied to any vector in the invariant subspace must necessarily land in that same subspace from which it started.

Thm. 4.1 (Shur's Lemma)**.** *If $D(g)$ is an irreducible complex representation with vector space V and $L(V)$ is the set of all linear maps, then*

$$\{A \in L(V);\ AD(g) = D(g)A\ \forall g \in G\} = \{c\mathbb{1};\ c \in \mathbb{C}\}. \tag{4.2}$$

4.3 BCS Hilbert Space

We define the BCS Hilbert space as the Hilbert space upon which BCS-type potentials operate. Specifically this is a reduced form of the two-state fermionic product Hilbert space $\mathcal{H}_2 = \mathcal{H} \otimes \mathcal{H}$ where $\mathcal{H} = \text{span}\{|\boldsymbol{k}, s\rangle\}$ and we only consider states that have opposite momentum. Thus this Hilbert space is given by

$$\mathcal{B} = \text{span}\{|\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle\}, \tag{4.3}$$

and the identity operator in this space can be written

$$\hat{\mathbb{1}} = \sum_{\boldsymbol{k}\, s_1 s_2} |\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle\langle-\boldsymbol{k}, s_2|\langle\boldsymbol{k}, s_1|. \tag{4.4}$$

Acting on the arbitrary vector $|v\rangle \in \mathcal{B}$ with this identity operator, we find that in terms of this basis it can be written

$$|v\rangle = \sum_{\boldsymbol{k}\, s_1 s_2} v_{s_1 s_2}(\boldsymbol{k})|\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle, \tag{4.5}$$

where

$$v_{s_1 s_2}(\boldsymbol{k}) = \langle -\boldsymbol{k}, s_2 | \langle \boldsymbol{k}, s_1 \mid v \rangle. \tag{4.6}$$

The indices s_1 and s_2 can take on only two values each, namely $s_1, s_2 \in \{\uparrow, \downarrow\}$. In total there are thus 4 different realizations of pairs, $s_1 s_2$ e.g. $\uparrow\uparrow$ for $v_{s_1 s_2}(\boldsymbol{k})$. Putting these different realizations of $v_{s_1 s_2}(\boldsymbol{k})$ as elements in a 2×2 matrix we get

$$v_{s_1 s_2}(\boldsymbol{k}) = \begin{pmatrix} v_{\uparrow\uparrow}(\boldsymbol{k}) & v_{\uparrow\downarrow}(\boldsymbol{k}) \\ v_{\downarrow\uparrow}(\boldsymbol{k}) & v_{\downarrow\downarrow}(\boldsymbol{k}) \end{pmatrix}. \tag{4.7}$$

Any 2×2 matrix can be written in the conventional basis of the 4 Pauli matrices $\sigma^0 = \mathbb{1}_{2\times 2}$, σ^x, σ^y, and σ^z. This means that we could write the matrix in Eq. (4.7)

$$v_{s_1 s_2}(\boldsymbol{k}) = v^0_{\boldsymbol{k}} \sigma^0_{s_1 s_2} + v^i_{\boldsymbol{k}} \sigma^i_{s_1 s_2}. \tag{4.8}$$

It is however conventional to factor out a Pauli matrix $i\sigma^y$ to the right in the expansion since this results in nice transformation properties of the coefficients as we shall see. With the spin-indices expanded in this basis it is conventional to let the function of $\boldsymbol{k}$ that is in front of σ^0 be called $\psi_{\boldsymbol{k}}$. The three others are conventionally denoted $d_{\boldsymbol{k},i}$. Expanded in this conventional basis then $v_{s_1 s_2}(\boldsymbol{k})$ takes the form

$$v_{s_1 s_2}(\boldsymbol{k}) = (\psi_{\boldsymbol{k}} \sigma^0_{s_1 s'} + d_{\boldsymbol{k},i} \sigma^i_{s_1 s'}) i\sigma^y_{s' s_2}, \tag{4.9}$$

and finally the state $|v\rangle$ can be written

$$|v\rangle = \sum_{\boldsymbol{k}\, s_1 s_2} [(\psi_{\boldsymbol{k}} \sigma^0 + \boldsymbol{d}_{\boldsymbol{k}} \cdot \boldsymbol{\sigma}) i\sigma^y]_{s_1 s_2} |\boldsymbol{k}, s_1\rangle |-\boldsymbol{k}, s_2\rangle. \tag{4.10}$$

Going one step back and writing out the different combinations of $s_1 s_2$ in $v_{s_1 s_2}(\boldsymbol{k})$ as a matrix like we did in Eq. (4.7), but now multiplying out the Pauli matrices in Eq. (4.9) we get

$$\begin{pmatrix} v_{\uparrow\uparrow}(\boldsymbol{k}) & v_{\uparrow\downarrow}(\boldsymbol{k}) \\ v_{\downarrow\uparrow}(\boldsymbol{k}) & v_{\downarrow\downarrow}(\boldsymbol{k}) \end{pmatrix} = \begin{pmatrix} -d_{\boldsymbol{k},x} + i d_{\boldsymbol{k},y} & \psi_{\boldsymbol{k}} + d_{\boldsymbol{k},z} \\ -\psi_{\boldsymbol{k}} + d_{\boldsymbol{k},z} & d_{\boldsymbol{k},x} + i d_{\boldsymbol{k},y} \end{pmatrix}. \tag{4.11}$$

This set of linear relations is easily inverted to yield

$$\psi_{\boldsymbol{k}} = \frac{1}{2}(v_{\uparrow\downarrow}(\boldsymbol{k}) - v_{\downarrow\uparrow}(\boldsymbol{k})) \tag{4.12}$$

$$d_{\boldsymbol{k},x} = \frac{1}{2}(v_{\downarrow\downarrow}(\boldsymbol{k}) - v_{\uparrow\uparrow}(\boldsymbol{k})) \tag{4.13}$$

$$d_{\boldsymbol{k},y} = -\frac{i}{2}(v_{\uparrow\uparrow}(\boldsymbol{k}) + v_{\downarrow\downarrow}(\boldsymbol{k})) \tag{4.14}$$

$$d_{\boldsymbol{k},z} = \frac{1}{2}(v_{\uparrow\downarrow}(\boldsymbol{k}) + v_{\downarrow\uparrow}(\boldsymbol{k})). \tag{4.15}$$

Since the space $\mathcal{B}$ is fermionic we have the symmetry

$$|\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle = -|-\boldsymbol{k}, s_2\rangle|\boldsymbol{k}, s_1\rangle. \tag{4.16}$$

Using this symmetry transformation on the basis vectors in the expansion of $|v\rangle$ in Eq. (4.6), then renaming indices and finally equating coefficients term by term, we see that for the coefficients of $|v\rangle$, this symmetry takes the form

$$v_{s_1 s_2}(\boldsymbol{k}) = -v_{s_2 s_1}(-\boldsymbol{k}). \tag{4.17}$$

4.4 Application of group elements

When we are talking about applying some symmetry transformation to a state, this is synonymous with applying a group element to a vector. Even more specifically, the 'applying' part means that we have some natural representation of the group on the vector space we have defined states on, and we are using the linear transformation of the representation of the group element to act on the state vector. In this thesis we will use the notation $g : |\psi\rangle$ to refer to this procedure.

Let g be an arbitrary group element in the symmetry group G and D be a representation of G on the d-dimensional vector space V. Let V have a basis $\{\boldsymbol{b}_i\}_{i=1}^{d}$. The application of a group element to a basis vector is then defined as

$$g : \boldsymbol{b}_i = \sum_j \boldsymbol{b}_j D_{ji}(g). \tag{4.18}$$

The application of a group element to any vector in V then is calculated by expanding the vector in the basis and applying the representation D to each basis vector separately as a linear transformation:

$$g : \boldsymbol{v} = \sum_i v_i g : \boldsymbol{b}_i. \tag{4.19}$$

4.4.1 Active vector transformation

As g : has been defined in Eq. (4.18) it is defined in a passive perspective where the transformation is happening to the basis vectors. Applying g to a vector $\boldsymbol{v}$ in the basis $\{\boldsymbol{b}_i\}$ it is sometimes useful to consider this as an application not on the vectors themselves but on the expansion coefficients v_i of $\boldsymbol{v}$ in the basis. This is the active view of the transformation. Inserting Eq. (4.18) into Eq. (4.19) we get

$$\begin{aligned} g : \boldsymbol{v} &= \sum_i v_i \sum_j \boldsymbol{b}_j D_{ji}(g) = \sum_j \Big(\sum_i D_{ji}(g) v_i \Big) \boldsymbol{b}_j \\ &= \sum_i v_i' \boldsymbol{b}_i, \end{aligned} \tag{4.20}$$

where we have defined the transformed coefficients

$$v_i' = \sum_j D_{ij}(g) v_j. \tag{4.21}$$

From this calculation we see that we can consider the application of g as a transformation of the coefficients of the vector as

$$g : v_i = \sum_j D_{ij}(g) v_j. \tag{4.22}$$

This defines the active transformation of a vector v by a group element g.

4.4.2 Representation on product spaces

The product space $V \otimes V$ then has a basis $\{\boldsymbol{b}_i \boldsymbol{b}_j\}_{i,j=1}^d$. A derived representation can be constructed from D on this product space called the product representation $D^{(D\times D)}(g)$. This is defined through its application on the basis by

$$g : \boldsymbol{b}_i \boldsymbol{b}_j = \sum_{kl} \boldsymbol{b}_k \boldsymbol{b}_l \big[D^{(D\times D)}(g)\big]_{kl,ij} = \sum_{kl} \boldsymbol{b}_k \boldsymbol{b}_l D_{ki}(g) D_{lj}(g). \tag{4.23}$$

To apply group theory to physical problems, we need to know how the objects we are working with in the physical theory transform under group elements. The most important vector space in quantum mechanics is arguably the Hilbert space where particle states are determined by a momentum and spin quantum number. Each quantum

number has its own vector space defined by the basis vectors $|\boldsymbol{k}\rangle$ and $|s\rangle$ in the Dirac notation. The combination of both quantum numbers in the description of a particle state then gives a state in the product space of these vector spaces. A basis for this space is given by the vectors $|\boldsymbol{k}, s\rangle = |\boldsymbol{k}\rangle|s\rangle$. Given $\boldsymbol{k} \in \mathbb{R}^d$ and $s \in \{\uparrow, \downarrow\}$, these basis vectors transform according to the product representation of the representations on each vector space, given by

$$g : |\boldsymbol{k}', s'\rangle = \sum_{\boldsymbol{k}s} |\boldsymbol{k}, s\rangle D^{(\boldsymbol{k}\times s)}_{\boldsymbol{k}s;\,\boldsymbol{k}'s'} = \sum_{\boldsymbol{k}s} |\boldsymbol{k}, s\rangle D_{g\;ss'} \delta_{\boldsymbol{k}, g:\boldsymbol{k}'}, \tag{4.24}$$

under a group element g. Here $g\boldsymbol{k}'$ means application of g to the vector $\boldsymbol{k}'$ through the standard representation of g in $\mathbb{R}^d$. $D_{g\;ss'}$ is a representation on the spin-up spin-down vector space given by the matrix

$$D_{g\;ss'} = \sigma^0_{ss'} \cos(\phi/2) - i\hat{\boldsymbol{u}} \cdot \boldsymbol{\sigma}_{ss'} \sin(\phi/2), \tag{4.25}$$

where $\hat{\boldsymbol{u}}$ is the rotation axis unity vector, while ϕ is the angle that defines the proper rotation associated with g. $\boldsymbol{\sigma}$ is the vector notation for the 3 Pauli matrices and $\sigma^0_{ss'} = \delta_{ss'}$.

In the BCS Hilbert space which we discussed in more detail in Section 4.3, the basis vectors are outer products of the momentum spin basis vectors with opposite momentum: $\{|\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle\}$. The product representation on this vector space then transforms the basis vectors according to

$$g : |\boldsymbol{k}', s_1'\rangle|-\boldsymbol{k}', s_2'\rangle = \sum_{\boldsymbol{k}\,s_1 s_2} |\boldsymbol{k}, s_1\rangle|-\boldsymbol{k}, s_2\rangle D^{(D\times D)}_{\boldsymbol{k}\,s_1 s_2;\,\boldsymbol{k}'\,s_1' s_2'}(g), \tag{4.26}$$

where

$$D^{(D\times D)}_{\boldsymbol{k}\,s_1 s_2;\,\boldsymbol{k}'\,s_1' s_2'}(g) = D_{g\,s_1 s_1'} \delta_{\boldsymbol{k}, g:\boldsymbol{k}'} D_{g\,s_2 s_2'} \delta_{-\boldsymbol{k}, g:-\boldsymbol{k}'}. \tag{4.27}$$

Since group representations on $\boldsymbol{k}$ is a linear transformation then $\delta_{-\boldsymbol{k}, g:-\boldsymbol{k}'} = \delta_{\boldsymbol{k}, g:\boldsymbol{k}'}$, such that the last Kronecker delta function becomes superfluous.

4.4.3 Representation on ψ-$\boldsymbol{d}$ functions

The coefficients of the basis expansion of a vector in the BCS Hilbert space were given in the conventional ψ-$\boldsymbol{d}$ notation in Eq. (4.10). Taking

the active view of group transformations we can say that the expansion coefficients of arbitrary states in the BCS Hilbert space $\mathcal{B}$ transforms like the v_i in Eq. (4.22) but where now the representation matrix D is given by the matrix $D^{(D\times D)}_{\boldsymbol{k}\, s_1 s_2;\, \boldsymbol{k}'\, s_1' s_2'}$ above in Eq. (4.27). Written out then, the coefficients transform according to

$$g : v_{s_1 s_2}(\boldsymbol{k}) = \sum_{\boldsymbol{k}'\, s_1' s_2'} D^{(D\times D)}_{\boldsymbol{k}\, s_1 s_2;\, \boldsymbol{k}'\, s_1' s_2'} v_{s_1' s_2'}(\boldsymbol{k}'). \tag{4.28}$$

Let now $|v\rangle$ be a state that is even in space, meaning that its expansion only consists of coefficients $\psi(\boldsymbol{k})$ in the ψ-$\boldsymbol{d}$ notation. Then we see from Eq. (4.11) that $\psi(\boldsymbol{k})$ can be written $\psi(\boldsymbol{k}) = v_{\uparrow\downarrow}(\boldsymbol{k})$. The transformation properties of $\psi(\boldsymbol{k})$ are thus given by

$$\begin{aligned} g : \psi(\boldsymbol{k}) = g : v_{\uparrow\downarrow}(\boldsymbol{k}) &= \sum_{\boldsymbol{k}'\, s_1' s_2'} D^{(D\times D)}_{\boldsymbol{k}\uparrow\downarrow;\, \boldsymbol{k}'\, s_1' s_2'} \psi(\boldsymbol{k}')(i\sigma^y)_{s_1' s_2'} \\ &= \psi(g^{-1} : \boldsymbol{k})(i\sigma^y)_{\uparrow\downarrow} = \psi(g^{-1} : \boldsymbol{k}). \end{aligned} \tag{4.29}$$

In this calculation we inserted the expression of $D^{(D\times D)}_{\boldsymbol{k}\, s_1 s_2;\, \boldsymbol{k}'\, s_1' s_2'}$ in Eq. (4.27) and used the equation $D_g i\sigma^y D_g^{\mathsf{T}} = i\sigma^y$, where D_g are the spin representation matrices given in Eq. (4.25).

To find the transformation properties of $\boldsymbol{d}_{\boldsymbol{k}}$ the principle is the same as above for $\psi(\boldsymbol{k})$ but the calculations become more involved. We assume that the spin-momentum basis expansion of state $|v\rangle$ consists of only odd coefficients so that $v_{s_1 s_2}(\boldsymbol{k}) = \boldsymbol{d}_{\boldsymbol{k}} \cdot (\boldsymbol{\sigma} i\sigma^y)_{s_1 s_2}$. Inserting this into the active transformation of the coefficients of $|V\rangle$ in Eq. (4.28) and also inserting the expression for $D^{(D\times D)}_{\boldsymbol{k}\, s_1 s_2;\, \boldsymbol{k}'\, s_1' s_2'}$ as we did before yields

$$\begin{aligned} g : v_{s_1 s_2}(\boldsymbol{k}) &= \sum_{\boldsymbol{k}'\, s_1' s_2'} \delta_{\boldsymbol{k}, g:\boldsymbol{k}'} D_{g\, s_1 s_1'} D_{g\, s_2 s_2} \boldsymbol{d}_{\boldsymbol{k}'} \cdot (\boldsymbol{\sigma} i\sigma^y)_{s_1' s_2'} \\ &= \sum_s \left(D_g \boldsymbol{\sigma} \sigma^y D_g^{\mathsf{T}} \sigma^y \right)_{s_1 s} \cdot \boldsymbol{d}_{g^{-1}:\boldsymbol{k}}\, i\sigma^y_{s s_2}. \end{aligned} \tag{4.30}$$

Since a group transformation (aka. the linear transformation given by a group representation) can not bring a state that was even to be odd or vice versa, then the resulting state given by the transformed coefficients $g : v_{s_1 s_2}(\boldsymbol{k})$ has to remain odd, and thus they can be expanded in terms of a new $\boldsymbol{d}'$ such that

$$g : v_{s_1 s_2}(\boldsymbol{k}) = \sum_s \boldsymbol{d}'_{g^{-1}:\boldsymbol{k}} \cdot \boldsymbol{\sigma}_{s_1 s} i\sigma^y_{s s_2}. \tag{4.31}$$

Having expanded both sides of the transformed coefficients with a common factor $i\sigma^y$ to the right we can equate the remaining 2×2 spin matrices which gives an expression for $\boldsymbol{d}'_{g^{-1}:\boldsymbol{k}} \cdot \boldsymbol{\sigma}$ by comparing Eq. (4.30) and Eq. (4.31). Furthermore, using the anti-commutation property $\{\sigma^i, \sigma^j\} = 2\delta_{ij}\sigma^0$ of Pauli matrices we find that

$$d'_{\boldsymbol{k},i} = \frac{1}{4} \operatorname{Tr}\left(\{\sigma^i, \boldsymbol{d}'_{\boldsymbol{k}} \cdot \boldsymbol{\sigma}\}\right). \tag{4.32}$$

Inserting the expression for $\boldsymbol{d}'_{g^{-1}:\boldsymbol{k}} \cdot \boldsymbol{\sigma}$ and the full expression of the $SU(2)$ spin-representation matrices D_g, which is found in Eq. (4.25), yields after some algebra

$$\begin{aligned} d'_{g^{-1}:\boldsymbol{k},i} &= \frac{1}{4} \operatorname{Tr}\left(\{\sigma^i, \boldsymbol{d}_{g^{-1}:\boldsymbol{k}} \cdot D_g \boldsymbol{\sigma} \sigma^y D_v^{\mathsf{T}} \sigma^y\}\right) \\ &= R_{ij}(\hat{\boldsymbol{u}}, \phi) d_{g^{-1}:\boldsymbol{k},j}, \end{aligned} \tag{4.33}$$

where we have defined the matrix

$$\begin{aligned} R_{ij}(\hat{\boldsymbol{u}}, \phi) &= \delta_{ij} \cos\phi + \hat{u}_i \hat{u}_j (1 - \cos\phi) - \epsilon_{ijk} \hat{u}_k \sin\phi \\ &= \begin{pmatrix} \cos\phi + \hat{u}_x^2(1-\cos\phi) & \hat{u}_x\hat{u}_y(1-\cos\phi) - \hat{u}_z \sin\phi & \hat{u}_x\hat{u}_z(1-\cos\phi) + \hat{u}_y \sin\phi \\ \hat{u}_y\hat{u}_x(1-\cos\phi) + \hat{u}_z \sin\phi & \cos\phi + \hat{u}_y^2(1-\cos\phi) & \hat{u}_y\hat{u}_z(1-\cos\phi) - \hat{u}_x \sin\phi \\ \hat{u}_z\hat{u}_x(1-\cos\phi) - \hat{u}_y \sin\phi & \hat{u}_z\hat{u}_y(1-\cos\phi) + \hat{u}_x \sin\phi & \cos\phi + \hat{u}_z^2(1-\cos\phi) \end{pmatrix}. \end{aligned} \tag{4.34}$$

This matrix is in fact the rotation matrix of a vector in $\mathbb{R}^3$ by an angle ϕ about a unit vector $\hat{\boldsymbol{u}}$. Since the coefficients of an odd state $|v\rangle$ are fully determined by the vector $\boldsymbol{d}$, their transformation can be regarded just as a tranformation of $\boldsymbol{d}$ itself which thus takes the form

$$g : d_{\boldsymbol{k},i} = R_{ij}(\hat{\boldsymbol{u}}, \phi) d_{g^{-1}:\boldsymbol{k},j}, \tag{4.35}$$

where $\hat{\boldsymbol{u}}$ and ϕ gives the unit vector and angle respectively, of the proper rotation associated with g. The conclusion is thus that $\boldsymbol{d}$ transforms as a vector by the proper rotation associated with g.

4.4.4 Projection Operators

Let us assume that we are in a vector space V that can be divided into possibly several different irreducible representations $D^{(\alpha)}$ of some symmetry group G. Further, let the basis vectors of these irreducible

representations be denoted by $\boldsymbol{b}_m^{(\alpha)}$ where m thus counts the number of basis vectors in each irrep. Then an arbitrary vector $\boldsymbol{f} \in V$ can be written in terms of these basis vectors as

$$\boldsymbol{f} = \sum_\alpha \sum_m c_m^{(\alpha)} \boldsymbol{b}_m^{(\alpha)}. \tag{4.36}$$

A projection operator can be used to extract any combination of constant $c_m^{(\alpha)}$ multiplied by a basis vector $\boldsymbol{b}_n^{(\alpha)}$, where m and n can in general be different. Denoting the projection operator that picks out the mth constant multiplied by the lth basis vector in the irrep. β of the expansion of $\boldsymbol{f}$: $P_{lm}^{(\beta)}$, then

$$P_{lm}^{(\beta)} \boldsymbol{f} = c_m^{(\beta)} \boldsymbol{b}_l^{(\beta)}. \tag{4.37}$$

This is extremely useful in finding a bases for the irreducible representations. To acheive this, the projection operator is defined as

$$P_{l,m}^{(\beta)} = \frac{d_\beta}{|G|} \sum_{g \in G} D_{lm}^{(\beta)}(g)^* g :, \tag{4.38}$$

where d_β is the dimension of irrep. β, $D_{lm}^{(\beta)}(g)$ is the lm element of the matrix representation of the group element g and finally we have used the notation $g :$ to denote application on vectors by the relevant representation. An example is the application of g to the basis vectors $\boldsymbol{b}_m^{(\alpha)}$. Since the relevant representation of g in this case is the irreducible representation for which $\boldsymbol{b}_m^{(\alpha)}$ is a basis vector, the application becomes

$$g : b_m^{(\alpha)} = \sum_n b_n^{(\alpha)} D_{nm}^{(\alpha)}(g). \tag{4.39}$$

Usually, the full generality of the projection operators $P_{l,m}^{(\beta)}$ isn't needed and it suffices to consider the diagonal projection operators $P_{l,l}^{(\beta)} \equiv P_l^{(\beta)}$ or indeed their sum, in which case the resulting operator can be written only in terms of the irrep. characters $\chi^{(\alpha)}(g)$ since

$$P^{(\beta)} \equiv \sum_l P_l^{(\beta)} = \frac{d_\beta}{|G|} \sum_{g \in G} \sum_l D_{ll}^{(\beta)}(g)^* g := \frac{d_\beta}{|G|} \sum_{g \in G} \chi^{(\beta)}(g)^* g :. \tag{4.40}$$

4.5 Fermionic symmetry transformations

4.6 Time-reversal symmetry

za

4.7 Symmetries of the Square Lattice

The symmetry group of the square lattice is denoted C_{4v} in the Schönflies notation. It contains 8 elements in total:

e: The identity element (do nothing),

C_4: Rotation by 90° in the positive direction (ccw),

C_4^{-1}: Rotation by 90° in the negative direction (cw),

C_4^2: Rotation by 180°,

σ_x: Mirror about the zy-plane,

σ_y: Mirror about the zx-plane,

σ_{d_1}: Mirror about the downwards diagonal plane,[1]

σ_{d_2}: Mirror about the upwards diagonal plane.

This results in the group multiplication table in Table 4.1. We can check that this is correct by performing the group transformations in the top row followed by the one in the left column and seeing that this results in the group transformations where these two intersect. As an example consider the vector $(x, y)^\mathsf{T}$. Transforming this by the 90° counter clockwise rotation C_4 we get $(-y, x)^\mathsf{T}$. Then mirroring this result about the yz-plane yields $(y, x)^\mathsf{T}$. We now realize that this is the same as mirroring the original vector about the axis $y = x$, hence $\sigma_x C_4 = \sigma_{d_2}$ as the multiplication table says.

1. We are assuming the western bias of left-to-right movement here. More precisely it is the plane containing the z axis and the axis $y = -x$

Table 4.1: Group multiplication table of the group C_{4v}.

	e	C_4	C_4^2	C_4^{-1}	σ_x	σ_y	σ_{d_1}	σ_{d_2}
e	e	C_4	C_4^2	C_4^{-1}	σ_x	σ_y	σ_{d_1}	σ_{d_2}
C_4	C_4	C_4^2	C_4^{-1}	e	σ_{d_1}	σ_{d_2}	σ_y	σ_x
C_4^2	C_4^2	C_4^{-1}	e	C_4	σ_y	σ_x	σ_{d_2}	σ_{d_1}
C_4^{-1}	C_4^{-1}	e	C_4	C_4^2	σ_{d_2}	σ_{d_1}	σ_x	σ_y
σ_x	σ_x	σ_{d_2}	σ_y	σ_{d_1}	e	C_4^2	C_4^{-1}	C_4
σ_y	σ_y	σ_{d_1}	σ_x	σ_{d_2}	C_4^2	e	C_4	C_4^{-1}
σ_{d_1}	σ_{d_1}	σ_x	σ_{d_2}	σ_y	C_4	C_4^{-1}	e	C_4^2
σ_{d_2}	σ_{d_2}	σ_y	σ_{d_1}	σ_x	C_4^{-1}	C_4	C_4^2	e

4.7.1 Conjugation classes

The conjugation classes of a group is the sets of group elements that is conjugate to each other, meaning that there exists a group element g such that $gAg^{-1} = B$ between conjugate elements A and B. Since conjugation is an equivalence relation it subdivides the group elements exactly into conjugation classes. The conjugation classes of the group C_{4v} are $e = \{e\}$, $2C_4 = \{C_4, C_4^{-1}\}$, $C_4^2 = \{C_4^2\}$, $2\sigma_d = \{\sigma_{d_1}, \sigma_{d_2}\}$ and $2\sigma_v = \{\sigma_x, \sigma_y\}$. The character $\chi^\Gamma(g)$ of a representation Γ is the trace of the representation matrix $D^\Gamma(g)$ of a certain group element g. Since the trace is cyclic, then for conjugate group elements A and B

$$\begin{aligned}\chi^\Gamma(B) &= \text{Tr}\left(D^\Gamma(g)D^\Gamma(A)D^\Gamma(g^{-1})\right)\\ &= \text{Tr}\left(D^\Gamma(g^{-1})D^\Gamma(g)D^\Gamma(A)\right)\\ &= \text{Tr}\left(D^\Gamma(g^{-1}g)D^\Gamma(A)\right) = \chi^\Gamma(A).\end{aligned} \tag{4.41}$$

This means that the representation of all group elements in a certain conjugation class has the same character.

It is useful to list the characters of the different conjugation classes in a table of the different irreducible representations of a group. This is because the number of conjugation classes of a finite group is the same as the number of irreducible representations of that group. This table is known as the character table of the group. The character table of the group C_{4v} is shown in Table 4.2. This table can be derived with-

Table 4.2: Character table of the group C_{4v}.

C_{4v}	e	$2C_4$	C_4^2	$2\sigma_v$	$2\sigma_d$
Γ_1	1	1	1	1	1
Γ_2	1	1	1	-1	-1
Γ_3	1	-1	1	1	-1
Γ_4	1	-1	1	-1	1
Γ_5	2	0	-2	0	0

out knowing the details of the irreducible representations but in stead using character relations from basic group theory.[2]

4.7.2 Irreducible representations

The dimensionality of the irrep. can be found in Table 4.2 by looking up the first column, i.e. the column giving the character of the conjugation class $\{e\}$. Since the group element e maps to the identity transformation in all representations then its trace gives the dimension of the representation. From the table we see that all the irreducible representations are 1-dimensional except for Γ_5 which is 2 dimensional. All the 1-dimensional irreducible representation matrices are then completely determined by the character table since they are just given by the characters themselves, i.e. $D^{\Gamma_2}(\sigma_x) = -1$.

To find a 2-dimensional representation of C_{4v} we can imagine how a normal 2D vector $(x, y)^\mathsf{T} \in \mathbb{R}^2$ behaves under its transformations. We take again the example of a counter clockwise rotation by 90° which tranforms a vector

$$C_4 : \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{4.42}$$

Obviously, the matrix

$$D^{(\Gamma_5)}(C_4) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{4.43}$$

is the representation matrix of a two-dimensional representation of C_4. Continuing in this way for all the group tranformations yields the ma-

2. Many of these relations are derived from the great orthogonality theorem which can be found e.g. in [3].

trices

$$D^{(\Gamma_5)}(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad D^{(\Gamma_5)}(C_4) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{4.44a}$$

$$D^{(\Gamma_5)}(C_4^{-1}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad D^{(\Gamma_5)}(C_4^2) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.44b}$$

$$D^{(\Gamma_5)}(\sigma_x) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad D^{(\Gamma_5)}(\sigma_y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.44c}$$

$$D^{(\Gamma_5)}(\sigma_{d_1}) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad D^{(\Gamma_5)}(\sigma_{d_2}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{4.44d}$$

Taking the trace of these matrices, we see that this representations characters are the same as the ones for the irrep. Γ_5 in the character table (Table 4.2). This implies that $\sum_g |\chi^{\Gamma_5 u}(g)|^2 = |C_{4_v}|$ which implies in turn that the representation given by the matrices in Eq. (4.44) is irreducible. Thus this is indeed the Γ_5 irreducible representation as advertised, and we have completed the description of the representation matrices of all the irreducible representations of C_{4v}.

4.7.3 Representation on odd BCS functions d_k

When discussing the representation of general group elements on states in the BCS Hilbert space in Section 4.4.3 we learned from Eq. (4.35) that the coefficient $\boldsymbol{d_k} \in \mathbb{R}^3$ of odd states transforms by a proper rotation $R(\hat{\boldsymbol{u}}, \phi)$. A proper rotation of a group element g is the rotation obtained when writing g as a rotation followed by an inversion P or identity. We thus obtain the representation matrices of the 3D-representation of group elements $g \in C_{4v}$ directly from the 3D rotation matrix $R(\hat{\boldsymbol{u}}, \phi)$ by envisioning the combination of a rotation and $C_4^2 = P$ or e that leads to g. Because of this, we denote the representation matrices of this representation $R(g)$. As an example $\sigma_x = C_4^2 R(\hat{\boldsymbol{x}}, \pi)$, such that the representation $R(\sigma_x) = R(\hat{\boldsymbol{x}}, \pi)$. Written out in its full matrix form this yields the representation matri-

ces

$$R(C_4) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R(C_4^{-1}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.45a}$$

$$R(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R(C_4^2) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.45b}$$

$$R(\sigma_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad R(\sigma_y) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{4.45c}$$

$$R(\sigma_{d_1}) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad R(\sigma_{d_2}) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{4.45d}$$

4.8 Square Lattice Harmonics

Since $\psi(\boldsymbol{k})$ and $\boldsymbol{d}_{\boldsymbol{k}}$ are invariant with respect to translation by any reciprocal lattice vector $\boldsymbol{Q}$: $\psi(\boldsymbol{k}+\boldsymbol{Q}) = \psi(\boldsymbol{k})$, they can be expanded in a discrete Fourier transform over the real lattice, such that

$$\psi(\boldsymbol{k}) = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}} \psi_{\boldsymbol{R}} \cos \boldsymbol{R} \cdot \boldsymbol{k}, \tag{4.46}$$

and

$$\boldsymbol{d}_{\boldsymbol{k}} = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}} \boldsymbol{d}_{\boldsymbol{R}} \sin \boldsymbol{R} \cdot \boldsymbol{k}, \tag{4.47}$$

where the exponential of the Fourier transform has been reduced to trigonometric functions by the parity of the functions.

We are interested in the basis vectors $|\Gamma, q, m\rangle$ of the representations of the symmetry group C_{4v} of the $2D$ square lattice. In this ket notation, Γ gives the irrep., m enumerates the basis vectors in case the irrep. is multi-dimensional, while q gives the version of the irrep. in case the space of possible $|v\rangle$ permits multiple versions of the same irrep. In the active view of group transformations, the question of finding the basis vectors translates to finding the basis functions of the functions $\psi(\boldsymbol{k})$ and $\boldsymbol{d}_{\boldsymbol{k}}$ for even and odd bases respectively. In Section 4.4.3 we saw how these functions transformed under group transformations.

In Section 4.4.4 we saw how the projection operators could be used to extract individual basis vectors. We will now use these operators on the Fourier expansions of $\psi(\boldsymbol{k})$ and $\boldsymbol{d_k}$ to extract possible basis functions given the symmetry group of the square lattice.

4.8.1 Even basis functions

We remember from Eq. (4.29) that the function $\psi(\boldsymbol{k})$ transforms as $g : \psi(\boldsymbol{k}) = \psi(g^{-1} : \boldsymbol{k})$ under a group transformation g. Operating on the Fourier expansion of $\psi(\boldsymbol{k})$ in Eq. (4.46) with the projection operator defined in Eq. (4.38) by an arbitrary irrep. Γ yields

$$\begin{aligned} P^{(\Gamma)}_{l,l}\psi(\boldsymbol{k}) = \frac{d_\Gamma}{8\sqrt{N}}\sum_{\boldsymbol{R}}\psi_{\boldsymbol{R}}\Big[&(D^{(\Gamma)}_{ll}(e) + D^{(\Gamma)}_{ll}(C_4^2))\cos(\boldsymbol{R}\cdot\boldsymbol{k}) \\ &+(D^{(\Gamma)}_{ll}(C_4) + D^{(\Gamma)}_{ll}(C_4^{-1}))\cos(R_xk_y - R_yk_x) \\ &+(D^{(\Gamma)}_{ll}(\sigma_x) + D^{(\Gamma)}_{ll}(\sigma_y))\cos(R_xk_x - R_yk_y) \\ &+(D^{(\Gamma)}_{ll}(\sigma_{d_1}) + D^{(\Gamma)}_{ll}(\sigma_{d_2}))\cos(R_xk_y + R_yk_x). \end{aligned} \tag{4.48}$$

Since $\boldsymbol{k} \in \mathbb{R}^2$ we have in this calculation used the natural $2D$ representation given by the matrices in Eq. (4.44) of the group elements in C_{4v} to calculate the expressions $g^{-1} : \boldsymbol{k}$ inside the cosine functions. Inserting the matrix element of the different irreducible representations of C_{4v} which we discussed in Section 4.7 we get the projected functions

$$\begin{aligned} P^{(\Gamma_1)}_{1,1}\psi(\boldsymbol{k}) &\propto \sum_{\boldsymbol{R}}\psi_{\boldsymbol{R}}[\cos R_xk_x\cos R_yk_y + \cos R_xk_y\cos R_yk_x], \\ P^{(\Gamma_2)}_{1,1}\psi(\boldsymbol{k}) &\propto \sum_{\boldsymbol{R}}\psi_{\boldsymbol{R}}[\sin R_xk_y\sin R_yk_x - \sin R_xk_x\sin R_yk_y], \\ P^{(\Gamma_3)}_{1,1}\psi(\boldsymbol{k}) &\propto \sum_{\boldsymbol{R}}\psi_{\boldsymbol{R}}[\cos R_xk_x\cos R_yk_y - \cos R_xk_y\cos R_yk_x], \\ P^{(\Gamma_4)}_{1,1}\psi(\boldsymbol{k}) &\propto \sum_{\boldsymbol{R}}\psi_{\boldsymbol{R}}[\sin R_xk_x\sin R_yk_y + \sin R_xk_y\sin R_yk_x], \\ P^{(\Gamma_5)}_{1,1}\psi(\boldsymbol{k}) &\propto 0, \\ P^{(\Gamma_5)}_{2,2}\psi(\boldsymbol{k}) &\propto 0. \end{aligned} \tag{4.49}$$

Since the projection operators $P^{(\Gamma)}_{ll}$ projects onto the subspace of vectors belonging to the irreducible representation Γ, if we let the different

basis functions of the irreps. Γ be denoted $\psi^{(\Gamma),q,m}(\boldsymbol{k})$, then we expect the projection to produce the result

$$P^{(\Gamma)}_{l,l}\psi(\boldsymbol{k}) = \sum_q c_{q,l}\psi^{(\Gamma),q,l}(\boldsymbol{k}). \tag{4.50}$$

Here q again enumerates the version of the basis of Γ possible in the space of different $\psi(\boldsymbol{k})$, and $c_{q,l}$ are the coefficients of $\psi(\boldsymbol{k})$ in the basis of the irrep. basis functions. Comparing Eqs. (4.50) and (4.49), we see that different sets of basis vectors can be obtained for the irreducible representations by including different order terms in the $\boldsymbol{R}$-sum, i.e. different lattice neighbour sites. It is also worth noting that the irreducible representation Γ_5 does not exist in the space of possible $\psi(\boldsymbol{k})$, since it is an odd representation.

Including on-site, nearest neighbor and next-nearest neighbor sites in the $\boldsymbol{R}$ sum of the projected arbitrary function $\psi(\boldsymbol{k})$ on the Γ_1 subspace in Eq. (4.50) we see that any such function can be constructed from the three basis functions

$$\psi^{(\Gamma_1),1}(\boldsymbol{k}) = \frac{1}{2\pi}, \tag{4.51a}$$

$$\psi^{(\Gamma_1),2}(\boldsymbol{k}) = \frac{1}{2\pi}(\cos k_x + \cos k_y), \tag{4.51b}$$

$$\psi^{(\Gamma_1),3}(\boldsymbol{k}) = \frac{1}{\pi}\cos k_x \cos k_y. \tag{4.51c}$$

Each of these functions give a complete basis function set of the Γ_1 irreducible representation which can be checked by calculating how they transform under group elements $g \in C_{4v}$. In this case, since this is the trivial Γ_1 representation, the functions are symmetric under all group elements g which produces the character 1 for all conjugation classes (compare with first row in Table 4.2).

These basis functions are automatically mutually orthogonal since they belong to different irreducible representation version subspaces and their normalization coefficients have been chosen such that they are normal on the 1st Brillouin zone, i.e.

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathrm{d}k_x \mathrm{d}k_y \; \psi^{(\Gamma_1),q}(\boldsymbol{k})^* \psi^{(\Gamma_1),q'}(\boldsymbol{k}) = \delta_{qq'}. \tag{4.52}$$

Inserting up to next-nearest neighbor sites for $\boldsymbol{R}$ in the projected functions under the remaining irreducible representations we find the

orthonormal basis functions

$$\psi^{(\Gamma_3)}(\boldsymbol{k}) = \frac{1}{2\pi}(\cos k_x - \cos k_y), \tag{4.53a}$$

$$\psi^{(\Gamma_4)}(\boldsymbol{k}) = \frac{1}{\pi}\sin k_x \sin k_y, \tag{4.53b}$$

of the representations Γ_3 and Γ_4 respectively. The Γ_3 basis is found by expansion of $\boldsymbol{R}$ to nearest neighbor, while the one for Γ_4 is found at next-nearest neighbor. To get a basis vector for the representation Γ_2 we would need to expand beyond the next-nearest neighbor site.

These basis functions are also known as square lattice harmonics. The set of square lattice harmonic functions includes the set of functions found when expanding $\boldsymbol{R}$ to arbitrary sites. As we have seen, they can be grouped and found through consideration of the irreducible representations of the symmetry group of the square lattice. We have so far only considered even-in-$\boldsymbol{k}$ basis functions. In the next section we will complete our discussion of square lattice harmonics with the inclusion of odd functions.

4.8.2 Odd basis functions

Any state made out of odd basis functions is fully determined by the coefficients $\boldsymbol{d}_{\boldsymbol{k}}$. These coefficients transform as $g : \boldsymbol{d}_{\boldsymbol{k}} = R(g)\boldsymbol{d}_{g^{-1}:\boldsymbol{k}}$ under group elements $g \in C_{4v}$, see Eq. (4.35), where $R(g)$ are the representation matrices in Eq. (4.45). Acting on the Fourier expansion of the arbitrary function $\boldsymbol{d}_{\boldsymbol{k}}$ in Eq. (4.47) with the projection operators Eq. (4.38) down on the subspace of the irreducible representation Γ_5 of the symmetry group C_{4v} yields the results

$$P_{1,1}^{(\Gamma_5)}\boldsymbol{d}_{\boldsymbol{k}} = \frac{\hat{\boldsymbol{z}}}{\sqrt{N}}\sum_{\boldsymbol{R}} d_{\boldsymbol{R},z}\cos R_x k_x \sin R_y k_y, \tag{4.54a}$$

$$P_{2,2}^{(\Gamma_5)}\boldsymbol{d}_{\boldsymbol{k}} = \frac{\hat{\boldsymbol{z}}}{\sqrt{N}}\sum_{\boldsymbol{R}} d_{\boldsymbol{R},z}\sin R_x k_y \cos R_y k_y, \tag{4.54b}$$

for the two basis functions of Γ_5. As in the spin-singlet case we get different versions of the Γ_5 basis vectors depending on the order of our expansion in $\boldsymbol{R}$. Expanding to nearest neighbor sites and normalizing

such that the states are orthonormal produces the basis functions

$$\boldsymbol{d}_{\boldsymbol{k}}^{(\Gamma_5),1} = -\frac{\hat{\boldsymbol{z}}}{2\pi}\sin k_y, \tag{4.55a}$$

$$\boldsymbol{d}_{\boldsymbol{k}}^{(\Gamma_5),2} = \frac{\hat{\boldsymbol{z}}}{2\pi}\sin k_x, \tag{4.55b}$$

of the two-dimensional irreducible representation Γ_5.

4.9 Decomposition of the Potential

Let at first $\hat{V}$ be a general two-body operator that acts on an N-particle state which is a vector in $\mathcal{H}_N = \otimes_{i=1}^{N}\mathcal{H}$. The single particle Hilbert space $\mathcal{H}$ in question is quantified by momentum and spin such that $\mathcal{H} = \text{span}\{|\boldsymbol{k}, s\rangle\}$. Denoting specific combinations of $\boldsymbol{k}$ and s as α as a shorthand for the moment, then $\hat{V}$ acts on basis vectors in $\mathcal{H}_N$ as

$$\hat{V}|\alpha_1\rangle \dots |\alpha_N\rangle = \sum_{1\leq i<j\leq N} \hat{V}_{ij}|\alpha_1\rangle \dots |\alpha_N\rangle, \tag{4.56}$$

by definition of being a two-body operator. Here $\hat{V}_{ij}$ is an operator that only acts on the ith and jth ket. Even though $\hat{V}$ acts on $\mathcal{H}_N$, because of how it can be written in terms of $\hat{V}_{ij}$ and this only acts on two states at a time, it follows that $\hat{V}$ is completely determined by its action on the reduced Hilbert space $\mathcal{H}_2$. This implies that $\hat{V}$ is fully described by its matrix elements

$$\langle\alpha|\langle\alpha'|\hat{V}|\beta\rangle|\beta'\rangle. \tag{4.57}$$

Inserting back the $|\boldsymbol{k}, s\rangle$ notation, these matrix elements are referred to as

$$V_{\boldsymbol{k}_1\boldsymbol{k}_2\boldsymbol{k}_3\boldsymbol{k}_4;\, s_1s_2s_3s_4} = \langle\boldsymbol{k}_1 s_1|\langle\boldsymbol{k}_2 s_2|\hat{V}|\boldsymbol{k}_4 s_4\rangle|\boldsymbol{k}_3 s_3\rangle. \tag{4.58}$$

When $\hat{V}$ is a BCS operator acting on the BCS Hilbert space described in the previous section, these matrix elements are denoted

$$V_{\boldsymbol{k}\boldsymbol{k}';\, s_1s_2s_3s_4} = \langle\boldsymbol{k}s_1|\langle-\boldsymbol{k}s_2|\hat{V}|\boldsymbol{k}'s_4\rangle|-\boldsymbol{k}'s_3\rangle. \tag{4.59}$$

Since $\hat{V}$ is Hermitian, it must be diagonalizable in a basis of eigenfunctions. Barring accidental degeneracy, a basis for a d-degenerate

eigenvalue is also a basis for an irreducible representation of the symmetry group G of the Hamiltonian. In the case of accidental degeneracy then this d-dimensional vector space consists of several non-intersecting subspaces where each subspace is a basis for a (possibly different) irrep. Note that this does not mean that (barring accidental degeneracy) there exists one seperate eigenvalue for each irrep. of G since there might be several different eigenvalues with different eigenspace bases but where all of them are bases for the same irrep. Regardless of these details, the connection between irreducible representations and the eigenvalues of $\hat{V}$ is a great help in finding the bases for which it is diagonal.

We let the basis for a d_Γ-dimensional irrep. Γ be denoted $\{|\Gamma, q_\Gamma, m\rangle\}_{m=1}^{d_\Gamma}$, where $\hat{V}$ has an eigenvalue V_{Γ,q_Γ} for the vectors in this basis and q_Γ is an index enumerating the different versions of bases of Γ that $\hat{V}$ might have in its set of eigenspace bases. Since $\hat{V}$ then is diagonal in this set of bases then

$$\hat{V} = \sum_{\Gamma q_\Gamma} V_{\Gamma,q_\Gamma} \sum_{m=1}^{d_\Gamma} |\Gamma, q_\Gamma, m\rangle\langle\Gamma, q_\Gamma, m|. \tag{4.60}$$

Because of the potential for accidental degeneracy we can not guarantee that $V_{\Gamma,q_\Gamma} \neq V_{\Gamma',q_{\Gamma'}}$ for different Γ and Γ'. Inserting this expression for $\hat{V}$ into the matrix elements in Eq. (4.59) lets us write them in terms of irreducible representation basis vectors in the momentum spin function representation:

$$V_{\boldsymbol{k}\boldsymbol{k}';s_1s_2s_3s_4} = \sum_{\Gamma} V_{\Gamma,q_\Gamma} \sum_{m=1}^{d_\Gamma} \Psi_{s_1s_2}^{\Gamma,q_\Gamma}(\boldsymbol{k})\Psi_{s_3s_4}^{\Gamma,q_\Gamma}(-\boldsymbol{k}')^\dagger, \tag{4.61}$$

where

$$\Psi_{s_1s_2}^{\Gamma,q_\Gamma}(\boldsymbol{k}) = \langle \boldsymbol{k}, s_1|\langle -\boldsymbol{k}, s_2 \mid \Gamma, q_\Gamma, m\rangle. \tag{4.62}$$

We can separate the set of different irreducible representation bases into bases that have vectors that transform either symmetrically or anti-symmetrically with respect to the group element of space inversion P. We call the representations of such bases even or odd representations. Even representations are those that map P to the identity operator $\mathbb{1}$ and as a consequence have $\Psi_{s_1s_2}^{\Gamma,q_\Gamma,m}(-\boldsymbol{k}) = \Psi_{s_1s_2}^{\Gamma,q_\Gamma,m}(\boldsymbol{k})$. Writing the spin-indices in these functions in terms of Pauli matrices

by using the expansion in Eq. (4.9) and using the fermionic symmetry then even representations a have

$$\Psi^{a,q_a,m}_{s_1 s_2}(\boldsymbol{k}) = \psi^{a,q_a,m}_{\boldsymbol{k}} i\sigma^y_{s_1 s_2}. \tag{4.63}$$

Odd representations b map P to the inversion operator I such that $\Psi^{b,q_b,m}_{s_1 s_2}(-\boldsymbol{k}) = -\Psi^{b,q_b,m}_{s_1 s_2}(\boldsymbol{k})$. Expanding in Pauli matrices then yields

$$\Psi^{b,q_b,m}_{s_1 s_2}(\boldsymbol{k}) = \boldsymbol{d}^{b,q_b,m}_{\boldsymbol{k}} \cdot (\boldsymbol{\sigma} i\sigma^y)_{s_1 s_2}. \tag{4.64}$$

Separating the sum over irreducible representations Γ into sums over even (a) and odd (b) representations in the potential operator matrix elements in Eq. (4.61), we arrive at the fully expanded expression

$$\begin{aligned} V_{\boldsymbol{k}\boldsymbol{k}';\, s_1 s_2 s_3 s_4} = &\sum_{a q_a} V_{a,q_a} \sum_{m=1}^{d_a} \psi^{a,q_a,m}_{\boldsymbol{k}} i\sigma^y_{s_1 s_2} (\psi^{a,q_a,m}_{-\boldsymbol{k}'} i\sigma^y_{s_3 s_4})^\dagger \\ &+ \sum_{b q_b} V_{b,q_b} \sum_{m=1}^{d_b} (\boldsymbol{d}^{b,q_b,m}_{\boldsymbol{k}} \cdot \boldsymbol{\sigma} i\sigma^y)_{s_1 s_2} \left[(\boldsymbol{d}^{b,q_b,m}_{-\boldsymbol{k}'} \cdot \boldsymbol{\sigma} i\sigma^y)_{s_3 s_4} \right]^\dagger. \end{aligned} \tag{4.65}$$

In this use of the dagger notation, the adjoint acts on both of the matrix indices such that $\boldsymbol{d}^\dagger_{-\boldsymbol{k}} = \boldsymbol{d}^*_{\boldsymbol{k}}$ and $\sigma^\dagger_{s_1 s_2} = \sigma^*_{s_2 s_1}$.

Chapter 5

Lattice Models

When we have a model for the free energy of a statistical mechanical system that is too complicated to calculate analytically one approach is to utilize computers and Monte-Carlo techniques. To make numerical methods more effective one common approach is to discretize continuous models down on a numerical lattice. The lattice can in principle be of any form as long as the continuum limit reproduces the original theory, however in this thesis we will exclusively focus on a square (cubic) numerical lattice due to its simplicity.

In this chapter we will introduce different aspects of discretizing a continuous free-energy model down on a square numerical lattice. If starting with a continuous model with a spatially dependent field $f(\boldsymbol{r})$, then the discretized model will have a corresponding field $f_{\boldsymbol{r}}$ only defined on the numerical lattice point sites at

$$\boldsymbol{r} = \sum_{\mu} r_{\mu}\hat{\mu} = \sum_{\mu} a_{\mu} n_{\mu}\hat{\mu} \tag{5.1}$$

where a_{μ} is the distance between lattice sites, $n_{\mu} \in [0, 1, \dots, N_{\mu} - 1]$ and N_{μ} is the total number of sites in the μ-direction. The length of the numerical lattice in this direction is $L_{\mu} = a_{\mu}N_{\mu}$. The cubic numerical lattice is specified by $a_{\mu} = a \ \forall \mu$ and $\mu \in \{x, y, z\}$. Any integrals $\int \mathrm{d}^3 r \ F[f(\boldsymbol{r})]$ will in such discretization have to be replaced with sums

such that

$$\int \mathrm{d}^3 r \mapsto a^3 \sum_{r_\mu = a}^{Na} . \tag{5.2}$$

If we are interested in bulk properties of the model in the thermodynamic limit, then specifying realistic boundary conditions of the numerical lattice are of less importance. In this case, periodic boundary conditions are from a computational perspective a convenient choice, which are defined by the requirement that $f_{\boldsymbol{r}+L_\mu \hat{\mu}} = f_{\boldsymbol{r}}$ for any direction $\hat{\mu}$.

5.1 Discretizing derivatives

In a model where fields only are defined at discrete points in space, any spatial gradient of the fields must take the form of discrete differences of the field values at these points. In a cubic grid of points with defined field-values such differences can be denoted by the forward difference operator Δ_μ, which acts on a spatially discrete function $f_{\boldsymbol{r}}$ as a forward difference in the direction of $\hat{\mu}$ such that

$$\Delta_\mu f_{\boldsymbol{r}} = f_{\boldsymbol{r}+a\hat{\mu}} - f_{\boldsymbol{r}}, \tag{5.3}$$

where a is the distance between lattice points. In a Euclidean geometry, then the natural discretization of a derivative ∂_μ is $\partial_\mu \mapsto \Delta_\mu / a$, which reproduces the continuum derivative in the limit $a \to 0$ with a fixed grid-size. Using an appropriate set of units, we in most cases can set $a = 1$.

5.1.1 Covariant derivatives

When discretizing continuous gauge theories, some extra care has to be taken when discretizing a covariant derivative. Because of the gauge field, the geometry is no longer naively Euclidean such that we have to rotate a field by a gauge group element to compare the field value at two spatially separate points. Given a $U(1)$ gauge symmetry with gauge field components $A_\mu(\boldsymbol{r})$, the appropriate way of discretizing a covariant derivative is the identification [4]

$$D_\mu f(\boldsymbol{r}) = [\partial_\mu + igA_\mu(\boldsymbol{r})] f(\boldsymbol{r}) \mapsto \frac{1}{a}(f_{\boldsymbol{r}+a\hat{\mu}} U_{\boldsymbol{r},\mu} - f_{\boldsymbol{r}}). \tag{5.4}$$

The value of f at $\boldsymbol{r} + a\hat{\mu}$ is parallel transported back to $\boldsymbol{r}$ by the $U(1)$ group element [5]

$$U_{\boldsymbol{r},\mu} = e^{igA_{\boldsymbol{r},\mu}}, \tag{5.5}$$

where g is the coupeling constant between f and the gauge field A, and

$$A_{\boldsymbol{r},\mu} \equiv \int_{\boldsymbol{r}}^{\boldsymbol{r}+a\hat{\mu}} \mathrm{d}\boldsymbol{r} \cdot \boldsymbol{A}(\boldsymbol{r}), \tag{5.6}$$

is a link-variable, linking $\boldsymbol{r}$ to its nearest neighbors.

In the limit of $a \to 0$ this identification reproduces the covariant derivative.[1] Furthermore it produces terms that transform in an analogous way to the continuum version under gauge transformations such that gauge invariant terms remain invariant after discretization. In the continuous fields, a gauge transformation is defined by

$$\begin{aligned} f(\boldsymbol{r}) &\to f(\boldsymbol{r})e^{i\phi(\boldsymbol{r})}, \\ A_\mu(\boldsymbol{r}) &\to A_\mu(\boldsymbol{r}) - \frac{1}{g}\partial_\mu\phi(\boldsymbol{r}). \end{aligned} \tag{5.7}$$

Then the covariant derivative transforms as $D_\mu f(\boldsymbol{r}) \to e^{i\phi(\boldsymbol{r})} D_\mu f(\boldsymbol{r})$ such that terms such as $|D_\mu f(\boldsymbol{r})|^2$ are invariant under gauge-transformations. Inserting the gauge transformation into the discretized field $f_{\boldsymbol{r}}$ and the definition of the link variables $A_{\boldsymbol{r},\mu}$ we see that these discretized fields transform as

$$\begin{aligned} f_{\boldsymbol{r}} &\to f_{\boldsymbol{r}}\, e^{i\phi_{\boldsymbol{r}}}, \\ A_{\boldsymbol{r},\mu} &\to A_{\boldsymbol{r},\mu} - \frac{1}{g}\Delta_\mu\phi_{\boldsymbol{r}}, \end{aligned} \tag{5.8}$$

where the field $\phi_{\boldsymbol{r}}$ is discretely defined on the same lattice points as $f_{\boldsymbol{r}}$. Inserting this into the discretization of the covariant derivative on the right hand side of Eq. (5.4) we see that indeed the right hand side transforms in the same way as the left, i.e. by picking up an overall factor $e^{i\phi_{\boldsymbol{r}}}$. This means that the discretized version of terms such as $|D_\mu f(\boldsymbol{r})|^2$, which were originally gauge-invariant, will remain invariant after discretization under the transformation in Eq. (5.8).

1. To show this, we see from Eq. (5.6) that $A_{\boldsymbol{r},\mu} \to aA_\mu(\boldsymbol{r})$, expand the exponential in $U_{\boldsymbol{r},\mu}$ to first order and insert on the right hand side of Eq. (5.4).

5.1.2 Reduction of symmetry

One word of caution in this connection is that the discretized gradient terms will not neccessarily have all the same spatial symmetries as the originating continuous terms because the bias of the forward direction in the forward difference and the cubic structure of the numerical lattice will in general break such symmetries. The effects of the cubic symmetry of the numerical lattice can be thought of as caused by implicit lattice potentials that increases in influence towards lower temperatures and higher field strengths when the model contains an external field . Such lattice potentials can e.g. cause topological defects to have preferred positions in discretizations of theories with translational symmetry.[2] As an example, consider again the discretization of the term $\int \mathrm{d}^3 r \ |\boldsymbol{D} f(\boldsymbol{r})|^2$ with discretized scalar field $f_{\boldsymbol{r}} = \rho_{\boldsymbol{r}} e^{i\theta_{\boldsymbol{r}}}$. The density term is rotationally symmetric[3] in 3D, however the discretized version can be written

$$2 \sum_{\boldsymbol{r}} \sum_{\mu} \rho_{\boldsymbol{r}}^2 [1 - \cos(\Delta_\mu \theta_{\boldsymbol{r}} + g A_{\boldsymbol{r},\mu})]. \tag{5.9}$$

From this form we can see that the term is only symmetric by rotation of 90° in the planes normal to the x, y and z directions. The discretized term contains cubic distortions when rotating in directions in-between these, hence the $SO(3)$ rotationall symmetry is broken down to the octahedral point group O.

The forward bias of the forward difference discretization scheme can also lead to breaking of symmetries that both the continuous model and the numerical lattice have in common when the scalar field consist of multiple components. Consider a density term of the form

$$\Re\Big[D_x \eta_x D_y \eta_y\Big], \tag{5.10}$$

where η_x and η_y are two scalar fields that transform as components of spin such that under a 90° $\hat{z}$-rotation (which is called a C_4 transformation), then $D_x \to D_y$, $D_y \to -D_x$, $\eta_x \to \eta_y$ and $\eta_y \to -\eta_x$. Inserting this into the continuous density term in Eq. (5.10) we see

2. More on this in Section 7.
3. By rotationally symmetric we mean that if we were to rotate the field configurations of $f(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$ in any direction, by any amount, the term would still yield the same value.

that it remains precicely the same i.e. invariant. Now consider the discretization of this term, which reads

$$\begin{aligned}
&\rho^x_{\boldsymbol{r}+\hat{x}}\rho^y_{\boldsymbol{r}+\hat{y}}\cos\left[\theta^x_{\boldsymbol{r}+\hat{x}}-\theta^y_{\boldsymbol{r}+\hat{y}}+g(A_{\boldsymbol{r},x}-A_{\boldsymbol{r},y})\right]\\
-&\rho^x_{\boldsymbol{r}+\hat{x}}\rho^y_{\boldsymbol{r}}\cos(\theta^x_{\boldsymbol{r}+\hat{x}}-\theta^y_{\boldsymbol{r}}+gA_{\boldsymbol{r},x})\\
-&\rho^x_{\boldsymbol{r}}\rho^y_{\boldsymbol{r}+\hat{y}}\cos(\theta^x_{\boldsymbol{r}}-\theta^y_{\boldsymbol{r}+\hat{y}}-gA_{\boldsymbol{r},y})\\
+&\rho^x_{\boldsymbol{r}}\rho^y_{\boldsymbol{r}}\cos(\theta^x_{\boldsymbol{r}}-\theta^y_{\boldsymbol{r}}),
\end{aligned} \tag{5.11}$$

where we have used the notation $\eta^a_{\boldsymbol{r}} = \rho^a_{\boldsymbol{r}}e^{i\theta^a_{\boldsymbol{r}}}$ for the discrete scalar fields. In terms of these scalar fields and link variables, a C_4 transformation consists of the mappings $\eta^x_{\boldsymbol{r}} \to \eta^y_{C_4\boldsymbol{r}}$, $\eta^y_{\boldsymbol{r}} \to -\eta^x_{C_4\boldsymbol{r}}$ such that e.g. $\rho^x_{\boldsymbol{r}+\hat{x}} \to \rho^y_{\boldsymbol{r}'+\hat{y}}$. The link-variables transform as $A_{\boldsymbol{r},\mu} \to A_{C_4\boldsymbol{r},C_4\mu}$ such that e.g. $A_{\boldsymbol{r},y} \to A_{\boldsymbol{r}',-x} = -A_{\boldsymbol{r}'-\hat{x},x}$. Using these transformations, and shifting the summation index of the external $\boldsymbol{r}$-sum, then the rotated discrete terms take the form

$$\begin{aligned}
&-\rho^x_{\boldsymbol{r}-\hat{x}}\rho^y_{\boldsymbol{r}+\hat{y}}\cos\left[\theta^x_{\boldsymbol{r}-\hat{x}}-\theta^y_{\boldsymbol{r}+\hat{y}}-g(A_{\boldsymbol{r},y}+A_{\boldsymbol{r}-\hat{x},x})\right]\\
+&\rho^x_{\boldsymbol{r}}\rho^y_{\boldsymbol{r}+\hat{y}}\cos(\theta^x_{\boldsymbol{r}}-\theta^y_{\boldsymbol{r}+\hat{y}}-gA_{\boldsymbol{r},y})\\
+&\rho^x_{\boldsymbol{r}-\hat{x}}\rho^y_{\boldsymbol{r}}\cos(\theta^x_{\boldsymbol{r}-\hat{x}}-\theta^y_{\boldsymbol{r}}-gA_{\boldsymbol{r}-\hat{x},x})\\
-&\rho^x_{\boldsymbol{r}}\rho^y_{\boldsymbol{r}}\cos(\theta^x_{\boldsymbol{r}}-\theta^y_{\boldsymbol{r}}),
\end{aligned} \tag{5.12}$$

which certainly is not the same as Eq. (5.11), i.e. the discretization of Eq. (5.10) is not invariant under a C_4 rotation. A more immediate way of seeing the problem is to recognize that the first term in Eq. (5.11) is a next-nearest neighbor coupling on the numeric lattice that only couples sites along one diagonal, but not the other as illustrated in Figure 5.1, thus rotational symmetry is broken by the discretization.

One remedy for this kind of problem is to re-establish the broken symmetry by an average over symmetry-transformed terms. In the case of the discretization of $\Re\left[D_x\eta_x D_y\eta_y\right]$ in Eq. (5.11), the C_4 symmetry can thus be re-established by taking the average of Eq. (5.11), Eq. (5.12), as well as the terms obtained by transforming the discretization in Eq. (5.11) by the rotations C_4^2 and C_4^3. Let $\mathcal{F}_{\boldsymbol{r}}$ denote the density terms in Eq. (5.11) and let $\mathcal{T}\mathcal{F}_{\boldsymbol{r}}$ be the terms that result when transforming $\mathcal{F}$ by a symmetry transformation $\mathcal{T}$. The average of symmetry-transformed terms that re-establishes the C_4 rotational symmetry can

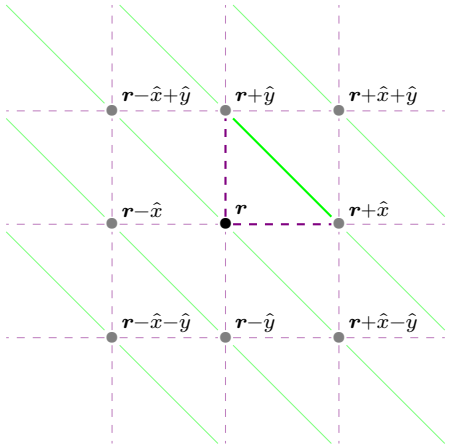

Figure 5.1: Couplings between sites on a single z-layer of the numerical lattice from the discretization in Eq. (5.11) of the term $\Re[D_x\eta_x D_y\eta_y]$. On-site terms are illustrated by a point ($\bullet$), while nearest neighbor and next-nearest neighbor couplings are illustrated by dashed and solid lines respectively. The couplings obtained by evaluating Eq. (5.11) at a point $\boldsymbol{r}$ are emphasized by being less transparent than the rest.

then be written

$$\begin{aligned}\mathcal{F}^{\text{sym}}_{\boldsymbol{r}} &= \frac{1}{4}\sum_{\mathcal{T}\in\{\mathbb{1},C_4,C_4^2,C_4^3\}} \mathcal{T}\mathcal{F}_{\boldsymbol{r}} \\ &= \frac{1}{4}\sum_{hh'=\pm 1} hh'\rho^x_{\boldsymbol{r}+h\hat{x}}\rho^y_{\boldsymbol{r}+h'\hat{y}}\cos\left[\theta^x_{\boldsymbol{r}+h\hat{x}} - \theta^y_{\boldsymbol{r}+h'\hat{y}} + g(A_{\boldsymbol{r},hx} - A_{\boldsymbol{r},h'y})\right].\end{aligned} \tag{5.13}$$

In this expression, the next-nearest neighbour couplings are along all diagonals around the point $\boldsymbol{r}$ such that it is rotationally symmetric under C_4 and thus does not explicitly break symmetries that both the original theory and the numerical lattice have in common. Such breaking of symmetries can, as we have shown above, result from the naive application of a forward difference discretization scheme[4] when discretizing terms with multiple gradient-directions and components. In

4. The symmetric expression in Eq. (5.13) can be more easily obtained by the use of a different discretization procedure than the discretization of the covariant derivative in Eq. (5.4). Taking the average of a forward and backward difference

models with such terms, the symmetry averaged expression can then be useful in diminishing the effect of meta-stable states and faster convergence when investigating such models by means of numerical computation. Finally we would like to stress that both versions yield the same theory in the continuum limit.

5.2 Including an external field

The interaction between superconductors and magnetic fields is an essential aspect in the study of superconductors and thus we will need to be able to add external magnetic fields to our models in order to study this interaction. An external field is usually thought of as a constant homogenous magnetic field in a certain direction that is a parameter of the problem rather than a variable. In other words we assume the magnetic flux to be the same everywhere and unchanging, and rather than ask what consequence the existance of a superconductor has on this field, we are interested in the effects the field has on the superconducting state. Physically this situation is relevant e.g. if a relatively small and thin sheet of superconducting material is placed in between two strong electromagnets as illustrated in Figure 5.2.

One way to introduce a constant magnetic field in a lattice model is simply to figure out what kind of vector potential $\boldsymbol{A}(\boldsymbol{r})$ would give a constant magnetic field $\boldsymbol{B}(\boldsymbol{r})$ through $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, and then set the link-variables $A_{\boldsymbol{r},\mu}$ accordingly with the only caviat being that for a lattice-model with periodic boundary conditions, then the factor $e^{igA_{\boldsymbol{r},\mu}}$ has to satisfy periodic boundary conditions as well. This implies the condition

$$\forall \nu \quad A_{\boldsymbol{r},\mu} = A_{\boldsymbol{r}+L_\nu \hat{\nu},\mu} + 2\pi m_\nu / g, \tag{5.14}$$

where $m_\nu \in \mathbb{Z}$ and ν gives a direction on the lattice.

---

that respects gauge-transformations we get the discretization mapping

$$D_\mu f(\boldsymbol{r}) \mapsto (e^{igA_{\boldsymbol{r},\mu}} f_{\boldsymbol{r}+\hat{\mu}} - e^{igA_{\boldsymbol{r},-\mu}} f_{\boldsymbol{r}-\hat{\mu}})/2.$$

Applying this symmetrized covariant discretization mapping to the density term in Eq. (5.10), yields Eq. (5.13).

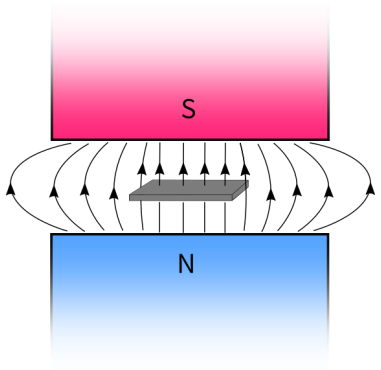

Figure 5.2: A thin sheet of superconducting material in the magnetic field produced by two magnets pointing in the same direction above and below the superconductor.

5.2.1 Landau Gauge

As an example, lets say we are interested in having an external field in the $\hat{z}$-direction with magnitude B. The vector-potential components A_x and A_y then have to satisfy the equation

$$\partial_x A_y - \partial_y A_x = B. \tag{5.15}$$

One configuration of the vector potential, which is called the *Landau gauge*, that satisfies this condition is $A_y = Bx$ with the other vector potential components set to zero. Inserting this into the definition of the link-variables in Eq. (5.6) yields $A_{\boldsymbol{r},\mu} = ar_x B\delta_{\mu,y}$. Here x is a continuous variable while r_x is the x-component of a lattice vector. Periodic boundary conditions on the lattice implies the condition $A_{\boldsymbol{r},y} = A_{\boldsymbol{r}+L_x\hat{x},y} - 2\pi m/g$ which finally restricts the value of the field B such that the link-variables in the Landau gauge must take the form

$$A_{\boldsymbol{r},\mu} = \delta_{\mu,y} r_x \frac{2\pi m}{gL_x}, \quad m \in \mathbb{Z}, \tag{5.16}$$

where $L_x = N_x a$ and N_x is the number of lattice sites in the x-direction. With this link-variable configuration, the field strength becomes $B = 2\pi f/ga^2$ where we have defined the filling fraction $f = m/N_x$ which in terms of vortices gives the number of magnetic vortices pr. plaquette of the numerical lattice.

5.2.2 Symmetric Landau gauge

The Landau gauge has the disadvantage that it singles out a direction in the xy-plane since the vector potential is set to $\boldsymbol{A}(\boldsymbol{r}) = Bx\hat{y}$ and thus only spatially dependent in the x-direction. It could because of this be argued to break a rotational symmetry of the model in the xy-plane given that Aharenov-Bohm-like effects are significant to the results. To mitigate any such concern, one can consider a symmetric gauge given by the choice $\boldsymbol{A}(\boldsymbol{r}) = -\boldsymbol{r} \times B\hat{z}/2$, which is rotationally symmetric in the xy-plane and, like the Landau gauge, produces the field $\boldsymbol{B} = B\hat{z}$. Inserting this choice of vector potential into the link variables yields, using implicit summation over repeated indices, $A_{\boldsymbol{r},\mu} = \epsilon_{\mu z\nu} r_\nu aB/2$. Periodic boundary conditions in this case implies two restrictions on the field value B because the vector potential varies in both the x- and y-direction. Implementing these conditions we can write the link-variables as

$$A_{\boldsymbol{r},\mu} = \epsilon_{\mu z\nu} r_\nu \frac{2\pi m}{gL_x}, \tag{5.17}$$

where m is a number $m \in \mathbb{Z}$ chosen such that there exists some $n \in \mathbb{Z}$ such that $mN_y = nN_x$, i.e. m is some multiple of N_x/N_y. Then the field value is given by $B = 2\pi f/ga^2$ for filling fraction $f = 2m/N_x$.

This gauge is a specification of the more general extended Landau gauge [6, 7], which is borne purely out of the assumptions of a field $\boldsymbol{B} \| \hat{z}$ pluss periodic boundary conditions.

5.2.3 Fluctuating field

For a normal strongly type-II superconductor, the London penetration depth λ is much larger than the superconducting coherence length ξ. In this regime it is valid to neglect spatial fluctuations in the gauge field since any deviation around the extremal field configuration is strongly suppressed. This is called the frozen gauge approximation and makes the vector potential act only as a constraint on the value of the uniform magnetic induction given by one of the gauges presented in the above sections [6]. When the superconducting state consists of multiple components on the other hand, it becomes difficult to classify it simply in terms of type-I or type-II based solely on λ and ξ [8]. With multiple components, it becomes essential to fluctuate the gauge field,

because it mediates a significant indirect interaction between the components [9, 10].

Fluctuations of the gauge field imparts an energy cost on the system given in SI-units by the free energy[5]

$$F_A = \frac{1}{2\mu_0}(\nabla \times \boldsymbol{A})^2. \tag{5.18}$$

There are a couple of different ways of discretizing this energy for inclusion in a lattice model depending on whether one defines the link-variables compactly, i.e. $gA_{\boldsymbol{r},\mu} \in (-\pi, \pi)$, or noncompactly, i.e. $gA_{\boldsymbol{r},\mu} \in (-\infty, \infty)$. Both versions belong to the same universailty class and thus produce the same results in a renormalization group sense as long as the fluctuations are sufficiently small [4]. For noncompact link-variables, we simply replace the gradient with the lattice difference operator from Eq. (5.3) divided by the lattice spacing a, and the gauge-field components by their corresponding link-variables such that

$$\begin{aligned} \partial_\mu &\mapsto \Delta_\mu / a_\mu, \\ A_\mu(\boldsymbol{r}) &\mapsto A_{\boldsymbol{r},\mu}. \end{aligned} \tag{5.19}$$

The discretized free energy pr. lattice site then becomes

$$F_{A,\boldsymbol{r}} = \frac{(\boldsymbol{\Delta} \times \boldsymbol{A}_{\boldsymbol{r}})^2}{2\mu_0 a^2} = \frac{1}{2\mu_0 a^2} \sum_\mu (A^{\boxdot}_{\boldsymbol{r},\mu})^2, \tag{5.20}$$

where we have defined the link-variable plaquette-sum vector, with components given by

$$A^{\boxdot}_{\boldsymbol{r},\mu} = \epsilon_{\mu\alpha\beta}\Delta_\alpha A_{\boldsymbol{r},\beta} = \oint_{\boxdot_\mu} \mathrm{d}\boldsymbol{r}' \cdot \boldsymbol{A}(\boldsymbol{r}'). \tag{5.21}$$

In the line-integral on the right hand side, the curve $\boxdot_\mu$ is given by a plaquette[6] normal to the vector $\hat{\mu}$, starting at the lattice point at $\boldsymbol{r}$ and

5. One way of deriving said energy is to start with the sourceless Maxwell Lagrangian for a massless vector field $\mathcal{L}_M = -F^{\mu\nu}F_{\mu\nu}/4\mu_0$. In this relativistic notation $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, $A^0 = V$, $\partial_0 = \partial_t/c$ and we use the metric $g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Assuming time-independence and neglecting terms consisting only of Vsince they do not couple to the Higgs fields (e.g. the superconducting components) in minimal coupling, then the Lagrangian reduces to $\mathcal{L}_M \to -(\nabla \times \boldsymbol{A})^2/2\mu_0$ and the free energy in Eq. (5.18) results.
6. In this context a plaquette is a square given by 4 neighboring lattice points contained in some plane.

moving along the square following the right hand rule. The integration curve given by the plaquette $\boxdot_z$ is shown in Figure 5.3.

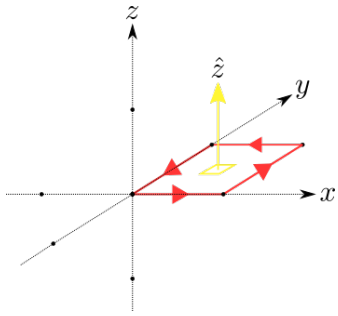

Figure 5.3: Integration path defined as $\boxdot_z$ along a plaquette of the numerical lattice in the xy-plane.

To impose an external field on a system with a fluctuating field we divide the link-variables into a fluctuating part $A^f_{\boldsymbol{r},\mu}$ with periodic boundary conditions and a constant part $A^0_{\boldsymbol{r},\mu}$ such that $A_{\boldsymbol{r},\mu} = A^f_{\boldsymbol{r},\mu} + A^0_{\boldsymbol{r},\mu}$. The field is then imposed by setting the constant part such that there is a net field induction through the system, e.g. by setting it to one of the gauges in Section 5.2.1 and 5.2.2.

Chapter 6

Monte-Carlo Techniques

In this chapter we discuss some techniques useful in Monte-Carlo (MC) simulations of systems in statistical-physics. In such systems these techniques wil be used to calculate thermal averages using random numbers. Let Z denote the partition function and $\mathcal{H}$ the Hamiltonian of the system. Then the thermal average of an observable $\mathcal{O}$ is defined as

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \sum_{\psi} \mathcal{O}(\psi) e^{-\beta \mathcal{H}(\psi)}, \tag{6.1}$$

where ψ denotes states of the system and we thus sum over all possible states. In the case of a quantum system, then this sum turns into a multi-dimensional integral over quantum coherent states. Now any attempt at estimating these integrals through an interpolation scheme is destined to fail because if we divide a 1-dimensional integral into M pieces and the error of the interpolation scheme scales as $\sim M^{-\kappa}$, then applied to a d-dimensional integral, its error will scale as $M^{-\kappa/d}$. What MC techniques then provides is a way of using random numbers in calculating Eq. (6.1) without actually summing over all the states. We do this by drawing random states ψ_i from a carefully selected probability distribution and using statistics to estimate how close the resulting thermal average is likely to be to the true thermal average. Letting M be the number of samples, then the error scales as $M^{-1/2}$ and is independent of the number of dimensions of the integral.

As in the case of the stationary phase approximation, the calculation of the sum in Eq. (6.1) can be made much more effective by considering which terms give large contributions. If we have a probability distribution $\pi(\psi)$ of sampled states ψ that has peaked at states that gives large contributions to $\langle\mathcal{O}\rangle$, then our estimate will converge much quicker to the true value than if we were to sample states uniformely. In a sense we are interested in sampling only the important states, and hence this is called importance sampling. Let $\{\tilde{\psi}_i\}$ be a set of states that are uniformly sampled, while $\{\psi_i\}$ are sampled with probability distribution $\pi(\psi_i)$. The statistical estimator $\langle\bar{\mathcal{O}}\rangle$ of the thermal average of the observable $\mathcal{O}$ is then

$$\langle\bar{\mathcal{O}}\rangle = \sum_i \mathcal{O}(\tilde{\psi}_i)\frac{e^{-\beta\mathcal{H}(\tilde{\psi}_i)}}{Z(\{\tilde{\psi}_i\})} = \sum_i \mathcal{O}(\psi_i)\frac{e^{-\beta\mathcal{H}(\psi_i)}}{\pi(\psi_i)Z(\{\psi_i\})}. \tag{6.2}$$

Now assuming that the state-dependence of the observaible is less important than the exponential, then the largest contributions to the sum will come from states that are such that $e^{-\beta\mathcal{H}}/Z$ is large. We thus want to pick states such that

$$\pi(\psi_i) = e^{-\beta\mathcal{H}(\psi_i)}/Z. \tag{6.3}$$

Then, given M states sampled according to this probability distribution, the statistical estimator reduces to the arithmetic average

$$\langle\bar{\mathcal{O}}\rangle = \frac{1}{M}\sum_i \mathcal{O}(\psi_i). \tag{6.4}$$

6.1 Markov-Chain Monte-Carlo method

Markov-Chain Monte-Carlo (MCMC) is a strategy of obtaining a sample of random states ψ_k where the the states are drawn sequentially in such a way that the probability $P_k(\psi)$ of drawing a new state $\psi_k = \psi$ is only dependent on what the last state ψ_{k-1} was. The chain developed by drawing states in this way thus has no memory of the rest of the content of the chain, except for its last link ψ_{k-1}. A chain with this property is called a Markov-Chain, hence the name.

We want the sampled states to be drawn according to the probability distribution $\pi(\psi_k)$ discussed above. This is assured with the criteria of ergodicity and detailed balance. Ergodicity means in this context that

the states are drawn in such a way that if we were to draw infinitely many states, then we would have drawn all possible states ψ in the original sum in Eq. (6.1).

The criterion of detailed balance comes from the idea that we want the probability that a certain state is drawn to be independent of when the state is drawn in the chain. Let $P_k(\psi)$ be the probability that ψ is drawn at the kth point in the chain. Because of the Markov-chain property, this probability is fully determined by the probability that the previous state in the chain transitions into the state ψ. Let $\mathcal{T}(\psi' \to \psi)$ denote the probability that state ψ' transitions into state ψ, i.e. that the state ψ is drawn given a previously drawn state ψ'. Then the probability that the state drawn at the point $k+1$ in the chain is ψ, is given by

$$\begin{aligned} P_{k+1}(\psi) &= \sum_{\psi'} P(\psi_k = \psi' \ \wedge \ \psi' \text{ transitions to } \psi) \\ &= \sum_{\psi'} P_k(\psi')\mathcal{T}(\psi' \to \psi) \\ &= P_k(\psi) + \sum_{\psi'} \Big[P_k(\psi')\mathcal{T}(\psi' \to \psi) - P_k(\psi)\mathcal{T}(\psi \to \psi') \Big], \end{aligned} \tag{6.5}$$

where we have used that $\sum_{\psi'} \mathcal{T}(\psi \to \psi') = 1$ since the state must transition to some state. Now since the probability should be invariant of the point in the chain and be given by our desired probability density, we demand $P_{k+1}(\psi) = P_k(\psi) = \pi(\psi)$ such that the sum vanishes. Because the probability density $\mathcal{T}$ is arbitrary, the sum needs to vanish term-wise, yielding the final condition of detailed balance:

$$\pi(\psi')\mathcal{T}(\psi' \to \psi) = \pi(\psi)\mathcal{T}(\psi \to \psi'). \tag{6.6}$$

This states that for the process to be invariant of the point in the chain, it has be reversible.

6.2 Metropolis-Hastings method

The Metropolis-Hastings (MH) method is an algorithm for drawing states in a Markov-Chain that spesifies a transition probability $\mathcal{T}(\psi \to \psi')$ between states ψ and ψ' that satisfies the detailed balance criterion. The algorithm proceeds as follows:

1. Given a state ψ_k, pick a new state ψ_p where the process of picking this state has an, as of now, arbitrary probability distribution denoted $q(\psi_p \,|\, \psi_k)$ with the only requirement being that it leads to ergodic selection.

2. Accept this new proposed state ψ_p, with the probability $\alpha(\psi_p \,|\, \psi_k)$, defined as

$$\alpha(\psi_p \,|\, \psi_k) = \min\left\{1,\ \frac{\pi(\psi_p)q(\psi_k \,|\, \psi_p)}{\pi(\psi_k)q(\psi_p \,|\, \psi_k)}\right\}. \tag{6.7}$$

3. If ψ_p is accepted we set $\psi_{k+1} = \psi_p$. If not, then $\psi_{k+1} = \psi_k$. Finally return to 1. to pick the next state in the chain.

By this procedure, then the probability of transitioning between a state ψ at point k to a state ψ' at point $k+1$ is given by probability that the state ψ' is picked and that ψ' is accepted such that

$$\mathcal{T}(\psi \to \psi') = \alpha(\psi' \,|\, \psi)\, q(\psi' \,|\, \psi). \tag{6.8}$$

This transition probability satisfies detailed balance since inserting Eq. (6.8) and (6.7) yields

$$\begin{aligned}\pi(\psi')\mathcal{T}(\psi' \to \psi) &= \min\{\pi(\psi)q(\psi' \,|\, \psi),\ \pi(\psi')q(\psi \,|\, \psi')\} \\ &= \pi(\psi)\mathcal{T}(\psi \to \psi').\end{aligned} \tag{6.9}$$

6.2.1 Practical considerations

Usually, the above is a bit too general for practical implementation since we would have to calculate, or know, the probability distribution $q(\psi' \,|\, \psi)$ used in picking new proposed states. If we assume q to be symmetric such that

$$q(\psi' \,|\, \psi) = q(\psi \,|\, \psi'), \tag{6.10}$$

then we don't need to calculate it explicitely since it cancels out of the equation for α in Eq. (6.8).

A further simplification can be acheived by inserting the expression for $\pi(\psi)$ in Eq. (6.3) into the q symmetric version of α, which in this case reduces to

$$\alpha(\psi' \,|\, \psi) = \min\left\{1,\ e^{-\beta\left[\mathcal{H}(\psi') - \mathcal{H}(\psi)\right]}\right\}. \tag{6.11}$$

The significance of this form is that we see the transition probability is only dependent on the difference between the energy of the updated state and the original state. If the state of the system ψ is a collection of site-dependent sub-states $\phi(\boldsymbol{r}_j)$, e.g. how the state of an Ising chain is given by a collection of site-dependent spins, then the calculation of $\mathcal{H}(\psi)$ has to take into account all the sites. If we update only a single site $\boldsymbol{r}_j$ of ψ to get ψ', which we call a local MC update, then all the sites that do not have an interaction with $\boldsymbol{r}_j$ cancels out in the difference $\mathcal{H}(\psi') - \mathcal{H}(\psi)$. Then we only need to calculate the the difference in the sub-states that are affected by $\boldsymbol{r}_j$ to calculate the energy-difference. This is an essential property to have when creating a parallelized version of this algorithm since different parts of the lattice of sites then can be updated in an asynchronuous manner without affecting each other. In other words: by simplifying to the energy difference, the update scheme becomes local which makes local MC updates grid-parallelizable.

To use pseudo-random numbers to accept a new state ψ' with probability α we pick a uniformely distributed number $r \in (0, 1]$. Then we use the fact that

$$P[r \leq \alpha(\psi' \,|\, \psi)] = \alpha(\psi' \,|\, \psi), \tag{6.12}$$

so that updating the state if $r \leq \alpha$ is equivalent to updating the state with probability α. Given the form of α in Eq. (6.11) then

$$r \leq \alpha(\psi' \,|\, \psi) \;\Leftrightarrow\; \ln r \leq -\beta[\mathcal{H}(\psi') - \mathcal{H}(\psi)]. \tag{6.13}$$

To update the state with probability α we thus simply take the natural logarithm of r and update the state if the right hand side of Eq. (6.13) is true.

To obtain good statistics, we want, as a rule of thumb, the acceptance rate to be about $30-60\%$ for high temperature states.[1] The acceptance rate is defined as the number of proposed states ψ' that are accepted, divided by the total number of proposed states, and will in general be proportional to the transition probability $\alpha(\psi' \,|\, \alpha)$. This can be adjusted by changing the way new states ψ' are proposed. Let ψ be composed of site specific substates $\phi(\boldsymbol{r}_j)$ and let a state ψ' be proposed by

1. High temperature states refers to states that are well above any transition temperature of the system.

changing the values of the sub-state $\phi(\boldsymbol{r}_0)$. Choosing values closer to the original sub-state $\phi(\boldsymbol{r}_o)$, the difference $\mathcal{H}(\psi') - \mathcal{H}(\psi)$ decreases such that $\alpha(\psi' \mid \psi)$ in Eq. (6.11) approaches 1 and the acceptance rate increases.

Proposing states such that the acceptance rate is very high, in this case means that the states do not change very much with each MC update. This can lead to freezing of the simulation, where the measurements do not change even after a significant number of MC updates because a large number of MC updates in a certain direction is needed to significantly change the measurements. On the other hand, too low of a acceptance rate will also freeze the simulation since then obviously states are very unlikely to change, leading to the same measurements repeatedly. Ultimately, whether the acceptance rate should be considered too high or too low, should be guided by the physics of the system since in the case that the system has reached a gobal minimum in the energy-lanscape and has low temperature, the proper statistics is obtained by an update scheme that gives a low acceptance rate. It is not advisable to change the acceptance rate during a measurement run over decreasing temperatures as this has tended to freeze the measurements at varying temperatures leading to confusion when trying to find a transition point.

6.3 Thermalization procedures

Thermalization in a MCMC simulation refers to the process of discarding a number of MC updates before starting to measure the states in the Markov-chain. The reason for doing this is because the first states in the chain will usually be very unlikely states in the ensamble of states, and thus give these states an artificially high statistical weight, unless we measure long enough. That time could be very long indeed if the starting states are very unlikely, thus to get measurements in a reasonable time, these unlikely starting states are discarded.

How many states to discard is usually estimated with the help of an energy vs. Monte-Carlo sweep (MCS)[2] plot as shown in Figure 6.1. Since the initial state usually has a different energy than the average

2. A Monte-Carlo sweep is a term used for attempting to update all the different sites of a system once.

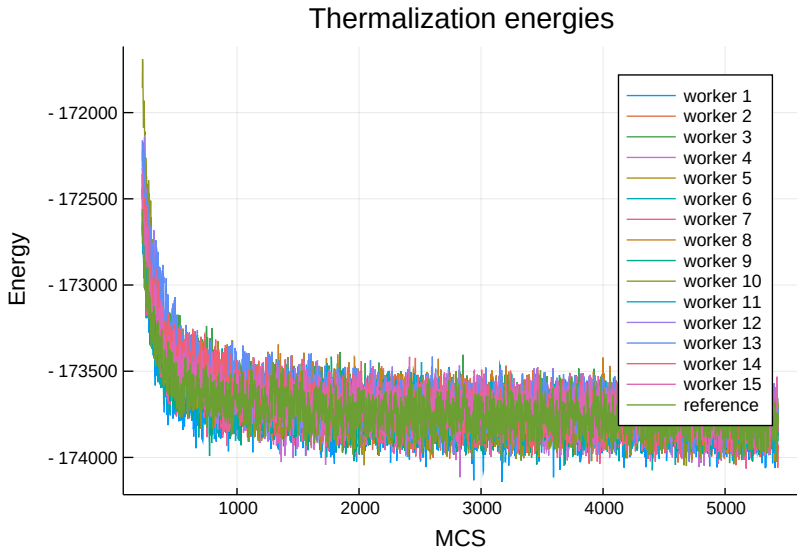

Figure 6.1: Thermalization of a 64^3 single-component XY-system from random initial states to a numerical temperature $T = 0.5$. The different curves represent the energy of different realizations of the same system initialized at different random states over several MCSs. The random initial states have a relatively high energy that stabilizes to the same value for all realizations in an exponential fashion.

energy in the Markov-chain, the energy can be seen to rapidly stabilize to the average value in such a plot.

Whether the energy stabilizes from above or below will depend on what the initial state is and what the temperature of the simulation is set to. For an ab-iniito state, in which the values of the sub-states are set to uniformly distributed random values in their validity range, this normally corresponds to a very high temperature and high energy state, hence thermalizing from such a state will have the energy stabilize from above. Another possibility for an ab-inito state is some kind of mean-field minimal solution of the Hamiltonian where sub-states at different sites are correlated. In this case the energy will usually be low and the thermalization energy will stabilize from below. This option has the disadvantage that if the mean field solution lies inside some local energy minimum then simulations based on it might not be able to get out and find the global miniomum, as opposed to states that are

thermalized from high energy where some might find this local minimum while others fall down in the global one. In general it is best to thermalize several independent systems from different initial conditions that yield quantitatively similar results to make it less likely that the results come from a local minimum.

A last suggestion for an initial state of the system, is the last state of a previous simulation. In this case the thermalization will stabilize depending on the relative temperature of the two simulations. To be sure that the measurements are not correlated with the measurements of the last simulation one should discard This is a useful practice if gathering results over an extended temperature range where the systems need a large thermalization time in order to stabilize. One would then typically start measuring at a high temperature and then decrease the temperature sucessively in steps with a thermalization period and measure period for each step.

For systems prone to fall into local minima it was found that a more careful thermalization period analogous to the measurement procedure described above, decreased the probability of freezing into such minima. Instead of thermalizing from a high energy - high temperature state directly down to the desired temperature, which we call quenching, a cooldown period was added. During the cooldown period, the temperature was lowered stepwise from a high temperature T_0 to a target temperature T with intermediate temperatures

$$T_k = \left(\frac{T}{T_0}\right)^{\frac{k}{N}} T_0. \tag{6.14}$$

The intermediate temperatures were geometrically distributed to ensure highest density of intermediate steps at lower temperatures. At each temperature step, a fixed number of MCSs were preformed such that more MCSs were done towards the lower temperature than higher. This was done because the simulations in general took longer to thermalize when the temperature decreased. An example of how the energy changed during such a thermalization period is shown in Figure 6.2.

The cooldown period was then followed by a conventional thermalization stage where the temperature was held constant at T. In most cases the energy had already stabilized at this point such that the en-

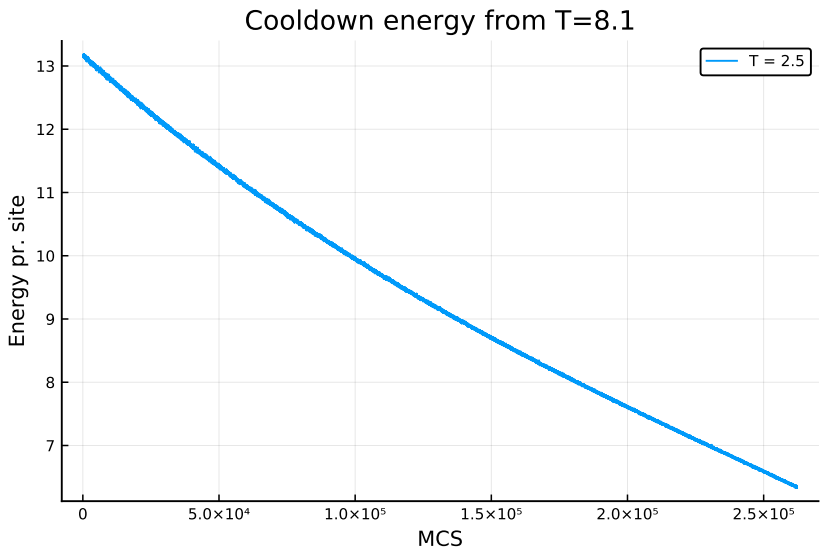

Figure 6.2: Energy pr. site of a 64^3 site model system of a $p + ip$ chiral superconductor during the cooldown stage. The temperature is lowered as a geometric sequence and a fixed number of MCSs are done at each temperature step. Comparing with the thermalization in Figure 6.1 we see that the cooldown period gives a significantly more gradual thermalization.

ergy measurements during this extra thermalization stage typically only showed fluctuations around the mean.

6.4 Parallel tempering

Parallel tempering is a method of simulating multiple systems over a range of different temperatures where the systems can exchange positions with their neighbors in this temperature range according to a MH-like update step. Since all the different systems all have the same parameters except for temperature, when viewed from the perspective of a single temperature, this leads to a normal Metropolis-Hastings MCMC simulation with an occational global update of all sites of the system whenever the current system exchanges with the system at a neighboring temperature. From the dual perspective of a single system, parallel tempering (PT) allows the system to move in temperature space.

This global updating, or movement in temperature space, has the advantage that it can prevent systems from getting stuck in local minima by allowing them to move to a higher temperature where it is easier to fluctuate to a more favorable configuration. In systems that has a jagged energy-landscape with lots of local minima this can be of great benefit and can reduce the required time it takes to measure observables with a certain accuracy by several orders of magnitude [11].

To implement parallel tempering MCMC in a temperature centric perspective, let $\{T_i\}_{i=1}^M$ be a sorted list of M ascending temperatures, and let $\{\lambda_i\}_{i=1}^M$ be a list of indices λ, that refer to replica states $\{\psi_\lambda\}_{\lambda=1}^M$ of the system such that the replica with temperature T_i is given by ψ_{λ_i} and its energy is given by E_{λ_i}. The simulation then proceeds according to the algorithm

1. Preform Δt normal Monte-Carlo updates on all replica states, e.g. using the MH method.

2. For each replica state ψ_λ calculate the corresponding energy E_λ.

3. For each pair of neighboring temperatures T_i and T_{i+1} where $T_i < T_{i+1}$:

   a) calculate the quantity

   $$\Delta = (E_{\lambda_{i+1}} - E_{\lambda_i})\Big(\frac{1}{T_{i+1}} - \frac{1}{T_i}\Big). \tag{6.15}$$

   b) Then swap the indices $\lambda_i \leftrightarrow \lambda_{i+1}$ with probability $\min\{1, e^\Delta\}$. This can as in the MH method be done by generating a random number $r \in [0, 1)$ and then swapping indices if $\ln r \leq \Delta$.

4. If the replicas ψ_λ have internal knowledge of their temperatures, then distribute T_i to ψ_{λ_i}, for all temperatures T_i.

5. Sample observables and return to 1.

This algorithm is very efficiently parallelizable since the bulk of computing time will be going to doing the Δt MC updates which can be performed in parallel by having each replica state ψ_λ be assigned to a separate thread / processor. If each thread in addition keeps track of

the replicas energy at the end of the MC updates, the only information that needs to be transferred between worker processes and the process doing the PT update step is the values of the energies to the PT process, and afterwards: the set of new temperatures back to the worker processes. The PT process itself only has to calculate $M-1$ values and move around the indices in an array.

For the parallel tempering method to generate good statistics efficiently, some care should be taken in the distribution of the temperatures T_i. A rule of thumb is to distribute them geometrically, i.e. according to

$$T_i = \left(\frac{T_M}{T_1}\right)^{\frac{k-1}{M-1}} T_1, \tag{6.16}$$

with the argument that lower temperatures generally have a lower relaxation rate. With geometric distribution, the temperatures are denser towards the low end such that the accpetance rate of swaps of replicas at neighboring temperatures would in general become more flat and independent of temperature. Should the specific heat diverge at some point in the temperature range as in the case of a phase transition, then this distribution would no longer be optimal since the acceptance probability of temperature swaps is inversely proportional to C_v and thus the accpetance rate would no longer be flat.

From the perspective of an individual replica, the overall goal with the distribution of temperatures is to maximize the number of the replica moves from the lowest temperature T_1, up to the highest temperature T_M and back to the lowest temperature again. This will then maximize the number of statistically independent visits of the system at each temperature. The hope then is that a flat accpetance rate with respect to the distributed temperatures will provide a good number of round-trips. With more advanced methods such as the feedback-optimized parallel-tempering Monte-Carlo method the number of round-trips can be optimized [12].

6.5 Grid parallelization

A simple way of utlizing multiple processor cores, or cpus on a multi-cpu system, is to run indipendent monte-carlo simulation on each processor. This is usually very efficient if a parameter of the system such as temperature is to be varied over some intervall. Then each simulation

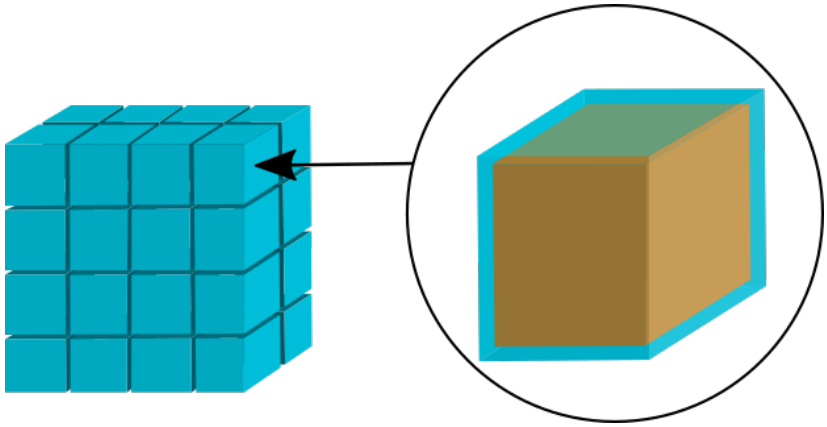

Figure 6.3: Illustration of the sub-division of the numerical grid into sub-cuboids. The right side illustrates that the sites in each sub-cuboid can be categorized as either internal sites, which can be updated asynchronuously on different sub-cuboids and is illustrated in the figure by the internal orange cube, or sites in a border region.

could have a different value of this parameter. In this case it is recommended to also implement parallel tempering since the extra overhead is minimal and the speedup of the simulations can be significant.

Alternatively, all the simulations can run with the same value on different processes and the samples from the individual simulations can then, be combined to a super-sample. This has the advantage that the individual sampling runs can be shortened and the implementation of the parallelization is straight-forward. A drawback with this method is however that separate simulations has to be thermalized individually, such that the more the super-sample is split on different processes, the more processor time is wasted on thermalization. Additionally, if the individual simulation runs depend on some sort of freezing like how vorticies freeze to the numeric lattice, then individual simulations could freeze at different angles such that they can not be combined to form good statistics.

A solution to the above issues with this simple parallelization is provided by grid-parallelization. This parallelization method is suited for simulations that consist of interacting sites either in 2D or 3D where local Monte-Carlo updates is to be preformed. We will focus on the

case of a 3D simulation where the sub-sites are organized in a numerical cubic lattice. The idea is to split the cube into different sub-cuboids as illustrated on the left side of Figure 6.3. Each sub-cuboids internal sites can then be updated in the normal Monte-Carlo fashion in parallell with the internal sites of other sub-cuboids. If the individual sites of the original system interacts with neighboring sites,[3] then it will be necessary for each sub-cuboid to have a border whose thickness depends on the range of the interaction. The sites in this border-region will then have an interaction with sites in the border-region of other sub-cuboids such that care has to be taken not to update a site based on the value of a neighboring site that is no longer valid. One solution to this is simply updating all the sites in the border-regions serially, i.e. update the border sites in a single sub-cuboid, communicate the updated values to the affected neighboring sub-cuboids, then move to the next sub-cuboid, etc. Or, it could be even more effective to parallelize the updates of the border sites as well by further sub-dividing the border-sites into different categories. As an example, one could define internal border-sites on the face of a sub-cuboid to be the border-sites that only depend on sites in the single neighboring sub-cuboid that is in the direction of the face-plane normal vector. Then all the right-facing internal border-sites can be done in parallel on all sub-cuboids, followed by all the top-facing internal border-sites, etc.

6.6 Reweighting

Reweigting techniques are methods of finding estimates of parameter-dependent observables at specific parameter values based on previously obtained samples of these observables from MC-simulations that have been done at parameter values independent or the reweighting parameters. Given a set of samples of some observable $\{o_i\}_{i=1}^{M}$ from previous MC simulation(s), these techniques provide a set of weights $\{w(\beta')_i\}_{i=1}^{M}$ that can be used to estimate the observable at a parameter value β' by the reweighting

$$\hat{o}(\beta') = \sum_i w(\beta')_i o_i. \tag{6.17}$$

3. In the case of a local MH update this can take the form of an energy-difference of the system at a proposed update-site that depends on the value of fields at neighboring sites.

This is very useful when estimating some observable over a temperature-range since then a single simulation can yield results not only for a single temperature, but for an extended region. If this region is close to a phase-transition then this can be used as a way of avoiding the critical slowing down of simulations at phase transitions by instead simulating at temperatures close to the phase-transition and then using reweighting techniques to estimate results at the phase-transition temperature. All this is possible because a simulation at a given temperature[4] produces a statistical distribution of energy values for the sampled states. This energy distribution will in general overlap with the energy-distribution produced when simulating at a temperature that is sufficiently close to the original. Because of this overlap, it is possible to statistically extrapolate the value of observables at the neighboring temperature.

Reweighting techniques are divided into single-histogram- and multi-histogram reweighting-techniques depending on whether they use statistical information from a single histogram or can combine histograms generated at multiple parameter values. The single-histogram techniques is used to estimate values of the observables at neighboring temperature-values and has the virtue of being comparatively simple to implement and understand. The multi-histogram techniques, on the other hand, has the advantage that the additional statistical information in general gives better estimates and is able to even give better estimates at the original temperatures the simulations were preformed at, but has a more involved implementation.

In our simulations we used a Julia implementation[5] of the multi-histogram technique originally developed by Bojesen, a technique which they used in [13], and updated for the current Julia release by us.

6.6.1 Ferrenberg-Swendsen single-histogram method

To derive the Ferrenberg-Swendsen single-histogram reweighting technique, let o_i be samples of an observable $\mathcal{O}$ done at sample states ψ_i of a system with Hamiltonian $\mathcal{H}$ at parameter value β, such that $\mathcal{O}(\psi_i) =$

4. It is possible to use rewighting techniques on other parameters of the simulation as well, as long as they are, like inverse-temperature, linear in the action.
5. This implementation is available at https://github.com/Sleort/FerrenbergSwendsenReweighting.jl/tree/1.0.3_update

o_i. Then, from out discussion of importance sampling, the estimate of the average of the observable when the states are sampled according to a probability distribution $\pi(\psi_i)$ is given by

$$\langle \hat{\mathcal{O}} \rangle_\beta = \frac{\sum_i o_i e^{-\beta \mathcal{H}(\psi_i)} / \pi(\psi_i)}{\sum_i e^{-\beta \mathcal{H}(\psi_i)} / \pi(\psi_i)}. \tag{6.18}$$

Using a simulation with importance sampling at parameter-value β to generate the sampled states, the probability distribution was

$$\pi(\psi_i) = \frac{e^{-\beta \mathcal{H}(\psi_i)}}{Z_\beta} = \frac{e^{-\beta \mathcal{H}(\psi_i)}}{\sum_\psi e^{-\beta \mathcal{H}(\psi)}}, \tag{6.19}$$

where the sum in the denominator is over all possible states ψ and not sampled states. Inserting $\pi(\psi_i)$ into $\langle \mathcal{O} \rangle_\beta$ in Eq. (6.18), this equation reduces to the arithmetric average, however if we now imagine wanting an estimate of $\langle \mathcal{O} \rangle_{\beta'}$ at an arbitrary parameter value β', then insertion yields

$$\langle \hat{\mathcal{O}} \rangle_{\beta'} = \frac{\sum_i o_i e^{-(\beta' - \beta)\mathcal{H}(\psi_i)}}{\sum_i e^{-(\beta' - \beta)\mathcal{H}(\psi_i)}}. \tag{6.20}$$

This expression then gives an estimate of the average of the observable $\mathcal{O}$ at an arbitrary parameter value β' using states ψ_i that were sampled at a specific parameter value β. This is called the Ferrenberg-Swendsen single-histogram reweighting technique [14], and in terms of the reweighting expression in Eq. (6.17) we can read off that the weights of this technique are given by

$$w(\beta')_i = \exp\left[-\ln\left(\sum_j e^{-(\beta' - \beta)[\mathcal{H}(\psi_j) - \mathcal{H}(\psi_i)]} \right) \right]. \tag{6.21}$$

Although simple, this techniques ability to extract information about observables around the simulated paramter-value makes it extremely valuable in for instance the study of scaling relations or to accurately calculate the peak of thermodynamic variables when the MC-simulations themselves take a significant amount of computing time.

6.6.2 Numeric evaluation of exponential sums

The reason for introducing the extra exponential in the form of $w(\beta')_i$ in Eq. (6.21) is because sums of exponential numbers generally are

hard to do numerically using finite-precision floating point numbers, however the logarithm of such a sum can be found using an iterative scheme. Let $S^{(k)}$ be a sum of k exponential numbers decreasing in magnitude that presumably are too large to be stored individually such that

$$S^{(k)} = e^{a_1} + e^{a_2} + ... + e^{a_k}, \tag{6.22}$$

with $a_{i+1} \leq a_i$, is numerically hard to do. Assuming however that fractions of the numbers can be stored, then we can numerically calculate

$$\ln S^{(2)} = a_1 + \ln\left(1 + e^{a_2 - a_1}\right). \tag{6.23}$$

Following the iteration

$$\ln S^{(k)} = \ln S^{(k-1)} + \ln\left(1 + e^{a_k - \ln S^{(k-1)}}\right), \tag{6.24}$$

then $\ln S^{(k)}$ can be found for arbitrary k without ever storing a single exponential number, only fractions of such numbers that are close to each other.

6.6.3 Multi-histogram Ferrenberg-Swendsen method

Let $\{\psi_i^k\}_{i=1}^{N_k}$ be sets of states sampled at the N_0 inverse temperatures β_k of a system with Hamiltonian $\mathcal{H}$. The energy of these states are then given by E_i^k and samples of an observable $\mathcal{O}$ at these states are given by $\mathcal{O}(\psi_i^k) = o_i^k$. The energy-samples can then be used to construct N_0 histograms

$$h_k(E) = \sum_{i=1}^{N_k} \delta_{E,E_i^k}, \tag{6.25}$$

giving the number of sampled states at a certain energy in the simulation with parameter value β_k. The goal is to use these histograms to estimate the density of states of the system which we for the purpose of the derivation of this method will define $n(E) = \sum_\psi \delta_{E,\mathcal{H}(\psi)}$. The essential steps in this derivation can be found in the original paper in Ref. [15], as well as Ref. [16] and [17]. With these definitions, the energetic probability distribution of the system at an inverse temperature β is given by

$$W(\beta, E) = n(E)e^{-\beta E}/Z_\beta, \tag{6.26}$$

where $Z_\beta = \sum_\psi e^{-\beta \mathcal{H}(\psi)}$ is the parition function. Based on the sampled histograms, $W(\beta_k, E)$ at temperature β_k can be estimated by $\hat{p}_k(E) = h_k(E)/N_k$, i.e. $\langle\ \hat{p}_k(E)\ \rangle = W(\beta_k, E)$. This implies that $\langle\ h_k(E)\ \rangle = N_k W(\beta_k, E)$ and assuming for now that the samples of states ψ_i^k and ψ_j^k are statistically independent, it can be shown that

$$\langle\ h_k(E)^2\ \rangle = N_k W(\beta_k, E)[1 + (N_k - 1)W(\beta_k, E)]. \tag{6.27}$$

Inserting these cumulants of the histograms into the variance we get

$$\delta^2 h_k(E) = \langle\ h_k(E)^2\ \rangle - \langle\ h_k(E)\ \rangle^2 = g_k N_k W(\beta_k, E), \tag{6.28}$$

by assuming $W(\beta_k, E) \ll 1$. The factor $g_k = 1 + 2\tau_k$, where τ_k is the autocorrelation time of the samples at β_k, is included to generalize the result to samples where ψ_i^k and ψ_j^k are not statistically independent.

By solving Eq. (6.26) w.r.t. $n(E)$ and inserting the estimator of $W(\beta_k, E)$, an estimator of the density of states is given by

$$\hat{n}_k(E) = \hat{p}_k(E) Z_{\beta_k} e^{\beta_k E}, \tag{6.29}$$

where Z_{β_k} is assumed known, an assumption we have to reconcile later. By the error propagation formula, then the variance of this estimator is given by

$$\delta^2 \hat{n}_k(E) = (Z_{\beta_k} e^{\beta_k E}/N_k)^2 \delta^2 h_k(E). \tag{6.30}$$

The estimator $\hat{n}_k(E)$ is an estimator of $n(E)$ using only a single histogram. We combine the estimators of single histograms using a weighted sum

$$\hat{n}(E) = \sum_k r_k \hat{n}_k(E), \tag{6.31}$$

where the coefficients r_k has to satisfy the condition $\sum_k r_k = 1$ for the expectation value of $\hat{n}(E)$ to give the density of states. The coefficients r_k are determined by minimizing the variance $\delta^2 \hat{n}_k(E)$ subject to the constraint $\sum_k r_k = 1$ using Lagrange multiplier which yields the estimator

$$\hat{n}(E) = \frac{\sum_{k=1}^{N_0} g_k^{-1} h_k(E)}{\sum_{l=1}^{N_0} N_l g_l^{-1} e^{-\beta_l E} Z_{\beta_l}^{-1}}. \tag{6.32}$$

The assumption that Z_{β_k} is known is now reconciled. Since we can write the partition function using the density of states through

$$Z_\beta = \sum_\psi e^{-\beta \mathcal{H}(\psi)} = \sum_E n(E) e^{-\beta E}, \tag{6.33}$$

then we use the density of states estimator to estimate the parition function, and use this estimate of the partition function $\hat{Z}_{\beta_k}$ in the density of states estimator. This then creates an implicit equation for $\hat{Z}_{\beta_k}$ that has to be solved self-consistently. Inserting the definition of $h_k(E)$ and exchanging sums to remove the histograms, this equation takes the form

$$\hat{Z}_\beta = \sum_{k=1}^{N_0} \sum_{i=1}^{N_k} \frac{g_k^{-1} e^{-\beta E_i^k}}{\sum_{l=1}^{N_0} N_l g_l^{-1} e^{-\beta_l E_i^k} \hat{Z}_{\beta_l}^{-1}}, \tag{6.34}$$

which gives N_0 equations for N_0 unknows $\hat{Z}_{\beta_k}$ when evaluated at the different $\beta = \beta_m$.

Solving Eq. (6.34) is usually done with the help of an iterative solution method for non-linear equations such as the Newton-Raphson method. To numerically calculate a solution, it is inconventient to work with the full quantities $\hat{Z}_{\beta_m}$ since these are usually extremely large. Instead it is sufficient to calculate the variables

$$L_m \equiv \ln \hat{Z}_{\beta_m} - \ln \hat{Z}_{\beta_1}, \tag{6.35}$$

since the weights in the reweighting of observables can be written in terms of them. Dividing Eq. (6.34) by Z_{β_1}, we get that the $N_0 - 1$ equations we need to solve self consistently for the $N_0 - 1$ variables L_m are given by

$$L_m = \ln \left\{ \sum_{k=1}^{N_0} \sum_{i=1}^{N_k} \frac{g_k^{-1} e^{-\beta_m E_i^k}}{\sum_{l=1}^{N_0} g_l^{-1} e^{-\beta_l E_i^k - L_l}} \right\}. \tag{6.36}$$

This form has the big advantage that the overall logarithm allows us to not have to calculate the exponential sums directly, but instead only calculate logarithms of these sums. For each sum $\sum_i e^{a_i}$ containing exponentials, which are potentially too large to be stored numerically, we simply re-exponentiate the entire sum to $\exp \ln \sum_i e^{a_i}$ and then

use the method outlined in Section 6.6.2 to calculate $\ln \sum_i e^{a_i}$. Because of the overall logarithm, the exponential drops out in the last re-exponentiation such that we never have to store a single exponential number.

After finding self-consistent values for the $N_0 - 1$ vairables L_m, the weights w_i^k for reweighting the observable $\mathcal{O}$ can be found. In terms of the density of states $n(E)$, the thermal average of the observable is written

$$\langle \mathcal{O} \rangle_\beta = \frac{\sum_E \mathcal{O}(E) n(E) e^{-\beta E}}{\sum_E n(E) e^{-\beta E}}. \tag{6.37}$$

Inserting the reweighting estimate of $\hat{n}(E)$ in Eq. (6.32) for $n(E)$ we get the reweighting estimate

$$\langle \hat{\mathcal{O}} \rangle_\beta = \frac{\hat{Z}_{\beta_1}}{\hat{Z}_\beta} \sum_{k=1}^{N_0} \sum_{i=1}^{N_k} \frac{o_i^k g_k^{-1} e^{-\beta E_i^k}}{\sum_{l=1}^{N_0} N_l g_l^{-1} e^{-\beta_l E_i^k - L_l}}, \tag{6.38}$$

where $\hat{Z}_{\beta_1} / \hat{Z}_\beta$ is given by

$$\frac{\hat{Z}_\beta}{\hat{Z}_{\beta_1}} = \sum_{k=1}^{N_0} \sum_{i=1}^{N_k} \frac{g_k^{-1} e^{-\beta E_i^k}}{\sum_{l=1}^{N_0} N_l g_l^{-1} e^{-\beta_l E_i^k - L_l}}, \tag{6.39}$$

through Eq. (6.34). The two equations for $\langle \hat{\mathcal{O}} \rangle_\beta$ in Eqs. (6.38) and (6.39) together with the self-consistency equation in Eq. (6.36), is sufficient to describe the multi-histogram method. Notice that in these equations the histograms on which the method was derived do not figure but have been replaced by the more fundamental energy samples. This form makes the method more convenient to implement for systems with continuous energy distributions since it removes the need for a sum over all possible energies.

When calculating the exponential sums in the weights implied by Eqs. (6.38) and (6.39), numerical overflow can be avoided by first using logarithms to calculate the logarithm of a set of related un-normalized weights as before, then subtracting the maximum logarithmic value for each weight such that each weight is $\lesssim 1$ and then using the sum of these weights to properly normalize in the end.

6.6.4 Initial guess

An iterative non-linear solver usually needs an initial guess at the solution. In the case of the multi-histogram method equations, a good initial guess can be provided by the single-histogram Ferrenberg-Swendsen method. Since only fractions of partition function values are needed we may set that $\hat{Z}_{\beta_1} = 1$ and use the Ferrenberg-Swendsen method based on the β_1 energies to estimate the value $\hat{Z}^0_{\beta_2}$ of $\hat{Z}_{\beta_2}$ at neighboring inverse-temperature β_2 by the formula

$$\hat{Z}^0_{\beta_2} = \sum_{i=1}^{N_1} e^{-(\beta_2-\beta_1)E_i^1}. \tag{6.40}$$

In terms of the numerically convenient variables L_m then this first guess L_2^0 takes the form

$$L_2^0 = \ln\Big[\frac{1}{N_1}\sum_{i=1}^{N_1} e^{-(\beta_2-\beta_1)E_i^1}\Big]. \tag{6.41}$$

Continuing to estimate the partition function $\hat{Z}_{\beta_m}$ through the single-histogram Ferrenberg-Swendsen method based on the data at β_{m-1}, then we may find all subsequent L_m^0 by applying the iteration scheme

$$L_m^0 = L_{m-1}^0 + \ln\Big[\frac{1}{N_{m-1}}\sum_{i=1}^{N_{m-1}} e^{-(\beta_m-\beta_{m-1})E_i^{m-1}}\Big]. \tag{6.42}$$

Chapter 7

Vortices in superconductors

In conventional type-I superconductors, the Meissner effect prevents any magnetic field from penetrating the superconductor when it is in the superconducting state. In a type-II superconductor, the transition between the normal- and superconducting state is more gradual than in the type-I case due to an intermediate transitional state where topological defects in the superconducting field becomes stable allowing quanta of magnetic field to pass through the material. The transitional value of the external field strength below which no magnetic field penetrates the superconductor is called B_{c1}. The upper transitional field strength above which the material stops being superconducting altogether is called B_{c2}. The state with regions of topological defects through which magnetic field quanta can penetrate, which are interspersed in a sea of superconducing state then exists between these values. It is important to note that the Meissner effect is still present in this transitional state - preventing magnetic field lines from penetrating the superconducting state, however at topological defects, the material switches to the normal state and thus allows magnetic field lines to penetrate at these points. The final continuous transition to the normal state at B_{c2} is then caused by the proliferation of vortex-loops, sending the whole material to the normal state.

The regions of normal state containing a topological defect of the superconducting state and through which magnetic field quanta can penetrate are known as superconducting vortices because they are sur-

rounded by a circulating superconducting current. This current is set up by the presence of the magnetic field and shilds the rest of the superconducting condensate from its influence.

Whether a superconductor is considered type-I or type-II is conventionally given by the magnetic field penetration length λ and the superconducting coherence length ξ which come together to form the Ginzburg-Landau (GL) parameter $\kappa = \lambda/\xi$. These parameters come out of the description of the superconducting state given by the GL theory of a single-component complex field minimally coupled to a gauge field. If $\kappa \gg 1$ then we say we have a strongly type-II superconductor, while is $\kappa \ll 1$ the superconductor is strongly type-I. The transitional value between type-I and type-II have a theoretical mean-field value of $\kappa = 1/\sqrt{2}$, however numerical calculations has given it the value $\kappa = (0.76 \pm 0.04)/\sqrt{2}$ all within the conventional GL formalism.

In a type-II conventional superconductor without any structural defects, as we increase the field strength, we introduce more vortices into the material in order to carry the required number of magnetic field quanta. At first these vortices behave like a liquid where they mutually repel eachother if they get close. As more vortices are introduced to the system, the inter-portex repulsion leads to them forming a two-dimensional lattice with equidistant lattice-spacing. Since the triagonal lattice is the lattice with the highest packing fraction, i.e. the lattice that have the highest density of sites at a given lattice spacing, the lattice formed will be triagonal. Such a triagonal (hexagonal) lattice of single quanta vortices is known as the Abrikosov lattice since it consists of single quantua vortices which are known as Abrikosov vortices.

7.1 Vorticity observables

A condensate described by a complex field ψ with phase θ can have topological defects given by discontinuities in the field θ due to its compact nature ($\theta \in [0, 2\pi)$). These topologicals defects then lead to singularities in the field $\nabla\theta$ which allows a nonzero value of $\nabla \times \nabla\theta$ at these points. Integrating over the surface S system with surface normal vector $\hat{s}$ and using Stokes theorem then yields

$$\int_S \mathrm{d}^2 r (\nabla \times \nabla\theta) \cdot \hat{s} = \oint_{\partial S} \nabla\theta \cdot \mathrm{d}\boldsymbol{r} = 2\pi N_v, \quad N_v \in \mathbb{Z}, \tag{7.1}$$

where ∂S is a path around the boundary of S traversed counter-clockwise. The last equality comes from the observation that ∂S is far away from the singularity such that $\nabla\theta$ is continuous along the path and N_v thus counts the number of times the vector $\nabla\theta$ rotates counter-clockwise back to its initial position. N_v can then be interpreted as the total vorticity of the field θ over the surface S. Since N_v is the total vorticity, then from Eq. (7.1) we see that

$$\boldsymbol{n}_v = \frac{\nabla \times \nabla\theta}{2\pi} \tag{7.2}$$

must be interpreted as a vector of local vorticity density.

If the system ψ and θ described contains a gauge field that is coupled to them, then any meaningful observable needs to be gauge-invariant. We clearly see that the expression in Eq. (7.3) is gauge-dependent by sending $\theta \to \theta + \phi$. To make a gauge-invariant observable under the gauge-transformation in Eq. (5.7), we see that we need to modify the definition to

$$\boldsymbol{n}_v = \frac{\nabla \times (\nabla\theta + g\boldsymbol{A})}{2\pi}. \tag{7.3}$$

This expression then defines $\boldsymbol{n}_v$ as a gauge invariant vector of local vorticity density of the compact field θ.

In lattice models we want to discretize the vorticity density in Eq. (7.3) in order to effectively calculate it in Monte-Carlo simulations of the lattice model. In such a discrete model we have to take care to re-compactify the quantity $\nabla\theta + g\boldsymbol{A}$ to only be defined on some interval of length 2π. Using the discretization mapping of ∂_μ and $A_\mu(\boldsymbol{r})$ from Eq. (5.19), we want $\Delta_\mu\theta + gA_{\boldsymbol{r},\mu} \in [-\pi, \pi)$. Defining the operator

$$\hat{C}_\pi\, x = \mathrm{mod}(x + \pi, 2\pi) - \pi, \tag{7.4}$$

the discretized vorticity density can be written

$$\begin{aligned} \boldsymbol{n}_{v,\boldsymbol{r}} &= \frac{\hat{e}_\mu \epsilon_{\mu\nu\lambda} \Delta_\nu \hat{C}_\pi(\Delta_\lambda\theta_{\boldsymbol{r}} + gA_{\boldsymbol{r},\lambda})}{2\pi a^2} \\ &= \frac{1}{2\pi a^2} \sum_\mu \hat{e}_\mu \sum_{\boxdot_\mu} \hat{C}_\pi(\Delta_\lambda\theta_{\boldsymbol{r}} + gA_{\boldsymbol{r},\lambda}). \end{aligned} \tag{7.5}$$

Implicit summation over repeated indices is used on the first line while on the second, the components of the vector $\boldsymbol{n}_{v,\boldsymbol{r}}$ are written as plaquette-sums, which are sums of direction dependent quantities along a path

$\boxdot_\mu$, which is described below Eq. (5.21) and illustrated in Figure 5.3. In the plaquette-sum the directional quantity is always chosen along the path and the path is traversed according to the right hand rule with normal vector $\hat{e}_\mu$.

7.2 Double quanta vorticies

7.3 Vortex lattices

7.4 Vortex lattice transitions

Chapter 8

Example chapter

This chapter serves as a simple demonstration of what the thesis template looks like. It provides some simple examples of typical content in academic theses; see e.g. table 8.1, figure 8.1, and equation (8.2). In addition, it shows what the default chapters, sections, and margins look like. For completeness, we also include some references [18–20].

8.1 Here is an example section

Here is an example of a display equation: a self-consistency equation taken from the study of superconductivity. Note that all equations are left-justified instead of centered. In text with a large number of short display equations, this makes the text easier to follow with your eyes, since they don't have to jump large distances at a time. It also makes the page look more organized due to the constant indentation level.

$$\Delta(z) = \int_0^\infty \mathrm{d}\epsilon \ \mathrm{Re}\big[f_s(\epsilon)\big] \tanh\left(\frac{\pi}{2e^\gamma}\frac{\epsilon/\Delta_0}{T/T_\mathrm{c}}\right) \tag{8.1}$$

Here is another example, in the form of Maxwell's equations. Note the consistent indentation level compared to the equation above.

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho, \quad \nabla \times \mathbf{E} = 0 - \partial_t \mathbf{B}; \\ \nabla \cdot \mathbf{B} &= 0, \quad \nabla \times \mathbf{H} = \mathbf{J} + \partial_t \mathbf{D}. \end{aligned} \tag{8.2}$$

Finally, we will show some examples of tables and figures. Note how the width of table 8.1 matches the indentation level of the equations.

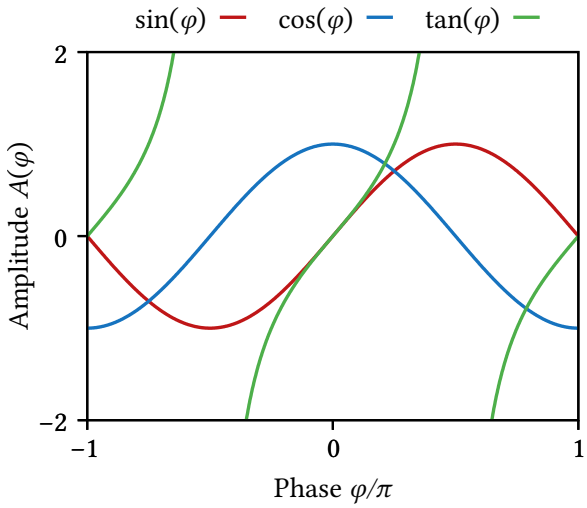

Figure 8.1: This is an example figure made by GNUPLOT. I have also included an intentionally long caption to show the margins.

The captions are formatted using a small font and extra margins, which helps separate the captions from the surrounding text.

8.2 Lorem ipsum

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra

Table 8.1: Test table with some mathematical constants.

Name	Symbol	Value
Euler constant	e	$2.71\ldots$
Circle constant	π	$3.14\ldots$
Imaginary identity	i	$\sqrt{-1}$

ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor.

Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Chapter 9

Conclusion

This would be a natural place to summarize your main results, and perhaps provide an outlook for future developments in the field.

Bibliography

1. **H.K. Onnes.**
   The disappearance of the resistivity of mercury.
   Commun. Phys. Lab. Univ. Leiden (May 1911).

2. **J.W. Negele, H. Orland.**
   Quantum Many-Particle Systems (1998).

3. **T. Inui, Y. Tanabe, Y. Onodera.**
   Group Theory and Its Applications in Physics (1990).
   DOI: 10.1007/978-3-642-80021-4

4. **A. Shimizu, H. Ozawa, I. Ichinose, T. Matsui.**
   Lattice Ginzburg-Landau model of a ferromagnetic p-wave pairing phase in superconducting materials and an inhomogeneous coexisting state.
   Phys. Rev. B 85, 144524 (14 Apr. 2012).
   DOI: 10.1103/physrevb.85.144524

5. **G. Münster, M. Walzl.**
   Lattice Gauge Theory - A short Primer.
   2000.
   arXiv: hep-lat/0012005 [hep-lat].

6. **A.K. Nguyen, A. Sudbø.**
   Topological phase fluctuations, amplitude fluctuations, and criticality in extreme type-II superconductors.
   Phys. Rev. B 60, 15307–15331 (22 Dec. 1999).
   DOI: 10.1103/physrevb.60.15307

7. **A.K. Nguyen, A. Sudbø.**
A new broken $U(1)$-symmetry in extreme type-II superconductors.
Eur. Phys. Lett. 46, 780–786 (6 June 1999).
DOI: https://doi.org/10.1209/epl/i1999-00332-7

8. **E. Babaev, M. Speight.**
Semi-Meissner state and neither type-I nor type-II superconductivity in multicomponent superconductors.
Phys. Rev. B 72, 180502 (18 Nov. 2005).
DOI: 10.1103/physrevb.72.180502

9. **E. Smørgrav, J. Smiseth, E. Babaev, A. Sudbø.**
Vortex Sublattice Melting in a Two-Component Superconductor.
Phys. Rev. Lett. 94, 096401 (9 Mar. 2005).
DOI: 10.1103/physrevlett.94.096401

10. **J. Smiseth, E. Smørgrav, A. Sudbø.**
Critical Properties of the N-Color London Model.
Phys. Rev. Lett. 93, 077002 (7 Aug. 2004).
DOI: 10.1103/physrevlett.93.077002

11. **H.G. Katzgraber.**
Introduction to Monte Carlo Methods.
2009.
eprint: arXiv:0905.1629.

12. **H.G. Katzgraber, S. Trebst, D.A. Huse, M. Troyer.**
Feedback-optimized parallel tempering Monte Carlo.
Journal of Statistical Mechanics: Theory and Experiment 2006, P03018–P03018 (Mar. 2006).
DOI: 10.1088/1742-5468/2006/03/p03018

13. **T.A. Bojesen.**
Multihistogram reweighting for nonequilibrium Markov processes using sequential importance sampling methods.
Phys. Rev. E 87, 045302 (4 Apr. 2013).
DOI: 10.1103/physreve.87.045302

14. **A.M. Ferrenberg, R.H. Swendsen.**
New Monte Carlo technique for studying phase transitions.
Phys. Rev. Lett. 61, 2635–2638 (23 Dec. 1988).
DOI: 10.1103/physrevlett.61.2635

15. **A.M. Ferrenberg, R.H. Swendsen.**
Optimized Monte Carlo data analysis.
Phys. Rev. Lett. 63, 1195–1198 (12 Sept. 1989).
DOI: 10.1103/physrevlett.63.1195

16. **M. Newman, G. Barkema.**
Monte Carlo Methods in Statistical Physics (1999).
ISBN: 9780198517979

17. **K. Rummukainen.**
Monte Carlo simulation methods.
University of Helsinki lecture notes.
eprint: https://www.mv.helsinki.fi/home/rummukai/lectures/montecarlo_oulu/.

18. **R. Feynman, R. Leighton, M. Sands.**
The Feynman Lectures on Physics 2nd edition, 36–38 (1963).
ISBN: 9780201021165

19. **M. Lipovaca.**
Learn You a Haskell for Great Good! 1st ed. (2011).
ISBN: 9781593272838

20. **J. Beringer *et al.***
Review of Particle Physics.
Physical Review D 86, 10–20 (2012).
DOI: 10.1103/physrevd.86.010001